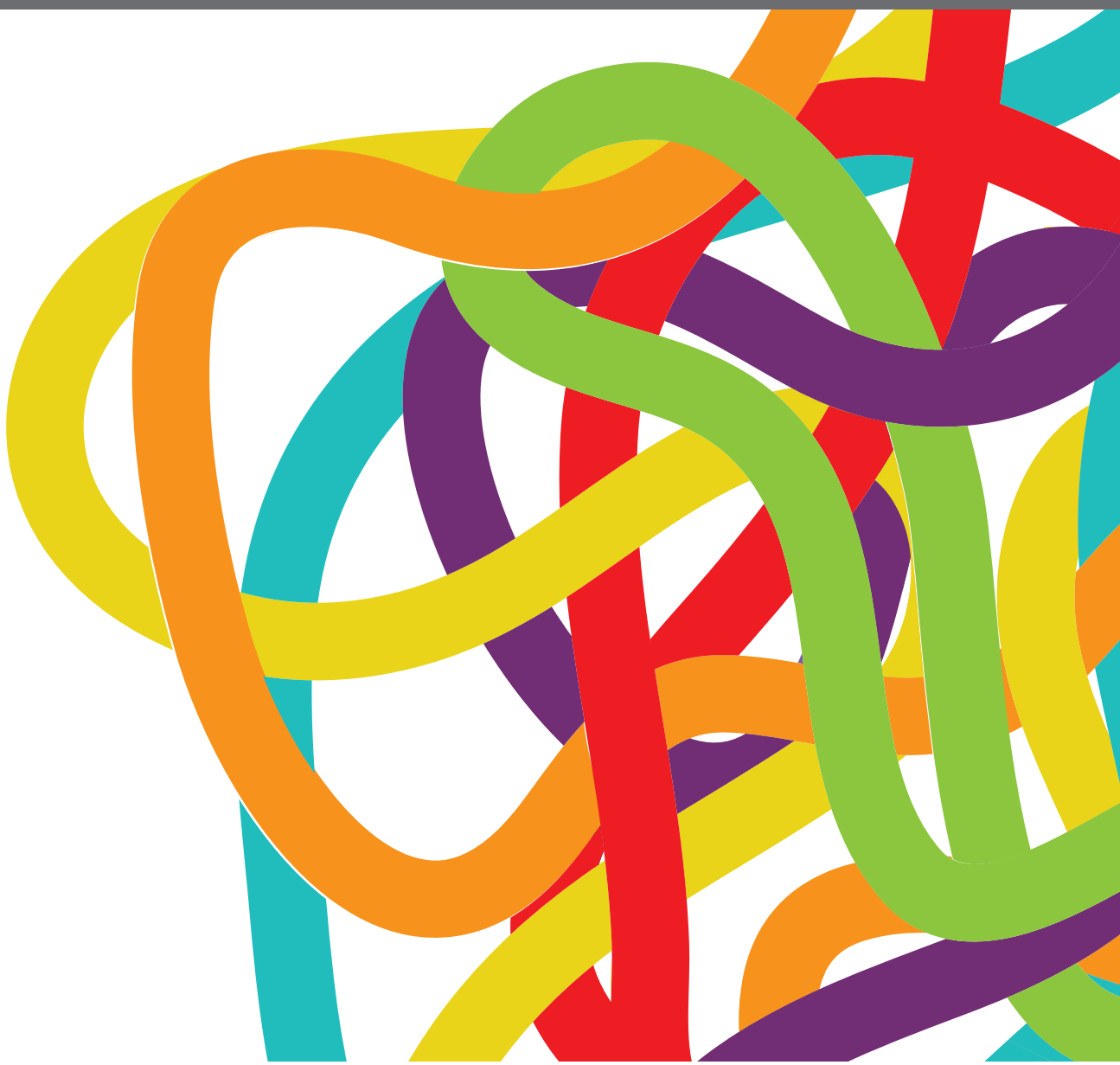
ADVANCES AND NEW INSIGHTS INTO CANCER CHARACTERIZATION: WHEN NOVEL IMAGING MEETS QUANTITATIVE IMAGING BIOMARKERS

EDITED BY: Guang Yang, Kyung Hyun Sung and Tingying Peng
PUBLISHED IN: Frontiers in Oncology



frontiers Research Topics

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-83250-751-3
DOI 10.3389/978-2-83250-751-3

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

ADVANCES AND NEW INSIGHTS INTO CANCER CHARACTERIZATION: WHEN NOVEL IMAGING MEETS QUANTITATIVE IMAGING BIOMARKERS

Topic Editors:
Guang Yang, Imperial College London, United Kingdom
Kyung Hyun Sung, Department of Radiology, UCLA Health System, United States
Tingying Peng, Helmholtz AI, Helmholtz zentrum München, Germany

Citation: Yang, G., Sung, K. H., Peng, T., eds. (2022). Advances and New Insights Into Cancer Characterization: When Novel Imaging Meets Quantitative Imaging Biomarkers. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83250-751-3

Table of Contents

06 ***Comparison of PET/CT and MRI in the Diagnosis of Bone Metastasis in Prostate Cancer Patients: A Network Analysis of Diagnostic Studies***
Fanxiao Liu, Jinlei Dong, Yelong Shen, Canhua Yun, Ruixiao Wang, Ganggang Wang, Jiyang Tan, Tao Wang, Qun Yao, Bomin Wang, Lianxin Li, Jingyi Mi, Dongsheng Zhou and Fei Xiong

23 ***Novel Non-Invasive Radiomic Signature on CT Scans Predicts Response to Platinum-Based Chemotherapy and Is Prognostic of Overall Survival in Small Cell Lung Cancer***
Prantesh Jain, Mohammadhadi Khorrami, Amit Gupta, Prabhakar Rajiah, Kaustav Bera, Vidya Sankar Viswanathan, Pingfu Fu, Afshin Dowlati and Anant Madabhushi

35 ***^{18}F-PSMA-1007 PET/CT Performance on Risk Stratification Discrimination and Distant Metastases Prediction in Newly Diagnosed Prostate Cancer***
Zhuonan Wang, Anqi Zheng, Yunxuan Li, Weixuan Dong, Xiang Liu, Wang Yuan, Fan Gao and Xiaoyi Duan

45 ***Visualization of microRNA-21 Dynamics in Neuroblastoma Using Magnetic Resonance Imaging Based on a microRNA-21-Responsive Reporter Gene***
Guangcheng Bao, Jun Sun, Helin Zheng, Jingxin Hou, Jie Huang, Jie Wei, Yuanqiao Fu, Jiawen Qiu, Xuefeng Zou, Bin Xiang and Jinhua Cai

57 ***Multimodal Imaging Technology Effectively Monitors HER2 Expression in Tumors Using Trastuzumab-Coupled Organic Nanoparticles in Patient-Derived Xenograft Mice Models***
Li Wen, Lei Xia, Xiaoyi Guo, Hai-Feng Huang, Feng Wang, Xian-teng Yang, Zhi Yang and Hua Zhu

67 ***Radiomics Analysis of Contrast-Enhanced CT for the Preoperative Prediction of Microvascular Invasion in Mass-Forming Intrahepatic Cholangiocarcinoma***
Fei Xiang, Shumei Wei, Xingyu Liu, Xiaoyuan Liang, Lili Yang and Sheng Yan

77 ***Automatic Sequence-Based Network for Lung Diseases Detection in Chest CT***
Jinkui Hao, Jianyang Xie, Ri Liu, Huaying Hao, Yuhui Ma, Kun Yan, Ruirui Liu, Yalin Zheng, Jianjun Zheng, Jiang Liu, Jingfeng Zhang and Yitian Zhao

91 ***Molecular Imaging of Ulex Europaeus Agglutinin in Colorectal Cancer Using Confocal Laser Endomicroscopy (With Video)***
Weijun Wang, Shuxin Tian, Xin Jiang, Suya Pang, Huiying Shi, Mengke Fan, Zeyu Wang, Weiwei Jiang, Weiqian Hu, Xueyan Xiao and Rong Lin

100 ***Deep Learning Enables Prostate MRI Segmentation: A Large Cohort Evaluation With Inter-Rater Variability Analysis***
Yongkai Liu, Qi Miao, Chuthaporn Surawech, Haoxin Zheng, Dan Nguyen, Guang Yang, Steven S. Raman and Kyunghyun Sung

109 ***Low-Dose 68 Ga-PSMA Prostate PET/MRI Imaging Using Deep Learning Based on MRI Priors***

Fuquan Deng, Xiaoyuan Li, Fengjiao Yang, Hongwei Sun, Jianmin Yuan, Qiang He, Weifeng Xu, Yongfeng Yang, Dong Liang, Xin Liu, Greta S. P. Mok, Hairong Zheng and Zhanli Hu

119 ***Feasibility of Shear Wave Elastography Imaging for Evaluating the Biological Behavior of Breast Cancer***

Chaoxu Liu, Jin Zhou, Cai Chang and Wenxiang Zhi

129 ***Application of Intravoxel Incoherent Motion Diffusion-Weighted Imaging in Predicting and Monitoring Early Efficacy of Anti-Angiogenic Therapy in the C6 Glioma Rat Model***

Weishu Hou, Yangyang Xue, Yinfeng Qian, Hongli Pan, Man Xu, Yujun Shen, Xiaohu Li and Yongqiang Yu

140 ***AttR2U-Net: A Fully Automated Model for MRI Nasopharyngeal Carcinoma Segmentation Based on Spatial Attention and Residual Recurrent Convolution***

Jiajing Zhang, Lin Gu, Guanghui Han and Xiujian Liu

153 ***WEENet: An Intelligent System for Diagnosing COVID-19 and Lung Cancer in IoMT Environments***

Khan Muhammad, Hayat Ullah, Zulfiqar Ahmad Khan, Abdul Khader Jilani Saudagar, Abdullah AlTameem, Mohammed AlKhathami, Muhammad Badruddin Khan, Mozaherul Hoque Abul Hasanat, Khalid Mahmood Malik, Mohammad Hijji and Muhammad Sajjad

166 ***Baseline Amide Proton Transfer Imaging at 3T Fails to Predict Early Response to Induction Chemotherapy in Nasopharyngeal Carcinoma***

Zhou Liu, Liyan Zou, Qian Yang, Long Qian, Tianran Li, Honghong Luo, Canwen Che, Yuanyuan Lei, Peng Chen, Chunyan Qiu, Xin Liu, Yin Wu and Dehong Luo

177 ***Establishment of a Novel Mouse Hepatocellular Carcinoma Model for Dynamic Monitoring of Tumor Development by Bioluminescence Imaging***

Xiangyi Cao, Yulong Zhang, Qianqian Zhou, Sujing Sun, Minwei He, Xiaohui Wang, Ping Ma, Xiaoang Yang, Liping Lv and Linsheng Zhan

189 ***CHAIMELEON Project: Creation of a Pan-European Repository of Health Imaging Data for the Development of AI-Powered Cancer Management Tools***

Luis Martí Bonmatí, Ana Miguel, Amelia Suárez, Mario Aznar, Jean Paul Beregi, Laure Fournier, Emanuele Neri, Andrea Laghi, Manuela França, Francesco Sardanelli, Tobias Penzkofer, Phillipe Lambin, Ignacio Blanquer, Marion I. Menzel, Karine Seymour, Sergio Figueiras, Katharina Krischak, Ricard Martínez, Yisroel Mirsky, Guang Yang and Ángel Alberich-Bayarri

200 ***RDAU-Net: Based on a Residual Convolutional Neural Network With DFP and CBAM for Brain Tumor Segmentation***

Jingjing Wang, Zishu Yu, Zhenye Luan, Jinwen Ren, Yanhua Zhao and Gang Yu

212 ***Comprehensive Risk System Based on Shear Wave Elastography and BI-RADS Categories in Assessing Axillary Lymph Node Metastasis of Invasive Breast Cancer—A Multicenter Study***

Huiting Zhang, Yijie Dong, Xiaohong Jia, Jingwen Zhang, Zhiyao Li, Zhirui Chuan, Yanjun Xu, Bin Hu, Yunxia Huang, Cai Chang, Jinfeng Xu, Fajin Dong, Xiaona Xia, Chengrong Wu, Wenjia Hu, Gang Wu, Qiaoying Li, Qin Chen, Wanyue Deng, Qiongchao Jiang, Yonglin Mou, Huannan Yan, Xiaojing Xu, Hongju Yan, Ping Zhou, Yang Shao, Ligang Cui, Ping He, Linxue Qian, Jinping Liu, Liying Shi, Yanan Zhao, Yongyuan Xu, Yanyan Song, Weiwei Zhan and Jianqiao Zhou

222 ***Establishment and Validation of a Novel MRI Radiomics Feature-Based Prognostic Model to Predict Distant Metastasis in Endemic Nasopharyngeal Carcinoma***

Hao-Jiang Li, Li-Zhi Liu, Ying Huang, Ya-Bin Jin, Xiang-Ping Chen, Wei Luo, Jian-Chun Su, Kai Chen, Jing Zhang and Guo-Yi Zhang

233 ***Circulating MicroRNAs as Quantitative Biomarkers for Diagnosis and Prognosis of Uveal Melanoma***

Wen-Da Zhou, Lei Shao, Li Dong, Rui-Heng Zhang, Yi-Fan Li, He-Yan Li, Hao-Tian Wu, Xu-Han Shi and Wen-Bin Wei

243 ***Automatic Tumor Grading on Colorectal Cancer Whole-Slide Images: Semi-Quantitative Gland Formation Percentage and New Indicator Exploration***

Shenlun Chen, Meng Zhang, Jiazhou Wang, Midie Xu, Weigang Hu, Leonard Wee, Andre Dekker, Weiqi Sheng and Zhen Zhang

254 ***Prostate Surface Distension and Tumor Texture Descriptors From Pre-Treatment MRI Are Associated With Biochemical Recurrence Following Radical Prostatectomy: Preliminary Findings***

Rakesh Shiradkar, Soumya Ghose, Amr Mahran, Lin Li, Isaac Hubbard, Pingfu Fu, Sree Harsha Tirumani, Lee Ponsky, Andrei Purysko and Anant Madabhushi

266 ***Comparing the Blood Oxygen Level–Dependent Fluctuation Power of Benign and Malignant Musculoskeletal Tumors Using Functional Magnetic Resonance Imaging***

Lisha Duan, Huiyuan Huang, Feng Sun, Zhenjiang Zhao, Mengjun Wang, Mei Xing, Yufeng Zang, Xiaofei Xiu, Meng Wang, Hong Yu, Jianling Cui and Han Zhang

frontiers in Oncology

ORIGINAL RESEARCH
published: 04 October 2021
doi: 10.3389/fonc.2021.736654

Comparison of PET/CT and MRI in the Diagnosis of Bone Metastasis in Prostate Cancer Patients: A Network Analysis of Diagnostic Studies

*Fanxiao Liu[1], Jinlei Dong[1], Yelong Shen[2], Canhua Yun[3], Ruixiao Wang[4], Ganggang Wang[5], Jiyang Tan[6], Tao Wang[6], Qun Yao[6], Bomin Wang[1], Lianxin Li[1], Jingyi Mi[6], Dongsheng Zhou[1] and Fei Xiong[6,7]**

[1] Department of Orthopaedics, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China, [2] Department of Medical Imaging, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China, [3] Department of Nuclear Medicine, The Second Hospital of Shandong University, Jinan, China, [4] Department of Urology Surgery, University Hospital of Munich, Ludwig-Maximilians-University, Munich, Germany, [5] Department of Urology Surgery, Shandong Provincial Hospital Affiliated to Shandong University, Jinan, China, [6] Department of Sports Medicine, Wuxi 9th People's Hospital Affiliated to Soochow University, Wuxi, China, [7] Orthopaedic Institute, Medical College, Soochow University, Suzhou, China

OPEN ACCESS

Edited by:
Kyung Hyun Sung,
UCLA Health System, United States

Reviewed by:
Celine Taglang,
University of California, San Francisco, United States
Daniel Margolis,
Weill Cornell Medical Center, United States

***Correspondence:**
Fei Xiong
wxsdjrmyy_xf@outlook.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 05 July 2021
Accepted: 10 September 2021
Published: 04 October 2021

Citation:
Liu F, Dong J, Shen Y, Yun C, Wang R, Wang G, Tan J, Wang T, Yao Q, Wang B, Li L, Mi J, Zhou D and Xiong F (2021) Comparison of PET/CT and MRI in the Diagnosis of Bone Metastasis in Prostate Cancer Patients: A Network Analysis of Diagnostic Studies. Front. Oncol. 11:736654. doi: 10.3389/fonc.2021.736654

Background: Accurate diagnosis of bone metastasis status of prostate cancer (PCa) is becoming increasingly more important in guiding local and systemic treatment. Positron emission tomography/computed tomography (PET/CT) and magnetic resonance imaging (MRI) have increasingly been utilized globally to assess the bone metastases in PCa. Our meta-analysis was a high-volume series in which the utility of PET/CT with different radioligands was compared to MRI with different parameters in this setting.

Materials and Methods: Three databases, including Medline, Embase, and Cochrane Library, were searched to retrieve original trials from their inception to August 31, 2019 according to the Preferred Reporting Items for Systematic Review and Meta-analysis (PRISMA) statement. The methodological quality of the included studies was assessed by two independent investigators utilizing Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2). A Bayesian network meta-analysis was performed using an arm-based model. Absolute sensitivity and specificity, relative sensitivity and specificity, diagnostic odds ratio (DOR), and superiority index, and their associated 95% confidence intervals (CI) were used to assess the diagnostic value.

Results: Forty-five studies with 2,843 patients and 4,263 lesions were identified. Network meta-analysis reveals that 68Ga-labeled prostate membrane antigen (68Ga-PSMA) PET/CT has the highest superiority index (7.30) with the sensitivity of 0.91 and specificity of 0.99, followed by 18F-NaF, 11C-choline, 18F-choline, 18F-fludeoxyglucose (FDG), and 18F-fluciclovine PET/CT. The use of high magnetic field strength, multisequence, diffusion-weighted imaging (DWI), and more imaging planes will increase the diagnostic value of MRI for the detection of bone metastasis in prostate cancer patients. Where available, 3.0-T high-quality MRI approaches 68Ga-PSMA PET/CT was performed in the detection of bone

metastasis on patient-based level (sensitivity, 0.94 *vs.* 0.91; specificity, 0.94 *vs.* 0.96; superiority index, 4.43 *vs.* 4.56).

Conclusions: 68Ga-PSMA PET/CT is recommended for the diagnosis of bone metastasis in prostate cancer patients. Where available, 3.0-T high-quality MRI approaches 68Ga-PSMA PET/CT should be performed in the detection of bone metastasis.

Keywords: prostate cancer, bone metastasis, diagnostic study, network meta-analysis, MRI, PET/CT

1 INTRODUCTION

Prostate cancer (PCa) is the second most frequently diagnosed cancer and the fifth leading cause of cancer death in men according to global cancer statistics in 2018 (1). Although the 5-year survival rate is fairly high, the common cause of death is bone metastasis (2–5), which is the second most site of metastases in PCa (6, 7). Patients with an early diagnosis of localized disease may benefit from radical localized curative treatment (8), but patients who suffer from bone metastasis may be only eligible for hormone therapy or chemotherapy (9, 10). Hence, assessment of bone metastasis status, especially the early detection, is an important issue in the management of PCa. To follow and quantify the metastasis extent, which is an independent prognostic factor (11), the use of non-invasive imaging modalities is essential (12).

For decades, European and US guidelines recommend bone scintigraphy (BS) for bone metastasis diagnosis, which, if necessary, can be complemented by targeted X-rays (TXR) (13, 14). This BS/TXR association is imperfect, which strikingly lacks diagnostic specificity (15, 16), although the use of single-photon emission computed tomography (SPECT) improves the resolution (8, 17). Equivocal imaging results of BS are required to be determined by the additional use of magnetic resonance imaging (MRI); however, this multiple approach can add to the cost and become inconvenient for patients (18). The accuracy of MRI to detect bone metastasis has been highlighted for almost 30 years, and the superiority than BS has also been repeatedly suggested (18–21). A meta-analysis conducted by Woo et al. demonstrated the excellent diagnostic performance of MRI for the detection of bone metastasis in PCa (22). The development of new technology, such as diffusion-weighted imaging (DWI) (23), and the application of whole-body MRI (WB-MRI) may further expand the potential of MRI (20, 24, 25). However, the use of MRI for one-step cancer tumor–node–metastasis (TMN) staging is often presented as not feasible due to costs and the limited study validating this modality (26).

In recent years, positron emission tomography/computed tomography (PET/CT) has emerged as a promising molecular imaging tool in the diagnosis, staging, restaging, and therapeutic evaluation of several malignancies, as PET provides metabolic information and morphological imaging techniques offer anatomical data (27, 28). 18F-fluorodeoxyglucose (18F-FDG) is the most widely used PET-imaging agent in oncology detection; however, the low glycolytic rate of most skeletal metastases in PCa and the influence of bladder activity limit the sensitivity for clinical detection (7). The European Nuclear Medical Association recommended PET/CT in their guidelines for bone imaging in 2015 (29), which can show areas of altered osteogenic activity by using 18F-sodium fluoride (18F-NaF), a bone-specific radiotracer (30). Additionally, 11C- or 18F-choline are designed to target tumor cells directly (31), and the European Association of Urology (EAU) once suggested referring patients for 11C- or 18F-choline PET/CT when the prostate-specific antigen (PSA) increases >1 ng/ml, and the result is expected to change patient management (32). Over the last 5 years, 68Ga-labeled prostate-specific membrane antigen (68Ga-PSMA) PET/CT has gained widespread use to assess PCa (33, 34), which could identify metastatic lesions in lymph node, bone, and soft tissue at low PSA levels (33–35). Therefore, 68Ga-PSMA PET/CT is regarded as a more specific modality for diagnosing osseous metastases in PCa (36).

Despite the increasing numbers of studies regarding PET/CT and MRI in the diagnosis procedure for bone metastases in PCa, the effectiveness of these two modalities still remains no consensus. Zhou et al. (37) compared PET/CT and MRI; the final conclusion was very general because of the limitation of traditional meta-methods, which could not directly compare the PET/CT using different radioligands and MRI with different parameters. Recently, Nyaga et al. (38) developed a Bayesian network meta-analysis using an arm-based model based on the assumption that the missing arms occur at random. This method has been applied in several studies (39, 40) because it could allow analysis of the variability in the accuracy of multiple tests within and between studies simultaneously (41).

The arm-based model is more appealing than traditional meta-analysis and the contrast-based model since the former not only permits more straightforward interpretation of the parameters, making use of all available data and yielding shorter credible intervals, but also provides more natural variance–covariance matrix structures. We adopted this model, which makes our results more convincing.

Thus, the primary aim of our meta-analysis is to compare the diagnostic accuracy of PET/CT and MRI in detecting bone metastases in PCa on a per-patient and per-lesion basis, respectively. Additionally, thanks to the establishment of the network meta-analysis, a direct comparison is performed between PET/CT with different radioligands and MRI with different magnet field strengths, coverage, and parameters to provide better evidence-based advice to physicians. The hypothesis is that

the multiparametric MRI equipped with high magnet field strength is an appropriate modality for the diagnosis of bone metastasis in PCa.

2 MATERIALS AND METHODS

2.1 Protocol and Guidance

This network meta-analysis was conducted in accordance with the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA) (42) statement and PRISMA (PROSPERO registration number CRD42020148865).

2.2 Inclusion Criteria

The included studies should meet all of the following inclusion criteria: clinical trials evaluating the diagnostic value of PET/CT or/and MRI for bone metastasis in prostate cancer patients; articles published providing data to calculate diagnostic parameters, including true positive (TP), false positive (FP), false negative (FN), and true negative (TN); and studies having a conclusive anatomical or morphological verification standard to prove or disprove the imaging study result, such as pathological examination or clinical confirmation.

2.3 Exclusion Criteria

We excluded studies if they were commentaries, letters, case reports, reviews or non-full-text studies; if all the participants were prostate patients with bone metastasis; and if they lack of the conclusive anatomical or morphological verification standard to prove or disprove the imaging study result, such as pathological examination or clinical confirmation.

2.4 Outcomes

The outcomes are the absolute sensitivity and specificity, relative sensitivity and specificity, diagnostic odds ratio (DOR), and superiority index of PET/CT with numerous tracers and MRI with different parameters for the diagnosis of bone metastasis in prostate cancer patients. **Supplementary Table S1** shows the definition of these outcomes.

2.5 Search Strategy

Three databases, including Medline, Embase, and Cochrane Library, were searched to retrieve original trials from their inception to August 31, 2019. We also searched ClinicalTrials.gov and the World Health Organization International Clinical Trials Registry Platform to identify ongoing or unpublished eligible trials. To maximize the search for relevant articles, a manual search of the references listed in all included trials and systematic reviews was performed to retrieve any relevant articles that were not listed in the databases. **Supplementary Table S2** shows the search strategy.

2.6 Study Selection

After removal of duplicates, two investigators performed a blind systematic screening for all titles and abstracts in duplicate. Then, the full texts of the remainders were downloaded to confirm their eligibility based on the above criteria. To maximize the sensitivity of the screen, disagreements at the title and abstract stages were resolved by automatic inclusion, whereas discrepancies at the full-text stage were resolved by consensus with input from a senior third investigator.

2.7 Data Collection Process

Two independent investigators used a predesigned Microsoft Excel spreadsheet (Version 2013, Microsoft, Redmond, WA, USA) to extract basic information from the included studies. The diagnostic data (TPs, FPs, FNs, and TNs) were extracted or calculated using the following methods, which are presented in **Supplementary Table S3**. The spreadsheets were combined, and each investigator checked a random selection of the other's entries for quality control. Any discrepancies were resolved by consensus.

2.8 Quality Assessment of Included Studies

The methodological quality of the included studies was assessed by two independent investigators utilizing Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool, which comprised of four key domains (patient selection, index test, reference standard, and flow and timing). Detailed information is shown in **Supplementary Table S4**.

2.9 Data Synthesis

A Bayesian network meta-analysis using an arm-based model, developed by Nyaga et al., was performed by running three chains in parallel until there is convergence. We used absolute sensitivity and specificity, relative sensitivity and specificity, DOR, and superiority index and their associated 95% confidence intervals (CI) to assess the diagnostic value of PET/CT and MRI. To assess the relative performance of diagnostic tests, the definitions of superior, inferior, equal, and not comparable were drawn. A diagnostic test, which is pairwise superior to a relatively large number of other tests and pairwise inferior to relatively few other tests, should have a high superiority value and be ranked higher than those tests that do not perform as well. In this network meta-analysis, superiority index was pooled to quantify rank probabilities of a PET/CT and MRI. All network meta-analyses were performed using R (v3.4.3; Comprehensive R Archive Network). **Supplementary Table S5** shows the main information of implementation process, software's packages, and models.

2.9.1 Subgroup Analyses

For different tracers of PET/CT, we performed several subgroup analyses according to clinical settings of prostate cancer (new diagnoses, mixed, and treated), number of patients (<50 and ≥50), patient age (60–70 and > 70), continent of origin (Europe and others), study design (prospective and retrospective), and methods of imaging analyses (visual and semiquantitative evaluation). For the parameters of MRI, we performed several subgroup analyses according to magnetic field strength (1.5 and 3.0 T), the number of sequences (single sequences, or ≥ multisequences), whether using DWI or not, number of imaging planes (1 or ≥2), and MRI coverage (pelvis skeleton, axial skeleton, or whole-body skeleton).

2.9.2 Sensitivity Analyses

The sensitivity analyses were performed by deleting studies involving only one diagnostic test for detecting bone metastasis, studies with low a QUADAS-2 score, studies with the maximum cases, studies with the minimum cases, and studies published before 2010.

2.10 Patient and Public Involvement

No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of the study. No patients were asked to advise on interpretation or writing up of results. There are no plans to disseminate the results of the research to study participants or the relevant patient community. It was not evaluated whether the studies included in the review had any patient involvement.

3 RESULTS

3.1 Eligible Studies and Study Characteristics

We initially identified 24,491 records and included 45 eligible studies (8, 43–86) in the final network meta-analysis (**Figure 1**). The studies comprised 2,843 participants (4,263 lesions), with 978 bone metastasis patients (2,186 lesions), and a total of 32 studies assessed the diagnostic value of PET/CT and 21 studies for MRI. **Table 1** shows the main characteristics of the included studies. **Supplementary Tables S6**, **S7** show the main technical parameters of MRI and PET/CT, respectively. **Supplementary Tables S8**, **9** show the diagnostic data of each included studies on patient- and lesion-based level, respectively. **Supplementary Table S10** shows the quality assessment of the included studies. Four (8.9%), 11 (24.4%), 18 (40.0%), 11 (4.5%), and 1 (2.2%) studies scored 11, 10, 9, 8, and 7, respectively.

3.2 Diagnostic Value of PET/CT on Per-Patient Analysis

Regarding the tracers of the included studies, nine used 18F-NaF, seven selected 18F-choline, seven applied 11C-choline, five used 68Ga-PSMA, three selected 18F-FDG, and two applied 18F-FACBC. As shown in **Figure 2**, 18F-NaF PET/CT has the highest sensitivity of 0.95 (95% CI, 0.91–0.99), followed closely by 68Ga-PSMA and 18F-choline PET/CT; 68Ga-PSMA PET/CT has the highest specificity of 0.99 (95% CI, 0.94-1.04), followed closely by 11C-choline and 18F-choline PET/CT. Network meta-analysis demonstrated that 68Ga-PSMA PET/CT has the highest diagnostic value with the highest superiority index of 7.30 (95% CI, 0.60–11.00), followed closely by 18F-NaF (3.33; 95% CI, 0.20–9.00), 11C-choline, 18F-choline, 18F-FACBC, and 18F-FDG PET/CT (**Table 2**).

3.2.1 Sensitivity Analysis

The sensitivity analyses were performed by deleting studies involving only one diagnostic test for detecting bone metastasis, studies with low QUADAS-2 score, studies with the maximum cases, studies with the minimum cases, and studies published before 2010, and the results were stable (**Supplementary Tables S11**–**S15**).

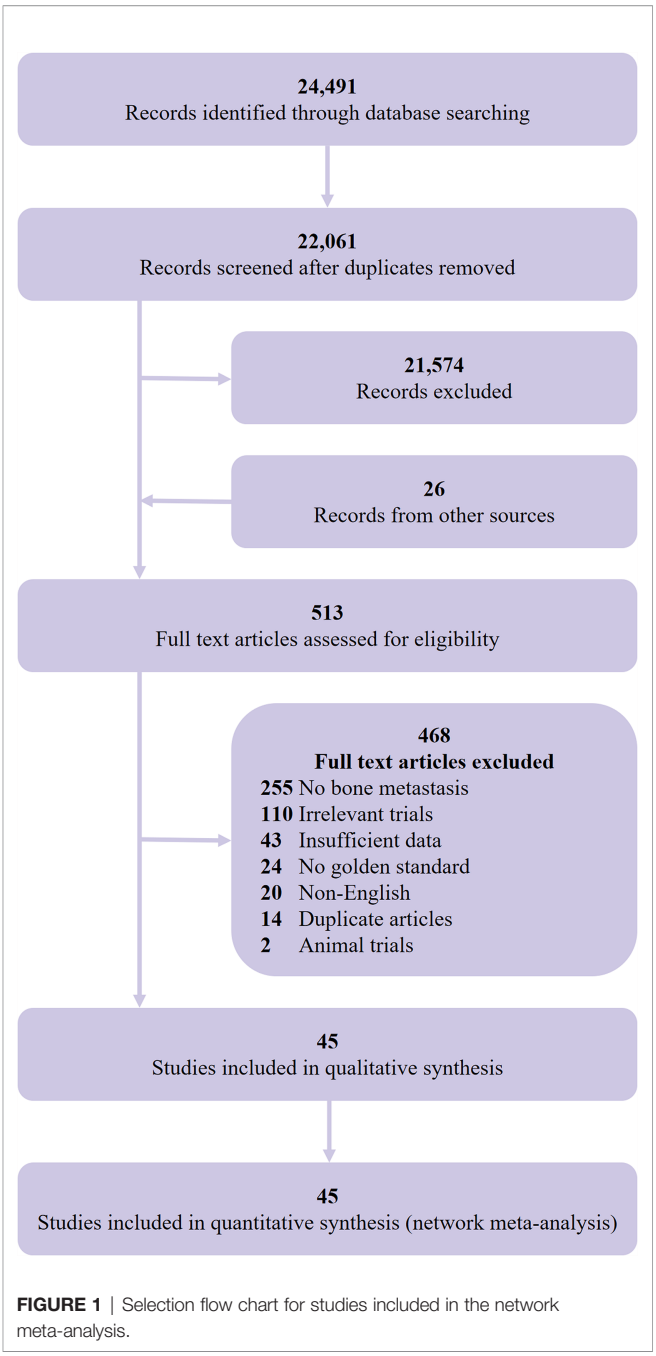


FIGURE 1 | Selection flow chart for studies included in the network meta-analysis.

3.2.2 Subgroup Analysis

The subgroup analyses were performed according to clinical settings of prostate cancer (newly diagnoses, mixed, and treated), number of patients (<50 and ≥50), patient age (60–70 and >70), continent of origin (Europe and others), study design (prospective and retrospective), and methods of imaging analyses (visual and semiquantitative evaluation), and the results were stable (**Supplementary Tables S16**–**S21**).

TABLE 1 | Main characteristics of the included studies.

Study, year	Country	No. of patients			Clinical setting	Age (years)		PSA (ng/ml)		Clinical T stage	Gleason score		Type of study	Inclusion interval
		Total (n)	Metastasis		New or treated	Mean ± SD	Range	Mean ± SD	Range	Range	Median	Range		
			n	%										
Even-Sapir et al. (8)	Israel	44	23	52.27	Mixed	71.6 ± 8.8	NR	NR	≥20	NR	NR	≥8	P	NR
Eschmann et al. (43)	Germany	44	44	100.00	Mixed	Median 64.1	51–79	Median 5.4	0.15–200	T1–T4	NR	NR	R	11.2004–01.2006
Lecouvet et al. (44)	USA	66	41	62.12	Mixed	74	46–85	NR	NR	NR	NR	NR	P	NR
Nemeth et al. (45)	USA	8	7	87.50	NR	NR	NR	NR	NR	NR	NR	NR	R	1998–2004
Beheshti et al. (46)	Austria	38	NR	–	Mixed	69 ± 8	NR	56 ± 64	NR	NR	NR	NR	P	NR
Beheshti et al. (6)	Austria	70	NR	–	Mixed	68 ± 7	NR	NR	≥10	NR	NR	≥7	P	NR
Venkitaraman et al. (47, 48)	UK	99	14	14.14	Newly diagnosed	Median 66	44–83	Median 26.5	2–1600	NR	7	6–10	P	12.2001–12.2005
Venkitaraman et al. (47, 48)[#]	UK	39	10	25.64	Newly diagnosed	Median 65	54–82	Median 34	5–1300	T1–T4	8	5–9	P	12.2001–07.2004
Fuccio et al. (50)	Italy	25	22	88.00	Treated	70.2	58–80	Median 6.3	0.2–37.7	T2N0/Nx M0-T4N1M0	7	6–9	R	NR
Iagaru et al. (51)	USA	18	9	50.00	Mixed	NR	NR	NR	NR	NR	NR	NR	P	09.2007–12.2010
Langsteger et al. (52)	Austria	40	22	55.00	Mixed	66	51–82	NR	0.38–617	NR	NR	4–9	P	01.2003–12.2009
Bortot et al. (53)	Brazil	9	2	22.22	NR	67.6	56–82	NR	NR	NR	NR	NR	P	NR
Jadvar et a (54)	USA	37	14	37.84	Treated	Median 71.1	53.5–86.9	Median 3.2	0.5–40.2	T1c–T3	NR	NR	P	22.09.2010–23.06.2011
Lecouvet et al. (55)	Belgium	100	51	51.00	Mixed	69	53–88	32	12–78	≥T3b	NR	≥8	P	03.2007–03.2010
Mosavi et al. (56)	Sweden	49	5	10.20	Newly diagnosed	Median 67	57–80	Median 14	1.3–950	T1c–T4	9	8–10	P	01.2009–03.2011
Picchio et al. (57)	Italy	78	27	34.62	Treated	69	47–82	21.1	0.2–500.0	T2N0–T4N0	NR	NR	R	03.2005–02.2010
Takesh et al. (58)	Germany	37	18	48.65	Treated	69 ± 7	NR	2.6	0.3–21	NR	7	3–9	R	NR
Damle et al. (59)	India	49	32	65.31	Mixed	65	50–84	NR	NR	T3/T4	NR	8–10	P	NR
Kitajima et al. (60)	USA	95	16	16.84	Treated	65.7	49–87	Median 2.5	0.58–68.3	T2N0–any T N1	7	2–10	R	12.2011–01.2013
Pasoglou et al. (61)	Belgium	30	9	30.00	Treated	Median 62.5	51.0–92.0	Median 30	1.7–4612.0	cT2–cT4	7.8	6–9	P	NR
Piccardo et al. (62)	Italy	21	6	28.57	Treated	77.2 ± 5.1	70–85	5.8 ± 3.4	2.2–13.4	NR	8	7–9	P	NR
Poulsen et al. (63)	Denmark	50	NR	–	Treated	73 ± 8.6	53–94	Median 84	4–5740	NR	7	5–10	P	05.2009–03.2012
Evangelista et al. (64)	Italy	48	11	22.92	Newly diagnosed	70	49–86	38.34 ± 90.12	2.80–581.0	T2–T4	NR	6–10	R	04.2010-04.2013
Pasoglou et al. (65)	Belgium	30	10	33.33	Mixed	69	NR	31 ± 28	NR	NR	NR	NR	P	02.2012–12.2012
Sampath et al. (66)	USA	38	22	57.89	NR	NR	NR	NR	NR	NR	NR	NR	R	09.2007–07.2013
Wieder et al. (67)	Germany	57	50	87.72	Treated	Median 86	54–80	29.9	1–670	NR	8	6–9	P	NR
Barchetti et al. (68)	Italy	152	70	46.05	Treated	NR	53–88	NR	NR	NR	NR	≥7	P	09.2011–01.2014
Conde-Moreno et al. (69)	Spain	35	17	48.57	Treated	Median 70 ± 6.77	52–80	Median 12	4.54–75.86	T1N0M0–T4N0M0	7	5–9	P	01.2014–03.2015
Nanni et al. (70)	Italy	89	6	6.74	Mixed	69	55–83	6.99	0.20–20.72	T1N0/Nx–T3N0/Nx	NR	NR	P	NR
Woo et al. (71)	Korea	308	21	6.82	Newly diagnosed	68.5 ± 7.8	38–91	30.9	1.2–955.5	NR	7	6–10	R	01.2013-12.2013

(Continued)

TABLE 1 | Continued

Study, year	Country	No. of patients			Clinical setting	Age (years)		PSA (ng/ml)		Clinical T stage	Gleason score		Type of study	Inclusion interval
		Total (n)	Metastasis		New or treated	Mean ± SD	Range	Mean ± SD	Range	Range	Median	Range		
			n	%										
Yi et al. (72)	China	26	12	46.15	Mixed	72.2	60–88	NR	≥20	T2–T4	NR	8–10	P	08.2010–11.2014
Fonager et al. (73)	Denmark	37	27	72.97	Mixed	71	46–87	Median 180	53–9708	T1–T4	9	7–10	R	02.2014–12.2015
Huysse et al. (74)	Belgium	64	62	96.88	Treated	NR	NR	NR	NR	NR	NR	NR	P	NR
Janssen et al. (75)	Germany	54	29	53.70	Mixed	69.6 ± 6.5	NR	38.4 ± 77.9	NR	NR	NR	NR	P	NR
Kitajima et al. (76)	Japan	21	11	52.38	Mixed	70.6 ± 10.8	47–90	342.9	0.2–5916	NR	NR	NR	P	01.2015–01.2017
Vargas et al. (77)	USA	228	57	25.00	Newly diagnosed	Median 63	36-83	Median 6.3	0.4–222	T1c–T4	7	6–≥8	R	01.2000–06.2014
Wondergem et al. (78)	Netherlands	104	61	58.65	Mixed	74.9	49–93	Median 88.7	2.5–13500	T1–T4	9	6–10	R	01.2011–04.2012
Dyrberg et al. (79)	Denmark	55	20	36.36	Newly diagnosed	75 ± 9	54–91	85	5–1000	NR	8	6–10	P	05.2016–06.2017
Kawanaka et al. (80)	Japan	30	17	56.67	Treated	71.3 ± 9.0	47–90	65.2 ± 177.4	0.23–946	NR	NR	NR	R	01.2015–07.2017
Larbi et al. (81)	Belgium	50	37	74.00	Newly diagnosed	67 ± 10	59–87	NR	≥20	NR	NR	≥ 8	R	01.2015–12.2015
Lengana et al. (82)	South Africa	113	26	23.01	Mixed	66.65	43–88	NR	NR	NR	NR	NR	P	NR
Zacho et al. (83)	Denmark	68	10	14.71	Treated	67.2	47–80	NR	0.2–11	M0–M1	7	5–9	P	NR
Chen et al. (84)	USA	106	14	13.21	Mixed	Median 70	47–80	Median 1.3	0–61	NR	8	6–10	R	01.2017–01.2018
Johnston et al. (85)	UK	56	5	8.93	Newly diagnosed	67.9	57.9–84.4	Median 20.05	10.07–61.20	NR	7	6–10	P	07.2012–11.2015
Uslu–besli et al. (86)	Turkey	28	11	39.29	Mixed	67.3 ± 7.4	49–82	25.49 ± 32.7	0.5–125.1	NR	7	6–9	R	03.2015-03.2016

SD, Standard deviation; NR, Not reported; PSA, Prostate specific antigen; P, Prospective; R, Retrospective. [#]*Represents different study.*

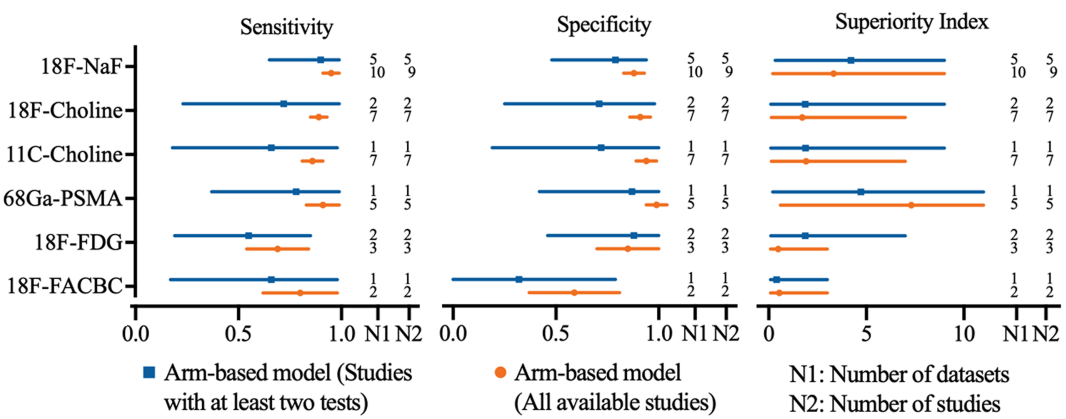


FIGURE 2 | Network meta-analysis results including sensitivity, specificity, and superiority index values of PET/CT with six commonly used tracers for the detection of bone metastasis in prostate cancer patients. Sensitivity, specificity, and superiority index are reported as mean (range) unless otherwise indicated. PET/CT, positron emission tomography/computed tomography; NaF, sodium fluoride; PSMA, prostate membrane antigen; FDG, fludeoxyglucose; FACBC, fluciclovine.

TABLE 2 | The network meta-analysis results of PET/CT with different tracers to detect bone metastasis in PCa.

Test	Absolute Sensitivity	Absolute Specificity	Diagnostic OR [Rank]	Superiority Index [Rank]	Relative Sensitivity	Relative Specificity	Datasets, n	Studies, n
18F-NaF	0.95 (0.91–0.99)	0.88 (0.83–0.93)	248.15 (34.14–799.75) [2]	3.33 (0.20–9.00) [2]	1.00 (1.00–1.00)	1.00 (1.00–1.00)	10	9
18F-Choline	0.89 (0.85–0.93)	0.91 (0.86–0.96)	123.18 (21.09–361.86) [4]	1.71 (0.14–7.00) [4]	0.94 (0.88–1.00)	1.03 (0.95–1.11)	7	7
11C-Choline	0.86 (0.81–0.91)	0.94 (0.89–0.99)	208.88 (21.37–780.88) [3]	1.92 (0.14–7.00) [3]	0.91 (0.84–0.98)	1.07 (0.98–1.16)	7	7
68Ga-PSMA	0.91 (0.83–0.99)	0.99 (0.94–1.04)	3379817.37 (49.99–5941029.19) [1]	7.30 (0.60–11.00) [1]	0.96 (0.87–1.05)	1.12 (1.04–1.20)	5	5
18F-FDG	0.69 (0.54–0.84)	0.85 (0.70–1.00)	81.17 (1.35–497.30) [6]	0.49 (0.09–3.00) [6]	0.73 (0.57–0.89)	0.96 (0.78–1.14)	3	3
18F-FACBC	0.80 (0.62–0.98)	0.59 (0.37–0.81)	92.40 (0.40–598.84) [5]	0.55 (0.09–3.00) [5]	0.84 (0.65–1.03)	0.67 (0.42–0.92)	2	2

Data are reported as mean (range) unless otherwise indicated.

PET/CT, positron emission tomography/computed tomography; PCa, prostate cancer; NaF, sodium fluoride; PSMA, prostate membrane antigen; FDG, fludeoxyglucose; FACBC, fluciclovine; CI, credible interval; OR, odds ratio.

3.3 Diagnostic Value of MRI on Per-Patient Analysis

3.3.1 Magnetic Field Strength of MRI

The pooled sensitivity and specificity of 1.5-T MRI were 0.82 (95% CI, 0.73–0.91) and 0.92 (95% CI, 0.87–0.97), respectively, while the pooled sensitivity and specificity of 3.0-T MRI were 0.89 (95% CI, 0.80–0.98) and 0.88 (95% CI, 0.79–0.97). Network meta-analysis demonstrated that 68Ga-PSMA PET/CT had the highest diagnostic value with the highest superiority index, followed closely by 18F-NaF PET/CT, 3.0-T MRI [1.76, (95% CI, 0.09–9.00)], 11C-choline PET/CT, 18F-choline PET/CT, and 1.5-T MRI [0.77, (95% CI, 0.09–5.00)] (**Supplementary Table S22** and **Figure 3**).

3.3.2 Sequence of MRI

The pooled sensitivity and specificity of multisequence MRI were 0.91 (95% CI, 0.87–0.95) and 0.93 (95% CI, 0.90–0.96), while the pooled sensitivity and specificity of single-sequence MRI were 0.64 (95% CI, 0.48–0.80) and 0.89 (95% CI, 0.76–1.02), respectively. Network meta-analysis demonstrated that 68Ga-PSMA PET/CT had the highest diagnostic value with the highest superiority index, followed closely by 18F-NaF PET/CT, multisequence MRI [2.04, (95% CI, 0.14–7.00)], 11C-choline PET/CT, 18F-choline PET/CT, and single-sequence MRI [0.36, (95% CI, 0.09–1.02)] (**Supplementary Table S23** and **Figure 3**).

3.3.3 Whether DWI Was Used

The pooled sensitivity and specificity of DWI MRI were 0.94 (95% CI, 0.89–0.99) and 0.93 (95% CI 0.88–0.98), while the pooled sensitivity and specificity of no-DWI MRI were 0.86 (95% CI, 0.71–1.01) and 0.86 (95% CI, 0.74–0.98), respectively. Network meta-analysis demonstrated that 68Ga-PSMA PET/CT had the highest diagnostic value with the highest superiority index, followed closely by DWI MRI [3.63, (95% CI 0.14–9.00)], 18F-NaF PET/CT, no-DWI MRI [1.44, (95% CI, 0.09–9.00)], 11C-choline PET/CT, and 18F-choline PET/CT (**Supplementary Table S24** and **Figure 3**).

3.3.4 Imaging Plane of MRI

The pooled sensitivity and specificity of ≥2 imaging planes MRI were 0.95 (95% CI, 0.90–1.00) and 0.93 (95% CI, 0.89–0.97), while the pooled sensitivity and specificity of 1 imaging plane MRI were 0.75 (95% CI, 0.69–0.81) and 0.92 (95% CI, 0.86–0.98), respectively. Network meta-analysis revealed that 68Ga-PSMA PET/CT had the highest diagnostic value with the highest superiority index, followed closely by two or more imaging planes MRI [4.27, (95% CI, 0.20–9.00)], 18F-NaF PET/CT, 11C-choline PET/CT, 18F-choline PET/CT, and one imaging plane MRI [0.25, (95% CI, 0.09–1.00)] (**Supplemental Table S25** and **Figure 3**).

3.3.5 MRI Coverage

The pooled results demonstrated that axial skeleton or whole-body skeleton MRI had similar sensitivity of 0.84 *vs.* 0.82 and specificity of 0.93 *vs.* 0.94, compared with pelvis MRI. Network meta-analysis demonstrated that 68Ga-PSMA PET/CT had the highest diagnostic value with the highest superiority index, followed closely by 18F-NaF PET/CT, 11C-choline PET/CT, pelvis skeleton MRI, 18F-choline PET/CT, and axial skeleton or whole-body skeleton MRI (**Supplementary Table S26** and **Figure 3**).

3.3.6 Sensitivity Analysis

In five subgroup analyses above, we performed sensitivity analysis by deleting studies involving only one test for detecting bone metastasis, which show the similar results (**Supplementary Tables S27–S31**).

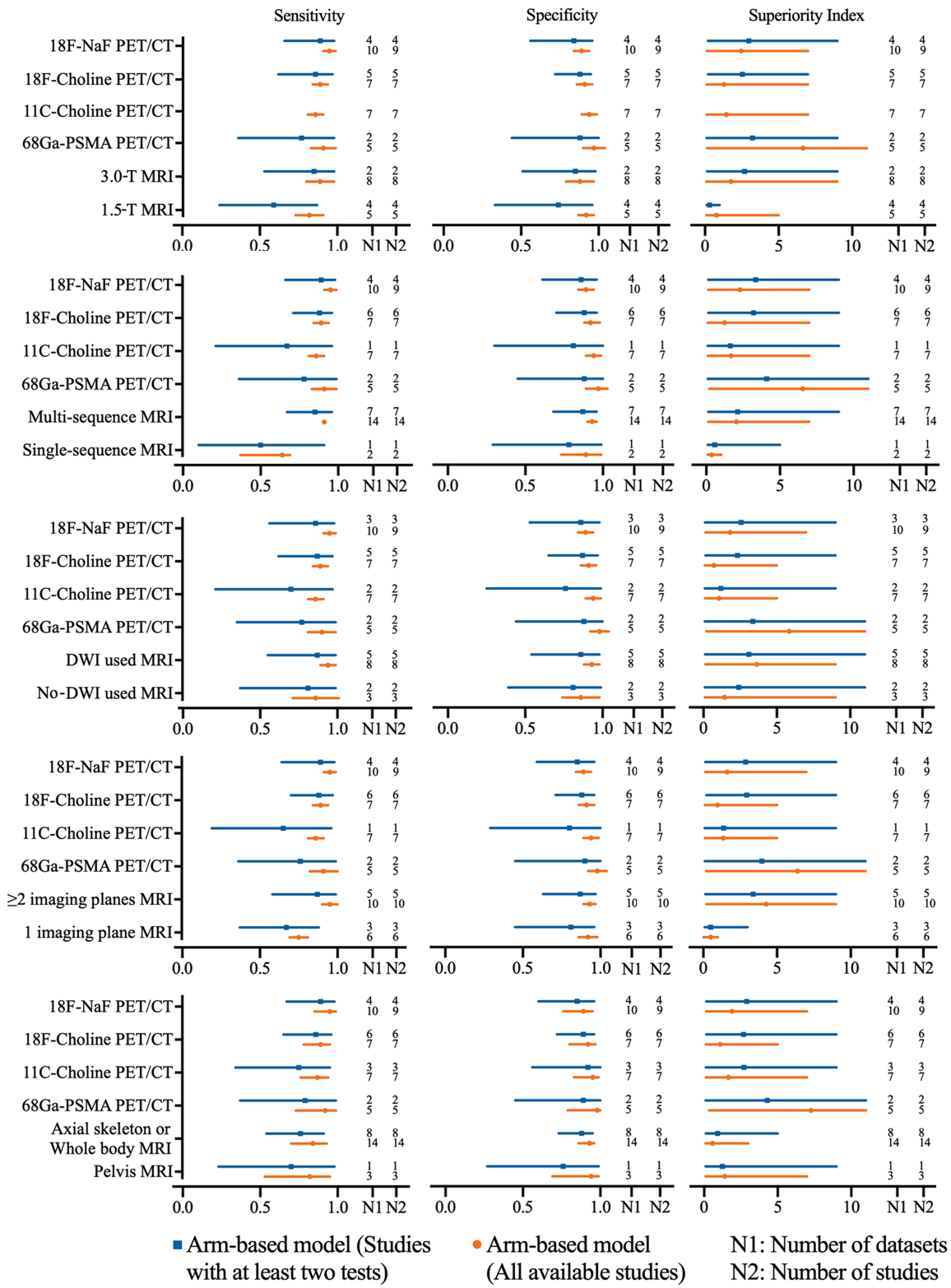


FIGURE 3 | Network meta-analysis results including sensitivity, specificity, and superiority index values of PET/CT with four commonly used tracers and MRI with numerous characteristics for the detection of bone metastasis. Sensitivity, specificity, and superiority index are reported as mean (range) unless otherwise indicated. PET/CT, positron emission tomography/computed tomography; NaF, sodium fluoride; PSMA, prostate membrane antigen; DWI, diffusion-weighted imaging.

3.4 Diagnostic Value of High-Quality MRI *vs.* PET/CT on Per-Patient Analysis

In order to achieve the highest accuracy of MRI, we define a high-quality MRI which is referred to the MRI equipped with multisequence, DWI used, and ≥2 imaging planes. High-quality MRI (1.5-T) has a sensitivity of 0.96 (95% CI, 0.90–1.02) and a specificity of 0.90 (95% CI, 0.81–0.99), while 3.0-T high-quality MRI has a sensitivity of 0.94 (95% CI, 0.86–1.02) and a specificity of 0.94 (95% CI, 0.86–1.02), respectively. Network meta-analysis demonstrates that 68Ga-PSMA PET/CT has the highest diagnostic value with the highest superiority index [4.56, (95% CI, 0.11–11.00)], followed closely by 3.0-T high-quality MRI [4.43, (95% CI, 0.14–11.00)], 1.5-T high-quality MRI [3.38, (95% CI, 0.11–9.00)], 18F-NaF PET/CT, 11C-choline PET/CT, and 18F-choline PET/CT (**Table 3**). The sensitivity analyses were performed by deleting studies involving only one test for detecting bone metastasis, and the results were stable (**Supplemental Table S32** and **Figure 4**).

3.5 Diagnostic Value of PET/CT *vs.* MRI on Per-Lesion Analysis

Regarding the tracers of the included studies, three studies used 18F-NaF, four selected 18F-choline, two applied 11C-choline, and one selected 68Ga-PSMA PET/CT. Only five studies reported MRI data. The pooled results demonstrated that 11C-choline PET/CT had the highest sensitivity of 0.84 (95% CI, 0.70–0.98), and 18F-choline PET/CT has the highest specificity of 0.91 (95% CI, 0.83–0.99). Network meta-analysis revealed that 11C-choline PET/CT had the highest diagnostic value with the highest superiority index, followed closely by 68Ga-PSMA PET/CT, 18F-choline PET/CT, MRI, and 18F-NaF PET/CT (**Supplementary eTable 33**). Subgroup analysis based on the analysis of high-quality MRI (only 1.5-T, insufficient data for 3.0-T high-quality MRI) showed similar results (**Supplementary Table S34**).

4 DISCUSSION

"In the current meta-analysis, we compared the diagnostic accuracy of MRI and PET/CT for the detection of bone metastasis in PCa. On patient-based level, network meta-analysis reveals that for numerous tracers, 68Ga-PSMA PET/CT has the highest superiority index, followed by 18F-NaF, 11C-choline, 18F-choline, 18F-FDG, and 18F-fluciclovine PET/CT; for the parameters of MRI, higher magnetic field strength, multisequence, more imaging planes, and MRI with DWI will increase the diagnostic value for bone metastasis in prostate cancer patients. Where available, 3.0-T high-quality MRI approaches 68Ga-PSMA PET/CT was performed in the detection of bone metastasis (sensitivity, 0.94 *vs.* 0.91; specificity, 0.94 *vs.* 0.96; superiority index, 4.43 *vs.* 4.56).

Given this hybrid method, as PET provides metabolic information and morphological imaging techniques offer anatomical data (27, 28, 87, 88), a more accurate delineation of bone metastases is allowed. As well-known, the radioactive tracer is one of the most important cores of nuclear medicine imaging. Antoch et al. (28) demonstrated that the selection of the appropriate radioligands could increase accuracy when detecting micrometastases. Hence, we conducted the analysis based on the several popular tracers and found that 68Ga-PSMA PET/CT possessed the highest diagnostic value [superiority index, 7.30 (95% CI, 0.60–11.00)]. Pyka et al. (89) performed a retrospective study and showed a higher sensitivity and specificity of 68Ga-PSMA PET/CT (100% and 100%) when

TABLE 3 | PET/CT with different tracers and high-quality MRI to detect bone metastasis in PCa.

Test	Absolute Sensitivity	Absolute Specificity	Diagnostic OR[Rank]	Superiority Index [Rank]	Relative Sensitivity	Relative Specificity	Datasets, n	Studies, n
18F-NaF PET/CT	0.95 (0.91–0.99)	0.89 (0.84–0.94)	256.89 (34.08–796.33) [4]	1.08 (0.11–5.00) [4]	1.00 (1.00–1.00)	1.00 (1.00–1.00)	10	9
18F-Choline PET/CT	0.89 (0.84–0.94)	0.91 (0.86–0.96)	121.19 (22.43–355.84) [6]	0.51 (0.09–3.00) [6]	0.94 (0.87–1.01)	1.03 (0.95–1.11)	7	7
11C-Choline PET/CT	0.86 (0.81–0.91)	0.94 (0.89–0.99)	212.63 (20.64–791.67) [5]	0.75 (0.09–5.00) [5]	0.91 (0.84–0.98)	1.07 (0.98–1.16)	7	7
68Ga-PSMA PET/CT	0.91 (0.83–0.99)	0.96 (0.87–1.05)	4,633,299.15 (14.51–4,438,033.15) [1]	4.56 (0.11–11.00) [1]	0.96 (0.86–1.06)	1.09 (0.96–1.22)	5	5
3.0-T high-quality MRI	0.94 (0.86–1.02)	0.94 (0.86–1.02)	6033.33 (27.09–39,719.47) [2]	4.43 (0.14–11.00) [2]	0.99 (0.89–1.09)	1.07 (0.96–1.18)	4	4
1.5-T high-quality MRI	0.96 (0.90–1.02)	0.90 (0.81–0.99)	2,056.83 (23.96–11,589.08) [3]	3.38 (0.11–9.00) [3]	1.02 (0.93–1.11)	1.02 (0.90–1.14)	4	4

Data are reported as mean (range) unless otherwise indicated.

PET/CT, positron emission tomography/computed tomography; MRI, magnetic resonance imaging; PCa, Prostate cancer; NaF, Sodium fluoride; PSMA, Prostate membrane antigen; T, Tesla; CI credible interval; OR, odds ratio.

High-quality MRI was referred to the MRI equipped with multisequence, DWI used, and ≥2 imaging planes.

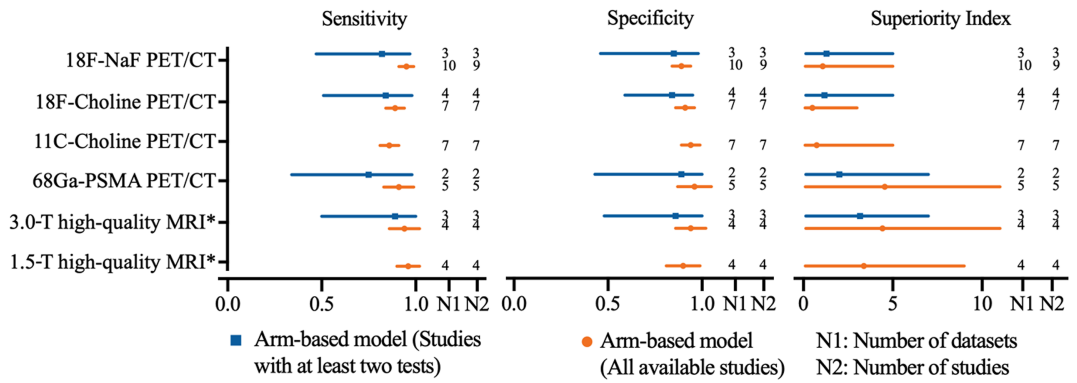


FIGURE 4 | Network meta-analysis results including sensitivity, specificity, and superiority index values of PET/CT with four commonly used tracers and high-quality MRI for the detection of bone metastasis in prostate cancer patients. Sensitivity, specificity, and superiority index are reported as mean (range) unless otherwise indicated. PET/CT, positron emission tomography/computed tomography; NaF, sodium fluoride; PSMA, prostate membrane antigen; DWI diffusion-weighted imaging; *high-quality MRI was referred to the MRI equipped with multi-sequence, DWI used, and ≥2 imaging planes.

compared to BS. Uslu-Besli et al. (86) conducted a cohort study of 28 patients and demonstrated that 68Ga-PSMA PET/CT changed management of seven patients by confirming the presence of bone metastasis, which was overlooked by BS in three patients and by excluding the false negative lesions on BS in four patients. Through binding to PSMA, a transmembrane protein expressed predominantly in prostate cells and especially in prostate cancer cells, 68Ga-PSMA leads to the internalization and accumulation in primary and metastatic cancer lesions (90), which is proven to have the higher yield of positive scans than obtained with other tracers at a low PSA level (91, 92). However, the half-life and yield of 68Ga significantly limited the ability 68Ga-PSMA to meet the demand for imaging in PCa. To this regard, 18F-PSMA was considered as the ideal PET/CT tracer (93), but we were not able to analyze it because of the insufficient study.

Although the recommendation of the use of 18F-NaF PET/CT in PCa was refrained in guidelines, it is routinely used worldwide (94, 95). Our results showed that 18F-NaF PET/CT had the highest sensitivity (0.95) but relative low specificity (0.88). The uptake of 18F-NaF, a bone-specific imaging radiotracer, is correlated to the blood flow and, especially, to the activity of local osteoblasts (30, 96–98), and the study by the National Oncologic PET Registry (NOPR) showed that the intended managements in approximately 44–53% of prostate cancer patients had been significantly impacted by the application of 18F-Na PET/CT (99, 100). However, 18F-NaF has the same shortcoming as the other bone-seeking agents, such as 99m-technetium used in BS, leading to confusion between benign lesion and metastases, and which creates more false positives and causes lower specificity (78, 101). Choline is an essential component of the phospholipids, whose increase is associated with the high proliferation of prostate cancer cells, and both 11C-Choline and 18F-choline have been investigated in particular for the detection of relapse and metastasis of PCa (102–104). Our results demonstrated that the urinary excretion of 18F-choline was slightly higher than 11C-Choline (superiority index, 1.92 *vs.* 1.71), which may affect the interpretation of findings in the pelvis and cause lower accuracy (105, 106). Additionally, our study validated that 11C-Choline PET/CT processed the highest sensitivity–specificity and superiority index on the per-lesion basis (preformed in a situation of the lack the data of 68GA-PMSA PET/CT). The commonly used tracer 18F-FDG appears to be less useful in PCa because of the low avidity of most prostate cancer cells and urinary activity (107). Osseous metastases in PCa are typically osteoblastic (7, 108), while 18F-FDG are more sensitive in osteolytic lesions than in osteogenic lesions (72).

According to different factors (magnetic field strength, coverage, sequence used, the participation of DWI, and the number of imaging planes), MRI was grouped and directly compared to PET/CT using four commonly used radioligands (68Ga-PSMA, 18F-NaF, 11C-choline, and 18F-choline). 68Ga-PSMA PET/CT was still predominating, which was followed by the MRI equipped with better options. It is possible for MRI to detect metastasis lesion at an early stage owing to the high soft-tissue resolution (109–111). Meanwhile, numerous clinical advantages of 3.0-T MRI over 1.5-T have been demonstrated (112), which was also confirmed in our analysis (sensitivity, 0.89 *vs.* 0.82; specificity, 0.88 *vs.* 0.92; superiority index, 1.76 *vs.* 0.77). However, Woo et al. (22) reported that there existed no significant heterogeneity among different magnetic field strength. Recently, DWI is of increasing interest for the detection of primary or metastatic cancers (113, 114), benefiting from the ability to differentiate malignant from benign prostatic tissues according to different water diffusivity (68). Our analysis, according to whether DWI was involved, indicated the usefulness of DWI in the evaluation of bone metastases in PCa (sensitivity, 0.94 *vs.* 0.86; specificity, 0.93 *vs.* 0.86; superiority index, 3.63 *vs.* 1.44). A previous meta-analysis reported similar results that of improved diagnostic performance for identifying tumor foci in PCa due to the useful complement from DWI (115). Barchetti et al. (68) also revealed that conventional imaging, including T1-weighted (T1W), T1-weighted images (T2W), and short tau inversion

recovery (STIR) sequence, may improve the specificity of DWI in detecting bone metastases. Regarding the coverage of MRI, whether a dedicated axial skeleton or whole-body MRI was utilized or only covered the pelvis, the results showed a similar performance (sensitivity, 0.44 *vs.* 0.82; specificity, 0.94 *vs.* 0.93). Although it is necessary to assess the extent of extra-prostatic extension, which is an independent prognostic factor (11), the possibility of distant bone metastasis without pelvic or lumbar spinal involvement is negligible (44, 116). The motion-related signal intensity also decreases, leading to the failure in depicting lesions in ribs, sternum, and scapula (117, 118), and the increased cost and acquisition time remain obstacles to the use of WB-MRI. However, the application of wider coverage enables the detection of extra-skeletal involvement, including lymph nodes (119, 120), allows their monitoring under therapy, and helps to assess the efficacy of many new drugs in advanced PCa (121–123). Additionally, the use of more imaging planes and sequences to determine bone metastases can achieve better diagnostic accuracy, attributed to the acquisition of more information. These analyses proved that the optimization of MRI parameters could significantly improve the diagnostic ability of bone metastasis in PCa.

The highlight of the current meta-analysis was to directly compare the high-quality MRI (multisequence, DWI involved, more imaging planes) with PET/CT using four commonly used tracers. The results showed that although the 3.0-T high-quality MRI did not surpass the 68Ga-PSMA PET/CT at the patient-based level, the diagnostic ability was very close (sensitivity, 0.94 *vs.* 0.91; specificity, 0.94 *vs.* 0.96; superiority index, 4.43 *vs.* 4.56). In recent decades, MRI and PET/CT compete for the single-step whole-body technique for assessing metastases and imaging of response to treatment in solid cancers. Lecouvet et al. (44) demonstrated that MRI was a highly sensitive and specific one-step modality to diagnose bone metastases in patients with high-risk PCa and leads to changes in treatment strategy in 22% of patients. Echmann et al. (43) proved the similarly high accuracy of 11C-Choline PET/CT and 1.5-T multiparametric WB-MRI, while Conde-Moreno et al. (124) found a significantly lower ability of DWI WB-MRI for detecting bone metastasis in recurrent PCa. Although previous studies have yielded inconsistent conclusions, our results demonstrated that 3.0-T multiparametric WB-MRI has a comparable high sensitivity, specificity, and superiority index at the patient-based level, following the 68Ga-PSMA PET/CT. Nevertheless, multiparametric WB-MRI seems to fulfill the requirements of no ionizing radiation and no intravenous injection of isotopes or any contrast medium, and it is also adept at depicting all metastatic bone lesions. T1W imaging can describe the infiltration of the bone marrow, T2W imaging has also been the mainstay of MRI due to its high tissue contrast resolution, while DWI can help discover the areas of increased cellularity and highly vascularized structures (71, 125, 126). Additionally, the efforts for standardization of prostate MRI acquisition and reporting (127, 128), including MR Prostate Imaging Reporting and Data System (PI-RADS) (129) and the Prostate Diagnostic Imaging Consensus Meeting (PREDICT) (130), have further contributed to the use and emphasized the importance of interpreting MRI in the context of clinical features.

4.1 Strengths and Limitations

The arm-based model is more appealing than traditional meta-analysis and the contrast-based model since the former not only permits more straightforward interpretation of the parameters, making use of all available data and yielding shorter credible intervals but also provides more natural variance–covariance matrix structures. We adopted this model, which makes our results more convincing.

We adopted numerous statistical indicators including absolute sensitivity and specificity, relative sensitivity and specificity, diagnostic odds ratio (DOR), and superiority index to compare the diagnostic value of PET/CT and MRI with different parameter systematically. It is still a challenge to rank competing diagnostic tests especially when a test does not outperform the others on both sensitivity and specificity. DOR is commonly used in traditional meta-analysis, but it cannot distinguish between tests with high sensitivity but low specificity or vice versa. Deutsch et al. (131) introduced a superiority index to quantify the superiority of a diagnostic test. The superiority index is designed to consider the joint performance of the assessment measures. Corresponding weight is given to diagnostic tests based on their performance.

Our network meta-analysis included a total of 45 studies involving 2,843 patients and 4,263 lesions, which, to our knowledge, is the largest among similar studies. While the data in our study can be used to demonstrate numerical superiority of one imaging modality in terms of sensitivity, specificity, etc., they are even richer in that these variables can be compared across all nine imaging modalities simultaneously to assess their relative accuracy. Using this large amount of diagnostic data, we compared these imaging modalities at patient- and lesion-based level. Furthermore, a series of subgroup analyses were conducted to explore the potential influencing factors. Other subgroup analysis suggested a similar direction and magnitude of effect for studies investigating diagnostic value. We also performed sensitivity analyses by removing studies involving only one diagnostic test to confirm the stability of our results. Our results provide a comprehensive overview of the existing evidence on the imaging diagnosis of bone metastasis and have implications for clinicians, researchers, radiologists, and guideline committees. While many oncologists already consider PET/CT as the preferred method for detecting bone metastasis, however, this study provide formal quantification of the relative value of MRI.

This study is not without weakness. The first limitation was the lack of a well-accepted reference standard; all included studies used best value comparator (BVC) or predominantly BVC as the reference standard, which is according to a combination of clinical, laboratory, imaging, and follow-up studies (22). Although it is a more accurate reference standard to obtain pathological results by biopsy or surgery when determining bone metastasis, it is neither feasible nor ethical to conduct such further examinations solely. Second, patients were

categorized at diagnosis to low or high risk based on clinical characteristics, such as PSA, Gleason score at biopsy, and clinical stage. It should be taken into account the potential value of the respective modalities according to the cancer grade and stage. For example, several studies reported that PSA values is strictly correlated with 11C-choline PET/CT sensitivity (102, 132). Hence, the specific imaging strategies should be adopted for PCa with different risk (133–135). Unfortunately, due to insufficient raw data, we are unable to do more subgroup analysis. Additionally, the same limitation made a comparison based on different metastatic sites difficult to accomplish. It is challenging for both PET/CT and MRI to diagnose rib metastases due to thoracic respiratory movements (103, 136–138), and further research on the diagnostic efficacy of different imaging modalities on thoracic sites will be interesting. Third, several other imaging modalities, such as BS, X-ray, BS/TXR, CT, PET, SPECT/CT, and especially, PET/MRI, a potentially disruptive technology synergizing PET and MRI, are also commonly used to assess bone metastasis. However, the purpose of this study was only to compare PET/CT and MRI in detecting bone metastasis in PCa; hence, caution is needed in applying our results to routine clinical practice. Fourth, studies assessing cost effectiveness should be performed to assist in making the clinical decision. In our study, insufficient data prevented us from assessing the economic effectiveness and social benefits systematically.

5 CONCLUSION

This systematic review and network meta-analysis of diagnostic tests, which included 45 studies involving 2,843 patients and 4,263 lesions, indicates that 68Ga-PSMA PET/CT is recommended for the diagnosis of bone metastasis in prostate cancer patients. Where available, 3.0-T high-quality MRI approaches 68Ga-PSMA PET/CT should be performed the detection of bone metastasis.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

FL, JD, and FX conceived and designed the project. FL, JD, and FX supervised the project. FL, JD, and FX performed the review and approval of the manuscript. FL, JD, DZ, YS, CY, RW, GW, JT, QY, BW, LL, JM, and FX contributed to the design of the study, writing the protocol, screening trials, data extraction, analysis and interpretation, and writing and final approval of the report. FL, JD, and FX generated the tables and figures. FL, JD, and FX assessed the quality of included trials. FL, JD, and FX performed the literature search. FL, JD, DZ, YS, CY, RW, GW, JT, QY, BW, LL, JM, and FX drafted the manuscript. FL, JD, DZ, YS, CY, RW, GW, JT, QY, BW, LL, JM, and FX participated in revising the manuscript before submission. FL, JD, DZ, YS, CY, RW, GW, JT, QY, BW, LL, JM, and FX participated in the formal revision, including data processing, statistical analysis, generating figures and tables, and text modification. All authors had full access to the data in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. FX is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by China Scholarship Council, Nos. 201808080126 and 201706920036; the National Natural Science Foundation of China, No. 81301556; and the Key R&D program in Shandong Province, No. 2016GSF201214. The sponsors or funders had no involvements in any parts of this study.

ACKNOWLEDGMENTS

We thank Chevalier Yan, Roland Klar, and Xiangyun Cheng, from Ludwig-Maximilians-Universität München for their helpful comments on this study. We thank Shen Lao, Hongxin Wang (Technische Universität München), and Daxia Feng (Department of Nuclear Medicine, Qilu Hospital, Shandong University) who provided valuable input on the manuscript. No compensation was received for these contributions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.736654/full#supplementary-material

Supplementary Table 1 | Definitions of outcomes.

Supplementary Table 2 | Search strategy.

Supplementary Table 3 | The extraction process of diagnostic data.

Supplementary Table 4 | The detail methods of QUADAS-2 score.

Supplementary Table 5 | Implementation methods and process using arm-based model for network meta-analysis of diagnostic test accuracy data.

Supplementary Table 6 | Main characteristics of MRI in the included studies.

Supplementary Table 7 | Main characteristics of PET/CT in the included studies.

Supplementary Table 8 | Diagnostic data of the included studies on patient-based level.

Supplementary Table 9 | Diagnostic data of the included studies on lesion-based level.

Supplementary Table 10 | The results of quality assessment of the included studies.

Supplementary Table 11 | PET/CT with different tracers for detecting bone metastasis in patients with PCa by deleting studies with only one diagnostic test.

Supplementary Table 12 | PET/CT with different tracers for detecting bone metastasis in patients with PCa by deleting studies with a QUADAS-2 score of 7.

Supplementary Table 13 | PET/CT with different tracers for detecting bone metastasis in patients with PCa by deleting study with the maximum sample size.

Supplementary Table 14 | PET/CT with different tracers for detecting bone metastasis in patients with PCa by deleting study with the minimum sample size.

Supplementary Table 15 | PET/CT with different tracers for detecting bone metastasis in patients with PCa by detecting studies published before 2010.

Supplementary Table 16 | PET/CT with different tracers for detecting bone metastasis in patients with PCa based on clinical settings of cancers.

Supplementary Table 17 | PET/CT with different trances for detecting bone metastasis in patients with PCa based on numbers of patients.

Supplementary Table 18 | PET/CT with different tracers for detecting bone metastasis in patients with PCa based on ages of patients.

Supplementary Table 19 | PET/CT with different tracers for detecting bone metastasis in patients with PCa based on continents of studies.

Supplementary Table 20 | PET/CT with different tracers for detecting bone metastasis in patients with PCa based on study design.

Supplementary Table 21 | PET/CT with different tracers for detecting bone metastasis in patients with PCa based on evaluation methods of imaging analyses.

Supplementary Table 22 | PET/CT with different tracers and MRI with different field strength for detecting bone metastasis in patients with PCa.

Supplementary Table 23 | PET/CT with different tracers and MRI with different numbers of sequences for detecting bone metastasis in patients with PCa.

Supplementary Table 24 | PET/CT with different tracers and MRI with/without DWI for detecting bone metastasis in patients with PCa.

Supplementary Table 25 | PET/CT with different tracers and MRI with different numbers of imaging planes for detecting bone metastasis in patients with PCa.

Supplementary Table 26 | PET/CT with different tracers and MRI with different coverage for detecting bone metastasis in patients with PCa.

Supplementary Table 27 | PET/CT with different tracers and MRI with different field strength for detecting bone metastasis in patients with PCa by deleting studies with only one diagnostic test.

Supplementary Table 28 | PET/CT with different tracers and MRI with different numbers of sequences for detecting bone metastasis in patients with PCa by deleting studies with only one diagnostic test.

Supplementary Table 29 | PET/CT with different tracers and MRI with/without DWI for detecting bone metastasis in patients with PCa by deleting studies with only one diagnostic test.

Supplementary Table 30 | PET/CT with different tracers and MRI with different numbers of imaging planes for detecting bone metastasis in patients with PCa by deleting studies with only one diagnostic test.

Supplementary Table 31 | PET/CT with different tracers and MRI with different coverage for detecting bone metastasis in patients with PCa by deleting studies with only one diagnostic test.

Supplementary Table 32 | PET/CT with different tracers and 3.0-T high-quality MRI for detecting bone metastasis in patients with PCa by deleting studies with only one diagnostic test.

Supplementary Table 33 | PET/CT with different tracers and MRI for detecting bone metastasis in patients with PCa on lesion-based level.

Supplementary Table 34 | PET/CT with different tracers and 1.5-T high-quality MRI for detecting bone metastasis in patients with PCa on lesion-based level.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492
2. Fizazi K, Carducci M, Smith M, Damiao R, Brown J, Karsh L, et al. Denosumab Versus Zoledronic Acid for Treatment of Bone Metastases in Men With Castration-Resistant Prostate Cancer: A Randomised, Double-Blind Study. *Lancet* (2011) 377:813–22. doi: 10.1016/S0140-6736(10)62344-6
3. Norgaard M, Jensen AO, Jacobsen JB, Cetin K, Fryzek JP, Sorensen HT. Skeletal Related Events, Bone Metastasis and Survival of Prostate Cancer: A Population Based Cohort Study in Denmark (1999 to 2007). *J Urol* (2010) 184:162–7. doi: 10.1016/j.juro.2010.03.034
4. Smith MR, Saad F, Oudard S, Shore N, Fizazi K, Sieber P, et al. Denosumab and Bone Metastasis-Free Survival in Men With Nonmetastatic Castration-Resistant Prostate Cancer: Exploratory Analyses by Baseline Prostate-Specific Antigen Doubling Time. *J Clin Oncol* (2013) 31:3800–6. doi: 10.1200/JCO.2012.44.6716
5. Choi J, Raghavan M. Diagnostic Imaging and Image-Guided Therapy of Skeletal Metastases. *Cancer Control* (2012) 19:102–12. doi: 10.1177/107327481201900204
6. Beheshti M, Langsteger W, Fogelman I. Prostate Cancer: Role of SPECT and PET in Imaging Bone Metastases. *Semin Nucl Med* (2009) 39:396–407. doi: 10.1053/j.semnuclmed.2009.05.003
7. Azad GK, Taylor B, Rubello D, Colletti PM, Goh V, Cook GJ. Molecular and Functional Imaging of Bone Metastases in Breast and Prostate Cancers: An Overview. *Clin Nucl Med* (2016) 41:e44–50. doi: 10.1097/RLU.0000000000000993
8. Even-Sapir E, Metser U, Mishani E, Lievshitz G, Lerman H, Leibovitch I. The Detection of Bone Metastases in Patients With High-Risk Prostate Cancer: 99mtc-MDP Planar Bone Scintigraphy, Single- and Multi-Field-of-View SPECT, 18F-Fluoride PET, and 18F-Fluoride PET/CT. *J Nucl Med* (2006) 47:287–97.
9. Zerbib M, Zelefsky MJ, Higano CS, Carroll PR. Conventional Treatments of Localized Prostate Cancer. *Urology* (2008) 72:S25–35. doi: 10.1016/j.urology.2008.10.005
10. Berry WR. The Evolving Role of Chemotherapy in Androgen-Independent (Hormone-Refractory) Prostate Cancer. *Urology* (2005) 65:2–7. doi: 10.1016/j.urology.2005.03.080
11. Rigaud J, Tiguert R, Le Normand L, Karam G, Glemain P, Buzelin JM, et al. Prognostic Value of Bone Scan in Patients With Metastatic Prostate Cancer Treated Initially With Androgen Deprivation Therapy. *J Urol* (2002) 168:1423–6. doi: 10.1016/S0022-5347(05)64465-5
12. Sanches PG, Peters S, Rossin R, Kaijzel EL, Que I, Lowik CW, et al. Bone Metastasis Imaging With SPECT/CT/MRI: A Preclinical Toolbox for Therapy Studies. *Bone* (2015) 75:62–71. doi: 10.1016/j.bone.2015.02.002
13. Thompson I, Thrasher JB, Aus G, Burnett AL, Canby-Hagino ED, Cookson MS, et al. Guideline for the Management of Clinically Localized Prostate

Cancer: 2007 Update. *J Urol* (2007) 177:2106–31. doi: 10.1016/j.juro.2007.03.003
14. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment With Curative Intent. *Eur Urol* (2017) 71:618–29. doi: 10.1016/j.eururo.2016.08.003
15. Roodman GD. Mechanisms of Bone Metastasis. *N Engl J Med* (2004) 350:1655–64. doi: 10.1056/NEJMra030831
16. Gibril F, Doppman JL, Reynolds JC, Chen CC, Sutliff VE, Yu F, et al. Bone Metastases in Patients With Gastrinomas: A Prospective Study of Bone Scanning, Somatostatin Receptor Scanning, and Magnetic Resonance Image in Their Detection, Frequency, Location, and Effect of Their Detection on Management. *J Clin Oncol* (1998) 16:1040–53. doi: 10.1200/JCO.1998.16.3.1040
17. Wagner A, Mahrholdt H, Holly TA, Elliott MD, Regenfus M, Parker M, et al. Contrast-Enhanced MRI and Routine Single Photon Emission Computed Tomography (SPECT) Perfusion Imaging for Detection of Subendocardial Myocardial Infarcts: An Imaging Study. *Lancet* (2003) 361:374–9. doi: 10.1016/S0140-6736(03)12389-6
18. Nakanishi K, Kobayashi M, Nakaguchi K, Kyakuno M, Hashimoto N, Onishi H, et al. Whole-Body MRI for Detecting Metastatic Bone Tumor: Diagnostic Value of Diffusion-Weighted Images. *Magn Reson Med Sci* (2007) 6:147–55. doi: 10.2463/mrms.6.147
19. Gosfield E3rd, Alavi A, Kneeland B. Comparison of Radionuclide Bone Scans and Magnetic Resonance Imaging in Detecting Spinal Metastases. *J Nucl Med* (1993) 34:2191–8.
20. Engelhard K, Hollenbach HP, Wohlfart K, von Imhoff E, Fellner FA. Comparison of Whole-Body MRI With Automatic Moving Table Technique and Bone Scintigraphy for Screening for Bone Metastases in Patients With Breast Cancer. *Eur Radiol* (2004) 14:99–105. doi: 10.1007/s00330-003-1968-7
21. Vande Berg BC, Lecouvet FE, Michaux L, Ferrant A, Maldague B, Malghem J. Magnetic Resonance Imaging of the Bone Marrow in Hematological Malignancies. *Eur Radiol* (1998) 8:1335–44. doi: 10.1007/s003300050548
22. Woo S, Suh CH, Kim SY, Cho JY, Kim SH. Diagnostic Performance of Magnetic Resonance Imaging for the Detection of Bone Metastasis in Prostate Cancer: A Systematic Review and Meta-Analysis. *Eur Urol* (2018) 73:81–91. doi: 10.1016/j.eururo.2017.03.042
23. Takahara T, Imai Y, Yamashita T, Yasuda S, Nasu S, Van Cauteren M. Diffusion Weighted Whole Body Imaging With Background Body Signal Suppression (DWIBS): Technical Improvement Using Free Breathing, STIR and High Resolution 3D Display. *Radiat Med* (2004) 22:275–82.
24. Gutzeit A, Doert A, Froehlich JM, Eckhardt BP, Meili A, Scherr P, et al. Comparison of Diffusion-Weighted Whole Body MRI and Skeletal Scintigraphy for the Detection of Bone Metastases in Patients With Prostate or Breast Carcinoma. *Skeletal Radiol* (2010) 39:333–43. doi: 10.1007/s00256-009-0789-4
25. Altehoefer C, Ghanem N, Hogerle S, Moser E, Langer M. Comparative Detectability of Bone Metastases and Impact on Therapy of Magnetic Resonance Imaging and Bone Scintigraphy in Patients With Breast Cancer. *Eur J Radiol* (2001) 40:16–23. doi: 10.1016/S0720-048X(01)00313-8
26. Schirrmeister H, Guhlmann A, Kotzerke J, Santjohanser C, Kuhn T, Kreienberg R, et al. Early Detection and Accurate Description of Extent of Metastatic Bone Disease in Breast Cancer With Fluoride Ion and Positron Emission Tomography. *J Clin Oncol* (1999) 17:2381–9. doi: 10.1200/JCO.1999.17.8.2381
27. Hamaoka T, Madewell JE, Podoloff DA, Hortobagyi GN, Ueno NT. Bone Imaging in Metastatic Breast Cancer. *J Clin Oncol* (2004) 22:2942–53. doi: 10.1200/JCO.2004.08.181
28. Antoch G, Vogt FM, Freudenberg LS, Nazaradeh F, Goehde SC, Barkhausen J, et al. Whole-Body Dual-Modality PET/CT and Whole-Body MRI for Tumor Staging in Oncology. *JAMA* (2003) 290:3199–206. doi: 10.1001/jama.290.24.3199
29. Beheshti M, Mottaghy FM, Paycha F, Behrendt FFF, Van den Wyngaert T, Fogelman I, et al. (18)F-NaF PET/CT: EANM Procedure Guidelines for Bone Imaging. *Eur J Nucl Med Mol Imaging* (2015) 42:1767–77. doi: 10.1007/s00259-015-3138-y
30. Bastawrous S, Bhargava P, Behnia F, Djang DS, Haseley DR. Newer PET Application With an Old Tracer: Role of 18F-NaF Skeletal PET/CT in Oncologic Practice. *Radiographics* (2014) 34:1295–316. doi: 10.1148/rg.345130061
31. Wondergem M, van der Zant FM, van der Ploeg T, Knol RJ. A Literature Review of 18F-Fluoride PET/CT and 18F-Choline or 11C-Choline PET/CT for Detection of Bone Metastases in Patients With Prostate Cancer. *Nucl Med Commun* (2013) 34:935–45. doi: 10.1097/MNM.0b013e328364918a
32. Heidenreich A, Bastian PJ, Bellmunt J, Bolla M, Joniau S, van der Kwast T, et al. EAU Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment With Curative Intent-Update 2013. *Eur Urol* (2014) 65:124–37. doi: 10.1016/j.eururo.2013.09.046
33. Afshar-Oromieh A, Holland-Letz T, Giesel FL, Kratochwil C, Mier W, Haufe S, et al. Diagnostic Performance of (68)Ga-PSMA-11 (HBED-CC) PET/CT in Patients With Recurrent Prostate Cancer: Evaluation in 1007 Patients. *Eur J Nucl Med Mol Imaging* (2017) 44:1258–68. doi: 10.1007/s00259-017-3711-7
34. Eiber M, Maurer T, Souvatzoglou M, Beer AJ, Ruffani A, Haller B, et al. Evaluation of Hybrid (6)(8)Ga-PSMA Ligand PET/CT in 248 Patients With Biochemical Recurrence After Radical Prostatectomy. *J Nucl Med* (2015) 56:668–74. doi: 10.2967/jnumed.115.154153
35. Calais J, Czernin J, Cao M, Kishan AU, Hegde JV, Shaverdian N, et al. (68) Ga-PSMA-11 PET/CT Mapping of Prostate Cancer Biochemical Recurrence After Radical Prostatectomy in 270 Patients With a PSA Level of Less Than 1.0 Ng/Ml: Impact on Salvage Radiotherapy Planning. *J Nucl Med* (2018) 59:230–7. doi: 10.2967/jnumed.117.201749
36. Janssen JC, Meissner S, Woythal N, Prasad V, Brenner W, Diederichs G, et al. Comparison of Hybrid (68)Ga-PSMA-PET/CT and (99m)Tc-DPD-SPECT/CT for the Detection of Bone Metastases in Prostate Cancer Patients: Additional Value of Morphologic Information From Low Dose CT. *Eur Radiol* (2018) 28:610–9. doi: 10.1007/s00330-017-4994-6
37. Zhou J, Gou Z, Wu R, Yuan Y, Yu G, Zhao Y. Comparison of PSMA-PET/CT, Choline-PET/CT, NaF-PET/CT, MRI, and Bone Scintigraphy in the Diagnosis of Bone Metastases in Patients With Prostate Cancer: A Systematic Review and Meta-Analysis. *Skeletal Radiol* (2019) 48(12):1915–24. doi: 10.1007/s00256-019-03230-z
38. Nyaga VN, Aerts M, Arbyn M. ANOVA Model for Network Meta-Analysis of Diagnostic Test Accuracy Data. *Stat Methods Med Res* (2018) 27:1766–84. doi: 10.1177/0962280216669182
39. Shepard C, Genevro R. *Citymakers : The Culture and Craft of Practical Urbanism*. New York: The Monacelli Press, In association with The Architecural League of New York (2017).
40. Owen RK, Cooper NJ, Quinn TJ, Lees R, Sutton AJ. Network Meta-Analysis of Diagnostic Test Accuracy Studies Identifies and Ranks the Optimal Diagnostic Tests and Thresholds for Health Care Policy and Decision-Making. *J Clin Epidemiol* (2018) 99:64–74. doi: 10.1016/j.jclinepi.2018.03.005
41. Menten J, Lesaffre E. A General Framework for Comparative Bayesian Meta-Analysis of Diagnostic Studies. *BMC Med Res Method* (2015) 15:70. doi: 10.1186/s12874-015-0061-7
42. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PMthe P-DTAG, et al. Preferred Reporting Items for a Systematic Review and Meta-Analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. *JAMA* (2018) 319:388–96. doi: 10.1001/jama.2017.19163
43. Eschmann SM, Pfannenberg AC, Rieger A, Aschoff P, Muller M, Paulsen F, et al. Comparison of 11C-Choline-PET/CT and Whole Body-MRI for Staging of Prostate Cancer. *Nuklearmedizin* (2007) 46:161–8; quiz N147-168. doi: 10.1160/nukmed-0075
44. Lecouvet FE, Geukens D, Stainier A, Jamar F, Jamart J, d'Othee BJ, et al. Magnetic Resonance Imaging of the Axial Skeleton for Detecting Bone Metastases in Patients With High-Risk Prostate Cancer: Diagnostic and Cost-Effectiveness and Comparison With Current Detection Strategies. *J Clin Oncol* (2007) 25:3281–7. doi: 10.1200/JCO.2006.09.2940
45. Nemeth AJ, Henson JW, Mullins ME, Gonzalez RG, Schaefer PW. Improved Detection of Skull Metastasis With Diffusion-Weighted MR Imaging. *AJNR Am J Neuroradiol* (2007) 28:1088–92. doi: 10.3174/ajnr.A0501
46. Beheshti M, Vali R, Waldenberger P, Fitz F, Nader M, Loidl W, et al. Detection of Bone Metastases in Patients With Prostate Cancer by 18F Fluorocholine and 18F Fluoride PET-CT: A Comparative Study. *Eur J Nucl Med Mol Imaging* (2008) 35:1766–74. doi: 10.1007/s00259-008-0788-z

47. Venkitaraman R, Cook GJ, Dearnaley DP, Parker CC, Huddart RA, Khoo V, et al. Does Magnetic Resonance Imaging of the Spine Have a Role in the Staging of Prostate Cancer? *Clin Oncol (R Coll Radiol)* (2009) 21:39–42. doi: 10.1016/j.clon.2008.09.006
48. Venkitaraman R, Cook GJ, Dearnaley DP, Parker CC, Khoo V, Eeles R, et al. Whole-Body Magnetic Resonance Imaging in the Detection of Skeletal Metastases in Patients With Prostate Cancer. *J Med Imaging Radiat Oncol* (2009) 53:241–7. doi: 10.1111/j.1754-9485.2009.02070.x
49. Beheshti M, Vali R, Waldenberger P, Fitz F, Nader M, Hammer J, et al. The Use of F-18 Choline PET in the Assessment of Bone Metastases in Prostate Cancer: Correlation With Morphological Changes on CT. *Mol Imaging Biol* (2010) 12:98–107. doi: 10.1007/s11307-009-0239-7
50. Fuccio C, Castellucci P, Schiavina R, Santi I, Allegri V, Pettinato V, et al. Role of 11C-Choline PET/CT in the Restaging of Prostate Cancer Patients Showing a Single Lesion on Bone Scintigraphy. *Ann Nucl Med* (2010) 24:485–92. doi: 10.1007/s12149-010-0390-x
51. Iagaru A, Mittra E, Dick DW, Gambhir SS. Prospective Evaluation of (99m) Tc MDP Scintigraphy, (18)F NaF PET/CT, and (18)F FDG PET/CT for Detection of Skeletal Metastases. *Mol Imaging Biol* (2011) 14:252–9. doi: 10.1007/s11307-011-0486-2
52. Langsteger W, Balogova S, Huchet V, Beheshti M, Paycha F, Egrot C, et al. Fluorocholine (18F) and Sodium Fluoride (18F) PET/CT in the Detection of Prostate Cancer: Prospective Comparison of Diagnostic Performance Determined by Masked Reading. *Q J Nucl Med Mol Imaging* (2011) 55:448–57.
53. Bortot DC, Amorim BJ, Oki GC, Gapski SB, Santos AO, Lima MC, et al. 18f-Fluoride PET/CT Is Highly Effective for Excluding Bone Metastases Even in Patients With Equivocal Bone Scintigraphy. *Eur J Nucl Med Mol Imaging* (2012) 39:1730–6. doi: 10.1007/s00259-012-2195-8
54. Jadvar H, Desai B, Ji L, Conti PS, Dorff TB, Groshen SG, et al. Prospective Evaluation of 18F-NaF and 18F-FDG PET/CT in Detection of Occult Metastatic Disease in Biochemical Recurrence of Prostate Cancer. *Clin Nucl Med* (2012) 37:637–43. doi: 10.1097/RLU.0b013e318252d829
55. Lecouvet FE, El Mouedden J, Collette L, Coche E, Danse E, Jamar F, et al. Can Whole-Body Magnetic Resonance Imaging With Diffusion-Weighted Imaging Replace Tc 99m Bone Scanning and Computed Tomography for Single-Step Detection of Metastases in Patients With High-Risk Prostate Cancer? *Eur Urol* (2012) 62:68–75. doi: 10.1016/j.eururo.2012.02.020
56. Mosavi F, Johansson S, Sandberg DT, Turesson I, Sorensen J, Ahlstrom H. Whole-Body Diffusion-Weighted MRI Compared With (18)F-NaF PET/CT for Detection of Bone Metastases in Patients With High-Risk Prostate Carcinoma. *AJR Am J Roentgenol* (2012) 199:1114–20. doi: 10.2214/AJR.11.8351
57. Picchio M, Spinapolice EG, Fallanca F, Crivellaro C, Giovacchini G, Gianolli L, et al. [11c]Choline PET/CT Detection of Bone Metastases in Patients With PSA Progression After Primary Treatment for Prostate Cancer: Comparison With Bone Scintigraphy. *Eur J Nucl Med Mol Imaging* (2012) 39:13–26. doi: 10.1007/s00259-011-1920-z
58. Takesh M, Odat Allh K, Adams S, Zechmann C. Diagnostic Role of (18)F-FECH-PET/CT Compared With Bone Scan in Evaluating the Prostate Cancer Patients Referring With Biochemical Recurrence. *ISRN Oncol* (2012) 2012:815234. doi: 10.5402/2012/815234
59. Damle NA, Bal C, Bandopadhyaya GP, Kumar L, Kumar P, Malhotra A, et al. The Role of 18F-Fluoride PET-CT in the Detection of Bone Metastases in Patients With Breast, Lung and Prostate Carcinoma: A Comparison With FDG PET/CT and 99mtc-MDP Bone Scan. *Jpn J Radiol* (2013) 31:262–9. doi: 10.1007/s11604-013-0179-7
60. Kitajima K, Murphy RC, Nathan MA, Froemming AT, Hagen CE, Takahashi N, et al. Detection of Recurrent Prostate Cancer After Radical Prostatectomy: Comparison of 11C-Choline PET/CT With Pelvic Multiparametric MR Imaging With Endorectal Coil. *J Nucl Med* (2014) 55:223–32. doi: 10.2967/jnumed.113.123018
61. Pasoglou V, Larbi A, Collette L, Annet L, Jamar F, Machiels JP, et al. One-Step TNM Staging of High-Risk Prostate Cancer Using Magnetic Resonance Imaging (MRI): Toward an Upfront Simplified "All-in-One" Imaging Approach? *Prostate* (2014) 74:469–77. doi: 10.1002/pros.22764
62. Piccardo A, Paparo F, Piccazzo R, Naseri M, Ricci P, Marziano A, et al. Value of Fused 18F-Choline-PET/MRI to Evaluate Prostate Cancer Relapse in Patients Showing Biochemical Recurrence After EBRT: Preliminary Results. *BioMed Res Int* (2014) 2014:103718. doi: 10.1155/2014/103718
63. Poulsen MH, Petersen H, Hoilund-Carlsen PF, Jakobsen JS, Gerke O, Karstoft J, et al. Spine Metastases in Prostate Cancer: Comparison of Technetium-99m-MDP Whole-Body Bone Scintigraphy, [(18) F]choline Positron Emission Tomography(PET)/computed Tomography (CT) and [(18) F]NaF PET/CT. *BJU Int* (2014) 114:818–23. doi: 10.1111/bju.12599
64. Evangelista L, Cimitan M, Zattoni F, Guttilla A, Zattoni F, Saladini G. Comparison Between Conventional Imaging (Abdominal-Pelvic Computed Tomography and Bone Scan) and [(18)F]choline Positron Emission Tomography/Computed Tomography Imaging for the Initial Staging of Patients With Intermediate- Tohigh-Risk Prostate Cancer: A Retrospective Analysis. *Scand J Urol* (2015) 49:345–53. doi: 10.3109/21681805.2015.1005665
65. Pasoglou V, Michoux N, Peeters F, Larbi A, Tombal B, Selleslagh T, et al. Whole-Body 3D T1-Weighted MR Imaging in Patients With Prostate Cancer: Feasibility and Evaluation in Screening for Metastatic Disease. *Radiology* (2015) 275:155–66. doi: 10.1148/radiol.14141242
66. Sampath SC, Sampath SC, Mosci C, Lutz AM, Willmann JK, Mittra ES, et al. Detection of Osseous Metastasis by 18F-NaF/18F-FDG PET/CT Versus CT Alone. *Clin Nucl Med* (2015) 40:e173–177. doi: 10.1097/RLU.0000000000000560
67. Wieder H, Beer AJ, Holzapfel K, Henninger M, Maurer T, Schwarzenboeck S, et al. 11C-Choline PET/CT and Whole-Body MRI Including Diffusion-Weighted Imaging for Patients With Recurrent Prostate Cancer. *Oncotarget* (2017) 8:66516–27. doi: 10.18632/oncotarget.16227
68. Barchetti F, Stagnitti A, Megna V, Al Ansari N, Marini A, Musio D, et al. Unenhanced Whole-Body MRI Versus PET-CT for the Detection of Prostate Cancer Metastases After Primary Treatment. *Eur Rev Med Pharmacol Sci* (2016) 20:3770–6.
69. Conde-Moreno AJ, Herrando-Parreno G, Muelas-Soria R, Ferrer-Rebolleda J, Broseta-Torres R, Cozar-Santiago MP, et al. Whole-Body Diffusion-Weighted Magnetic Resonance Imaging (WB-DW-MRI) vs Choline-Positron Emission Tomography-Computed Tomography (Choline-PET/CT) for Selecting Treatments in Recurrent Prostate Cancer. *Clin Transl Oncol* (2016) 19:553–61. doi: 10.1007/s12094-016-1563-4
70. Nanni C, Zanoni L, Pultrone C, Schiavina R, Brunocilla E, Lodi F, et al. (18) F-FACBC (Anti1-Amino-3-(18)F-Fluorocyclobutane-1-Carboxylic Acid) Versus (11)C-Choline PET/CT in Prostate Cancer Relapse: Results of a Prospective Trial. *Eur J Nucl Med Mol Imaging* (2016) 43:1601–10. doi: 10.1007/s00259-016-3329-1
71. Woo S, Kim SY, Kim SH, Cho JY. JOURNAL CLUB: Identification of Bone Metastasis With Routine Prostate MRI: A Study of Patients With Newly Diagnosed Prostate Cancer. *AJR Am J Roentgenol* (2016) 206:1156–63. doi: 10.2214/AJR.15.15761
72. Yi C, Yu D, Shi X, Zhang X, Luo G, He Q, et al. The Combination of 13N-Ammonia and 18F-FDG Whole-Body PET/CT on the Same Day for Diagnosis of Advanced Prostate Cancer. *Nucl Med Commun* (2016) 37:239–46. doi: 10.1097/MNM.0000000000000444
73. Fonager RF, Zacho HD, Langkilde NC, Fledelius J, Ejlersen JA, Haarmark C, et al. Diagnostic Test Accuracy Study of (18)F-Sodium Fluoride PET/CT, (99m)Tc-Labelled Diphosphonate SPECT/CT, and Planar Bone Scintigraphy for Diagnosis of Bone Metastases in Newly Diagnosed, High-Risk Prostate Cancer. *Am J Nucl Med Mol Imaging* (2017) 7:218–27.
74. Huysse W, Lecouvet F, Castellucci P, Ost P, Lambrecht V, Artigas C, et al. Prospective Comparison of F-18 Choline PET/CT Scan Versus Axial MRI for Detecting Bone Metastasis in Biochemically Relapsed Prostate Cancer Patients. *Diagnostics (Basel)* (2017) 7(4). doi: 10.3390/diagnostics7040056
75. Janssen JC, Meissner S, Woythal N, Prasad V, Brenner W, Diederichs G, et al. Comparison of Hybrid (68)Ga-PSMA-PET/CT and (99m)Tc-DPD-SPECT/CT for the Detection of Bone Metastases in Prostate Cancer Patients: Additional Value of Morphologic Information From Low Dose CT. *Eur Radiol* (2017) 28:610–9. doi: 10.1007/s00330-017-4994-6
76. Kitajima K, Fukushima K, Yamamoto S, Kato T, Odawara S, Takaki H, et al. Diagnostic Performance of (11)C-Choline PET/CT and Bone Scintigraphy in the Detection of Bone Metastases in Patients With Prostate Cancer. *Nagoya J Med Sci* (2017) 79:387–99. doi: 10.18999/nagjms.79.3.387

77. Vargas HA, Schor-Bardach R, Long N, Kirzner AN, Cunningham JD, Goldman DA, et al. Prostate Cancer Bone Metastases on Staging Prostate MRI: Prevalence and Clinical Features Associated With Their Diagnosis. *Abdom Radiol (NY)* (2017) 42:271–7. doi: 10.1007/s00261-016-0851-3
78. Wondergem M, van der Zant FM, Knol RJJ, Burgers AMG, Bos SD, de Jong IJ, et al. (99m)Tc-HDP Bone Scintigraphy and (18)F-Sodiumfluoride PET/CT in Primary Staging of Patients With Prostate Cancer. *World J Urol* (2018) 36:27–34. doi: 10.1007/s00345-017-2096-3
79. Dyrberg E, Hendel HW, Huynh THV, Klausen TW, Logager VB, Madsen C, et al. (68)Ga-PSMA-PET/CT in Comparison With (18)F-Fluoride-PET/CT and Whole-Body MRI for the Detection of Bone Metastases in Patients With Prostate Cancer: A Prospective Diagnostic Accuracy Study. *Eur Radiol* (2018) 29:1221–30. doi: 10.1007/s00330-018-5682-x
80. Kawanaka Y, Kitajima KD, Yamamoto S, Nakanishi Y, Yamada Y, Hashimoto T, et al. Comparison of 11C-Choline Positron Emission Tomography/Computed Tomography (PET/CT) and Conventional Imaging for Detection of Recurrent Prostate Cancer. *Cureus* (2018) 10: e2966. doi: 10.7759/cureus.2966
81. Larbi A, Omoumi P, Pasoglou V, Michoux N, Triqueneaux P, Tombal B, et al. Whole-Body MRI to Assess Bone Involvement in Prostate Cancer and Multiple Myeloma: Comparison of the Diagnostic Accuracies of the T1, Short Tau Inversion Recovery (STIR), and High B-Values Diffusion-Weighted Imaging (DWI) Sequences. *Eur Radiol* (2018) 29:4503–13. doi: 10.1007/s00330-018-5796-1
82. Lengana T, Lawal IO, Boshomane TG, Popoola GO, Mokoala KMG, Moshokoa E, et al. (68)Ga-PSMA PET/CT Replacing Bone Scan in the Initial Staging of Skeletal Metastasis in Prostate Cancer: A Fait Accompli? *Clin Genitourin Cancer* (2018) 16:392–401. doi: 10.1016/j.clgc.2018.07.009
83. Zacho HD, Nielsen JB, Afshar-Oromieh A, Haberkorn U, deSouza N, De Paepe K, et al. Prospective Comparison of (68)Ga-PSMA PET/CT, (18)F-Sodium Fluoride PET/CT and Diffusion Weighted-MRI at for the Detection of Bone Metastases in Biochemically Recurrent Prostate Cancer. *Eur J Nucl Med Mol Imaging* (2018) 45:1884–97. doi: 10.1007/s00259-018-4058-4
84. Chen B, Wei P, Macapinlac HA, Lu Y. Comparison of 18F-Fluciclovine PET/CT and 99mtc-MDP Bone Scan in Detection of Bone Metastasis in Prostate Cancer. *Nucl Med Commun* (2019) 40:940–6. doi: 10.1097/MNM.0000000000001051
85. Johnston EW, Latifoltojar A, Sidhu HS, Ramachandran N, Sokolska M, Bainbridge A, et al. Multiparametric Whole-Body 3.0-T MRI in Newly Diagnosed Intermediate- and High-Risk Prostate Cancer: Diagnostic Accuracy and Interobserver Agreement for Nodal and Metastatic Staging. *Eur Radiol* (2019) 29:3159–69. doi: 10.1007/s00330-018-5813-4
86. Uslu-Besli L, Sager S, Akgun E, Asa S, Sahin OE, Demirdag C, et al. Comparison of Ga-68 PSMA Positron Emission Tomography/Computerized Tomography With Tc-99m MDP Bone Scan in Prostate Cancer Patients. *Turk J Med Sci* (2019) 49:301–10. doi: 10.3906/sag-1807-4
87. Lardinois D, Weder W, Hany TF, Kamel EM, Korom S, Seifert B, et al. Staging of Non-Small-Cell Lung Cancer With Integrated Positron-Emission Tomography and Computed Tomography. *N Engl J Med* (2003) 348:2500–7. doi: 10.1056/NEJMoa022136
88. van Kruchten M, de Vries EGE, Brown M, de Vries EFJ, Glaudemans A, Dierckx R, et al. PET Imaging of Oestrogen Receptors in Patients With Breast Cancer. *Lancet Oncol* (2013) 14:e465–75. doi: 10.1016/S1470-2045(13)70292-4
89. Pyka T, Okamoto S, Dahlbender M, Tauber R, Retz M, Heck M, et al. Comparison of Bone Scintigraphy and (68)Ga-PSMA PET for Skeletal Staging in Prostate Cancer. *Eur J Nucl Med Mol Imaging* (2016) 43:2114–21. doi: 10.1007/s00259-016-3435-0
90. Eder M, Eisenhut M, Babich J, Haberkorn U. PSMA as a Target for Radiolabelled Small Molecules. *Eur J Nucl Med Mol Imaging* (2013) 40:819–23. doi: 10.1007/s00259-013-2374-2
91. Afshar-Oromieh A, Zechmann CM, Malcher A, Eder M, Eisenhut M, Linhart HG, et al. Comparison of PET Imaging With a (68)Ga-Labelled PSMA Ligand and (18)F-Choline-Based PET/CT for the Diagnosis of Recurrent Prostate Cancer. *Eur J Nucl Med Mol Imaging* (2014) 41:11–20. doi: 10.1007/s00259-013-2525-5
92. von Eyben FE, Picchio M, von Eyben R, Rhee H, Bauman G. (68)Ga-Labeled Prostate-Specific Membrane Antigen Ligand Positron Emission Tomography/Computed Tomography for Prostate Cancer: A Systematic Review and Meta-Analysis. *Eur Urol Focus* (2018) 4:686–93. doi: 10.1016/j.euf.2016.11.002
93. Giesel FL, Will L, Lawal I, Lengana T, Kratochwil C, Vorster M, et al. Intraindividual Comparison of (18)F-PSMA-1007 and (18)F-DCFPyL PET/CT in the Prospective Evaluation of Patients With Newly Diagnosed Prostate Carcinoma: A Pilot Study. *J Nucl Med* (2018) 59:1076–80. doi: 10.2967/jnumed.117.204669
94. Carroll PR, Parsons JK, Andriole G, Bahnson RR, Castle EP, Catalona WJ, et al. NCCN Guidelines Insights: Prostate Cancer Early Detection, Version 2.2016. *J Natl Compr Canc Netw* (2016) 14:509–19. doi: 10.6004/jnccn.2016.0060
95. Mohler JL, Armstrong AJ, Bahnson RR, D'Amico AV, Davis BJ, Eastham JA, et al. Prostate Cancer, Version 1.2016. *J Natl Compr Canc Netw* (2016) 14:19–30. doi: 10.6004/jnccn.2016.0004
96. Cook GJ, Fogelman I. The Role of Positron Emission Tomography in Skeletal Disease. *Semin Nucl Med* (2001) 31:50–61. doi: 10.1053/snuc.2001.18746
97. Cook GJ, Fogelman I. The Role of Positron Emission Tomography in the Management of Bone Metastases. *Cancer* (2000) 88:2927–33. doi: 10.1002/1097-0142(20000615)88:12+<2927::AID-CNCR8>3.0.CO;2-V
98. Hawkins RA, Choi Y, Huang SC, Hoh CK, Dahlbom M, Schiepers C, et al. Evaluation of the Skeletal Kinetics of Fluorine-18-Fluoride Ion With PET. *J Nucl Med* (1992) 33:633–42.
99. Hillner BE, Siegel BA, Hanna L, Duan F, Quinn B, Shields AF. 18F-Fluoride PET Used for Treatment Monitoring of Systemic Cancer Therapy: Results From the National Oncologic PET Registry. *J Nucl Med* (2015) 56:222–8. doi: 10.2967/jnumed.114.150391
100. Hillner BE, Siegel BA, Hanna L, Duan F, Shields AF, Coleman RE. Impact of 18F-Fluoride PET in Patients With Known Prostate Cancer: Initial Results From the National Oncologic PET Registry. *J Nucl Med* (2014) 55:574–81. doi: 10.2967/jnumed.113.130005
101. Liu FY, Chang JT, Wang HM, Liao CT, Kang CJ, Ng SH, et al. [18F] Fluorodeoxyglucose Positron Emission Tomography Is More Sensitive Than Skeletal Scintigraphy for Detecting Bone Metastasis in Endemic Nasopharyngeal Carcinoma at Initial Staging. *J Clin Oncol* (2006) 24:599–604. doi: 10.1200/JCO.2005.03.8760
102. Krause BJ, Souvatzoglou M, Tuncel M, Herrmann K, Buck AK, Praus C, et al. The Detection Rate of [11C]Choline-PET/CT Depends on the Serum PSA-Value in Patients With Biochemical Recurrence of Prostate Cancer. *Eur J Nucl Med Mol Imaging* (2008) 35:18–23. doi: 10.1007/s00259-007-0581-4
103. Picchio M, Messa C, Landoni C, Gianolli L, Sironi S, Brioschi M, et al. Value of [11C]Choline-Positron Emission Tomography for Re-Staging Prostate Cancer: A Comparison With [18F]Fluorodeoxyglucose-Positron Emission Tomography. *J Urol* (2003) 169:1337–40. doi: 10.1097/01.ju.0000056901.95996.43
104. de Jong IJ, Pruim J, Elsinga PH, Vaalburg W, Mensink HJ. 11C-Choline Positron Emission Tomography for the Evaluation After Treatment of Localized Prostate Cancer. *Eur Urol* (2003) 44:32–8; discussion 38-39. doi: 10.1016/s0302-2838(03)00207-0
105. Cimitan M, Bortolus R, Morassut S, Canzonieri V, Garbeglio A, Baresic T, et al. [18F]Fluorocholine PET/CT Imaging for the Detection of Recurrent Prostate Cancer at PSA Relapse: Experience in 100 Consecutive Patients. *Eur J Nucl Med Mol Imaging* (2006) 33:1387–98. doi: 10.1007/s00259-006-0150-2
106. Kwee SA, Wei H, Sesterhenn I, Yun D, Coel MN. Localization of Primary Prostate Cancer With Dual-Phase 18F-Fluorocholine PET. *J Nucl Med* (2006) 47:262–9.
107. Emonds KM, Swinnen JV, Mortelmans L, Mottaghy FM. Molecular Imaging of Prostate Cancer. *Methods* (2009) 48:193–9. doi: 10.1016/j.ymeth.2009.03.021
108. Vessella RL, Corey E. Targeting Factors Involved in Bone Remodeling as Treatment Strategies in Prostate Cancer Bone Metastasis. *Clin Cancer Res* (2006) 12:6285s–90s. doi: 10.1158/1078-0432.CCR-06-0813
109. Moulopoulos LA, Dimopoulos MA. Magnetic Resonance Imaging of the Bone Marrow in Hematologic Malignancies. *Blood* (1997) 90:2127–47. doi: 10.1182/blood.V90.6.2127
110. Costelloe CM, Rohren EM, Madewell JE, Hamaoka T, Theriault RL, Yu TK, et al. Imaging Bone Metastases in Breast Cancer: Techniques and

Recommendations for Diagnosis. *Lancet Oncol* (2009) 10:606–14. doi: 10.1016/S1470-2045(09)70088-9
111. Constantine G, Shan K, Flamm SD, Sivananthan MU. Role of MRI in Clinical Cardiology. *Lancet* (2004) 363:2162–71. doi: 10.1016/S0140-6736(04)16509-4
112. Willinek WA, Schild HH. Clinical Advantages of 3.0 T MRI Over 1.5 T. *Eur J Radiol* (2008) 65:2–14. doi: 10.1016/j.ejrad.2007.11.006
113. Koh DM, Collins DJ. Diffusion-Weighted MRI in the Body: Applications and Challenges in Oncology. *AJR Am J Roentgenol* (2007) 188:1622–35. doi: 10.2214/AJR.06.1403
114. Li S, Sun F, Jin ZY, Xue HD, Li ML. Whole-Body Diffusion-Weighted Imaging: Technical Improvement and Preliminary Results. *J Magn Reson Imaging* (2007) 26:1139–44. doi: 10.1002/jmri.21074
115. Tan CH, Wei W, Johnson V, Kundra V. Diffusion-Weighted MRI in the Detection of Prostate Cancer: Meta-Analysis. *AJR Am J Roentgenol* (2012) 199:822–9. doi: 10.2214/AJR.11.7805
116. Carmona Echeverria LM, Drudge-Coates L, Wilkins CJ, Muir GH. Bone Scan Is of Doubtful Value as a First Staging Test in the Primary Presentation of Prostate Cancer. *ISRN Oncol* (2012) 2012:585017. doi: 10.5402/2012/585017
117. Padhani AR, Koh DM, Collins DJ. Whole-Body Diffusion-Weighted MR Imaging in Cancer: Current Status and Research Directions. *Radiology* (2011) 261:700–18. doi: 10.1148/radiol.11110474
118. Vogler JB3rd, Murphy WA. Bone Marrow Imaging. *Radiology* (1988) 168:679–93. doi: 10.1148/radiology.168.3.3043546
119. Low RN. Abdominal MRI Advances in the Detection of Liver Tumours and Characterisation. *Lancet Oncol* (2007) 8:525–35. doi: 10.1016/S1470-2045(07)70170-5
120. Komori T, Narabayashi I, Matsumura K, Matsuki M, Akagi H, Ogura Y, et al. 2-[Fluorine-18]-Fluoro-2-Deoxy-D-Glucose Positron Emission Tomography/Computed Tomography Versus Whole-Body Diffusion-Weighted MRI for Detection of Malignant Lesions: Initial Experience. *Ann Nucl Med* (2007) 21:209–15. doi: 10.1007/s12149-007-0010-6
121. Lecouvet FE, Simon M, Tombal B, Jamart J, Vande Berg BC, Simoni P. Whole-Body MRI (WB-MRI) Versus Axial Skeleton MRI (AS-MRI) to Detect and Measure Bone Metastases in Prostate Cancer (PCa). *Eur Radiol* (2010) 20:2973–82. doi: 10.1007/s00330-010-1879-3
122. Lecouvet FE, Lhommel R, Pasoglou V, Larbi A, Jamar F, Tombal B. Novel Imaging Techniques Reshape the Landscape in High-Risk Prostate Cancers. *Curr Opin Urol* (2013) 23:323–30. doi: 10.1097/MOU.0b013e328361d451
123. Tombal B, Rezazadeh A, Therasse P, Van Cangh PJ, Vande Berg B, Lecouvet FE. Magnetic Resonance Imaging of the Axial Skeleton Enables Objective Measurement of Tumor Response on Prostate Cancer Bone Metastases. *Prostate* (2005) 65:178–87. doi: 10.1002/pros.20280
124. Conde-Moreno AJ, Herrando-Parreno G, Muelas-Soria R, Ferrer-Rebolleda J, Broseta-Torres R, Cozar-Santiago MP, et al. Whole-Body Diffusion-Weighted Magnetic Resonance Imaging (WB-DW-MRI) vs Choline-Positron Emission Tomography-Computed Tomography (Choline-PET/CT) for Selecting Treatments in Recurrent Prostate Cancer. *Clin Transl Oncol* (2017) 19:553–61. doi: 10.1007/s12094-016-1563-4
125. Mullerad M, Hricak H, Kuroiwa K, Pucar D, Chen HN, Kattan MW, et al. Comparison of Endorectal Magnetic Resonance Imaging, Guided Prostate Biopsy and Digital Rectal Examination in the Preoperative Anatomical Localization of Prostate Cancer. *J Urol* (2005) 174:2158–63. doi: 10.1097/01.ju.0000181224.95276.82
126. Kajihara H, Hayashida Y, Murakami R, Katahira K, Nishimura R, Hamada Y, et al. Usefulness of Diffusion-Weighted Imaging in the Localization of Prostate Cancer. *Int J Radiat Oncol Biol Phys* (2009) 74:399–403. doi: 10.1016/j.ijrobp.2008.08.017
127. Vargas HA, Hotker AM, Goldman DA, Moskowitz CS, Gondo T, Matsumoto K, et al. Updated Prostate Imaging Reporting and Data System (PIRADS V2) Recommendations for the Detection of Clinically Significant Prostate Cancer Using Multiparametric MRI: Critical Evaluation Using Whole-Mount Pathology as Standard of Reference. *Eur Radiol* (2016) 26:1606–12. doi: 10.1007/s00330-015-4015-6
128. Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, et al. ESUR Prostate MR Guidelines 2012. *Eur Radiol* (2012) 22:746–57. doi: 10.1007/s00330-011-2377-y
129. Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol* (2016) 69:16–40. doi: 10.1016/j.eururo.2015.08.052
130. Dickinson L, Ahmed HU, Allen C, Barentsz JO, Carey B, Futterer JJ, et al. Scoring Systems Used for the Interpretation and Reporting of Multiparametric MRI for Prostate Cancer Detection, Localization, and Characterization: Could Standardization Lead to Improved Utilization of Imaging Within the Diagnostic Pathway? *J Magn Reson Imaging* (2013) 37:48–58. doi: 10.1002/jmri.23689
131. Deutsch R, Mindt MR, Xu R, Cherner M, Grant I, Group atH. Quantifying Relative Superiority Among Many Binary-Valued Diagnostic Tests in the Presence of a Gold Standard. *J Data Sci* (2009) 7:161–77. doi: 10.6339/JDS.2009.07(2).450
132. Castellucci P, Fuccio C, Nanni C, Santi I, Rizzello A, Lodi F, et al. Influence of Trigger PSA and PSA Kinetics on 11C-Choline PET/CT Detection Rate in Patients With Biochemical Relapse After Radical Prostatectomy. *J Nucl Med* (2009) 50:1394–400. doi: 10.2967/jnumed.108.061507
133. Han M, Partin AW, Zahurak M, Piantadosi S, Epstein JI, Walsh PC. Biochemical (Prostate Specific Antigen) Recurrence Probability Following Radical Prostatectomy for Clinically Localized Prostate Cancer. *J Urol* (2003) 169:517–23. doi: 10.1016/S0022-5347(05)63946-8
134. Hricak H, Schoder H, Pucar D, Lis E, Eberhardt SC, Onyebuchi CN, et al. Advances in Imaging in the Postoperative Patient With a Rising Prostate-Specific Antigen Level. *Semin Oncol* (2003) 30:616–34. doi: 10.1016/S0093-7754(03)00359-2
135. Schoder H, Larson SM. Positron Emission Tomography for Prostate, Bladder, and Renal Cancer. *Semin Nucl Med* (2004) 34:274–92. doi: 10.1053/j.semnuclmed.2004.06.004
136. Nakai T, Okuyama C, Kubota T, Yamada K, Ushijima Y, Taniike K, et al. Pitfalls of FDG-PET for the Diagnosis of Osteoblastic Bone Metastases in Patients With Breast Cancer. *Eur J Nucl Med Mol Imaging* (2005) 32:1253–8. doi: 10.1007/s00259-005-1842-8
137. Du Y, Cullum I, Illidge TM, Ell PJ. Fusion of Metabolic Function and Morphology: Sequential [18F]Fluorodeoxyglucose Positron-Emission Tomography/Computed Tomography Studies Yield New Insights Into the Natural History of Bone Metastases in Breast Cancer. *J Clin Oncol* (2007) 25:3440–7. doi: 10.1200/JCO.2007.11.2854
138. Serpa Neto A, Tobias-Machado M, Esteves MA, Senra MD, Wroclawski ML, Fonseca FL, et al. A Systematic Review and Meta-Analysis of Bone Metabolism in Prostate Adenocarcinoma. *BMC Urol* (2010) 10:9. doi: 10.1186/1471-2490-10-9

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Liu, Dong, Shen, Yun, Wang, Wang, Tan, Wang, Yao, Wang, Li, Mi, Zhou and Xiong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 20 October 2021
doi: 10.3389/fonc.2021.744724

Novel Non-Invasive Radiomic Signature on CT Scans Predicts Response to Platinum-Based Chemotherapy and Is Prognostic of Overall Survival in Small Cell Lung Cancer

Prantesh Jain[1†], Mohammadhadi Khorrami[2†], Amit Gupta[3], Prabhakar Rajiah[4], Kaustav Bera[2], Vidya Sankar Viswanathan[2], Pingfu Fu[5], Afshin Dowlati[1] and Anant Madabhushi[2,6]*

[1] Department of Hematology and Oncology, University Hospitals Cleveland Medical Center, Cleveland, OH, United States, [2] Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH, United States, [3] Department of Radiology, University Hospitals Cleveland Medical Center, Cleveland, OH, United States, [4] Department of Radiology, Mayo Clinic Minnesota, Rochester, MN, United States, [5] Department of Population and Quantitative Health Sciences, Case Western Reserve University (CWRU), Cleveland, OH, United States, [6] Louis Stokes Cleveland Veterans Administration Medical Center, Cleveland, OH, United States

OPEN ACCESS

Edited by:
Guang Yang,
Imperial College London,
United Kingdom

Reviewed by:
Jianxin Wang,
Central South University, China
Yuming Jiang,
Stanford University, United States
Xuefeng Xia,
Geneplus Beijing Institute, China

***Correspondence:**
Anant Madabhushi
axm788@case.edu

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 20 July 2021
Accepted: 29 September 2021
Published: 20 October 2021

Citation:
Jain P, Khorrami M, Gupta A, Rajiah P, Bera K, Viswanathan VS, Fu P, Dowlati A and Madabhushi A (2021) Novel Non-Invasive Radiomic Signature on CT Scans Predicts Response to Platinum-Based Chemotherapy and Is Prognostic of Overall Survival in Small Cell Lung Cancer. Front. Oncol. 11:744724. doi: 10.3389/fonc.2021.744724

Background: Small cell lung cancer (SCLC) is an aggressive malignancy characterized by initial chemosensitivity followed by resistance and rapid progression. Presently, there are no predictive biomarkers that can accurately guide the use of systemic therapy in SCLC patients. This study explores the role of radiomic features from both within and around the tumor lesion on pretreatment CT scans to a) prognosticate overall survival (OS) and b) predict response to chemotherapy.

Methods: One hundred fifty-three SCLC patients who had received chemotherapy were included. Lung tumors were contoured by an expert reader. The patients were divided randomly into approximately equally sized training (S^{tr} = 77) and test sets (S^{te} = 76). Textural descriptors were extracted from the nodule (intratumoral) and parenchymal regions surrounding the nodule (peritumoral). The clinical endpoints of this study were OS, progression-free survival (PFS), and best objective response to chemotherapy. Patients with complete or partial response were defined as "responders," and those with stable or progression of disease were defined as "non-responders." The radiomic risk score (RRS) was generated by using the least absolute shrinkage and selection operator (LASSO) with the Cox regression model. Patients were classified into the high-risk or low-risk groups based on the median of RRS. Association of the radiomic signature with OS was evaluated on S^{tr} and then tested on S^{te}. The features identified by LASSO were then used to train a linear discriminant analysis (LDA) classifier (M^{Rad}) to predict response to chemotherapy. A prognostic nomogram ($N^{Rad+Clin}$) was also developed on S^{tr} by combining clinical and prognostic radiomic features and validated on S^{te}. The Kaplan–

Meier survival analysis and log-rank statistical tests were performed to assess the discriminative ability of the features. The discrimination performance of the $N^{Rad+Clin}$ was assessed by Harrell's C-index. To estimate the clinical utility of the nomogram, decision curve analysis (DCA) was performed by calculating the net benefits for a range of threshold probabilities in predicting which high-risk patients should receive more aggressive treatment as compared with the low-risk patients.

Results: A univariable Cox regression analysis indicated that RRS was significantly associated with OS in S^{tr} (HR: 1.53; 95% CI, [1.1–2.2; p = 0.021]; C-index = 0.72) and S^{te} (HR: 1.4, [1.1–1.82], p = 0.0127; C-index = 0.69). The RRS was also significantly associated with PFS in S^{tr} (HR: 1.89, [1.4–4.61], p = 0.047; C-index = 0.7) and S^{te} (HR: 1.641, [1.1–2.77], p = 0.04; C-index = 0.67). M^{Rad} was able to predict response to chemotherapy with an area under the receiver operating characteristic curve (AUC) of 0.76 ± 0.03 within S^{tr} and 0.72 within S^{te}. Predictors, including the RRS, gender, age, stage, and smoking status, were used in the prognostic nomogram. The discrimination ability of the $N^{Rad+Clin}$ model on S^{tr} and S^{te} was C-index [95% CI]: 0.68 [0.66–0.71] and 0.67 [0.63–0.69], respectively. DCA indicated that the $N^{Rad+Clin}$ model was clinically useful.

Conclusions: Radiomic features extracted within and around the lung tumor on CT images were both prognostic of OS and predictive of response to chemotherapy in SCLC patients.

Keywords: radiomics, computed tomography, small cell lung cancer (SCLC), chemotherapy, overall survival, progression-free survival

1 INTRODUCTION

Lung cancer remains the leading cause of cancer-related mortality worldwide. Traditionally, primary lung cancers have been divided into non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC), with majority of them (85%) being of the NSCLC subtype (1). SCLC is an aggressive neuroendocrine (NE) malignancy that accounts for 13% to 15% of all lung cancers and is strongly associated with smoking. Besides Tumor-Node-Metastasis (TNM) staging, SCLC can be classified into limited-stage disease (LS-SCLC—tumor confined to single radiation port with or without loco-regional adenopathy) and extensive-stage disease (ES-SCLC—tumor not confined to single radiation port) (2). This two-stage system has therapeutic and prognostic implications with 5-year relative survival rates of 31% for LS-SCLC and 2% for ES-SCLC (3).

The current treatment modalities for SCLC include systemic chemotherapy, immunotherapy, thoracic radiation, and prophylactic cranial irradiation, depending on the tumor stage. While SCLC is very responsive to initial treatment, most patients develop early resistance to conventional therapies, show rapid progression, and relapse with decreased sensitivity to further pharmacological treatment (4, 5); and fewer than 10% patients enjoy long-term survival (6–8). There has been little improvement in outcome over the past few decades with the addition of immunotherapy to chemotherapy in an unselected patient population, and thus, platinum-based chemotherapy remains the mainstay of systemic treatment (9). While younger age, good performance status (PS), normal creatinine level, and normal lactate dehydrogenase (LDH) are favorable prognostic factors (10), there are no reliable predictive biomarkers that can identify SCLC patients who will benefit from cytotoxic chemotherapy or patients at a high risk of relapse, although recently it has been shown that patients with wild-type retinoblastoma gene are chemotherapy-refractory (11).

CT is a routinely used clinical diagnostic tool for tumor staging and monitoring treatment response. The common presentation of SCLC on a CT scan is a centrally located large parenchymal mass or a mediastinal mass involving at least one hilum. In recent years, computational imaging approaches such as "radiomics" (12) can provide a more detailed feature-based characterization of the disease than possible by visual examination. Radiomics-based biomarkers have been shown to be prognostically useful in different types of therapies for various cancers, including NSCLC (13–16). However, there are no data regarding the role of radiomics in predicting response to chemotherapy or prognosticating outcome for SCLC.

In this study, we sought to identify chemotherapy response as well as prognostic biomarkers for overall survival (OS) and progression-free survival (PFS) in SCLC patients treated with chemotherapy by interrogating the tumor and tumor microenvironment on CT imaging. We hypothesize that quantitative subvisual phenotypic differences in SCLC tumors on CT imaging can be characterized non-invasively to develop predictive and prognostic biomarker signatures to improve decision support in SCLC treatment. With the use of a cohort of 153 patients with SCLC, 77 were used for training the classifier and constructing a radiomic risk score (RRS), whereas the

remaining 76 were used for the test set. The choice of an equal split of training and test sets has been employed in previous approaches involving radiomics (14, 15, 17). In addition, a novel prognostic nomogram was constructed by integrating RRS and clinical biomarkers, and its performance for predicting high-risk patients was evaluated by the decision curve analysis (DCA) model.

2 MATERIALS AND METHODS

2.1 Datasets and Patient Selection

This study was conducted in full accordance with the Health Insurance Portability and Accountability Act (HIPAA) regulations after approval from the Institutional Review Board (IRB) at Case Western Reserve University (Cleveland, OH), and the IRB waived the requirements for informed consent of all patients because of the retrospective, non-interventional, and non-therapeutic nature of this study. A total of 305 consecutive patients with LS-SCLC or ES-SCLC treated with platinum-based chemotherapy from April 2004 to March 2018 in University Hospitals Cleveland Medical Center (UHCMC) were identified. All patients that met the following criteria were included: a) availability of pathological confirmation of SCLC, b) presence of diagnostic thoracic CT scan in axial view, and c) presence of a solitary pulmonary nodule/mass. To this cohort of 180 patients, the exclusion criteria were applied to remove scans with CT artifacts and poor image quality not suitable for feature extraction. The final cohort had 153 patients. The training set imbalance has been discussed extensively as an issue in training machine classifiers. To reduce the possible impact of imbalanced data on the machine classifier, the patients were randomly divided into an equal number for training set (S^{tr}) that consisted of 77 patients (age: 39–87 years) and a test set (St^{e}) of 76 patients (age: 47–90 years). Patient selection and overall experimental design for this study are shown in **Figure 1**.

The non-contrast CT scan images were acquired from all participants at baseline before initiation of chemotherapy from Siemens, GE Medical Systems, Philips, or Toshiba machine utilizing a tube voltage of 100 to 120 kVp. The median time between cancer diagnosis and CT scan acquisition was 10.4 days, and that between diagnosis and first-line therapy was 33.5 days (range: 2–220 days). The dataset also had images acquired from multiple reconstruction kernels. It is known that the different slice thickness and reconstruction kernels of the CT image acquisition affect radiomic feature expression and measurements (18). Therefore, precaution was taken to sample cases from both S^{tr} and S^{te}. The slice thickness ranged from 1 to 5 mm (mean = 2.82 mm, SD = 0.71 mm), and the pixel sizes ranged from 0.42 × 0.42 mm to 0.97 × 0.97 mm with an average size of 0.73 × 0.73 mm. **Table 1** shows CT acquisition parameters involving all scanner types, slice thicknesses, kVp, and reconstruction kernels for both the training and test sets.

2.2 Clinical Endpoints

The primary endpoints of this study were response to chemotherapy, OS, and PFS. The OS was measured from the date of diagnosis to the date of death and censored at the date of last follow-up for survivors. PFS was defined as the length of time during the treatment that a patient survives with cancer but without evidence of disease progression or death, whichever occurred earlier. The objective response to chemotherapy was evaluated based on RECIST 1.1. The target lesions were evaluated to assess response, and the following definitions were used: complete response (CR), i.e., the disappearance of all the lesions; partial response (PR), i.e., ≥30% decrease in the sum of

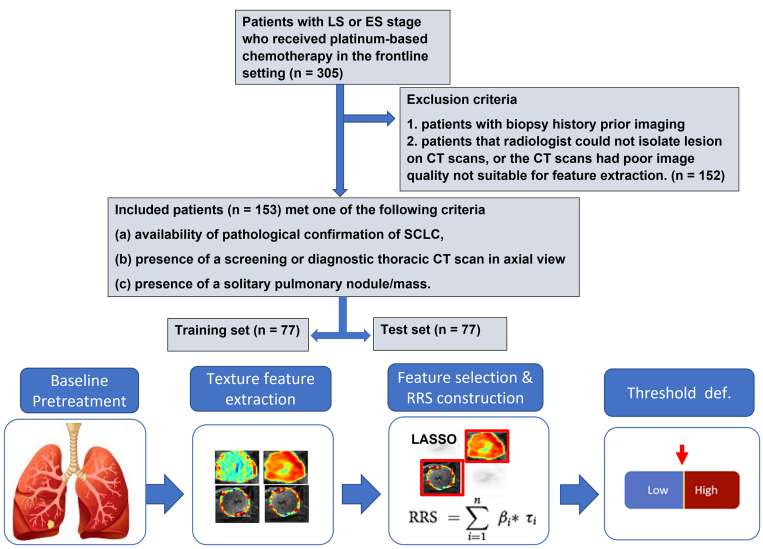


FIGURE 1 | Patient selection and overall experimental design for this study.

TABLE 1 | CT image acquisition parameter distribution over the training set and test set.

	Training set	Test set
Patients	77	76
Scanner		
GE	6	4
Siemens	43	45
Philips	23	21
Toshiba	5	6
Thickness		
1	1	2
1.5	2	0
2	10	12
2.5	0	2
3	10	11
3.2	2	0
5	42	49
Kernel		
GE Med		
Standard	6	4
Siemens		
B20f	1	2
B30f	0	0
B35f	5	7
B35s	0	0
B40f	13	10
B41f	0	0
B50f	10	7
B60f	3	1
B70f	14	18
Philips		
A	0	0
B	0	0
C	11	8
D	12	13
E	0	0
L	0	0
Toshiba		
FC02	0	0
FC08	5	6
FC10	0	0
kVp		
100	34	31
110	23	27
120	20	18

the longest diameters of target lesions compared with baseline; progressive disease (PD), i.e., at least 20% increase in the sum of the longest diameters of target lesions with an absolute increase of ≥5 mm; and stable disease (SD), i.e., neither PR nor PD (19). For our study, patients with CR or PR were classified as "responders," and those with SD or PD were classified as "non-responders."

2.3 Tumor Segmentation

The tumor was identified by a two-board-certified cardiothoracic radiologist (PR) with 20 years of experience and Reader 2 (VV, a physician with 2 years of experience in cardiothoracic radiology research), blinded to each other; and the region of interest (ROI) was manually segmented across all the 2D CT slices of the nodule *via* a hand-annotation tool in axial view on 3D-Slicer software. The radiologist was blinded to clinical data and given the option to vary the window and level setting within the software to efficiently annotate the nodule. The segmented nodules were used to extract the intratumoral texture and shape features. The intratumoral mask was then dilated out to a 20-mm peritumoral radius. The definition of the extent of the peritumoral zone was based on a previous publication (20), which showed that patients with SCLC exhibit a tendency to develop peritumoral edema to a region of 20-mm extent around the tumor. Peritumoral masks were inspected manually and adjusted to include only lung tissue when masks extended into chest wall soft tissues.

2.4 Radiomic Feature Extraction

For a given tumor, 2D manual segmentations were assessed in a slice-by-slice basis to pick all representative slices that had the tumor. Since there was an unequal slice spacing (ranging from 1 to 5 mm), 2D radiomic analysis provides for a more consistent approach as compared with 3D analysis, which in turn would have been impacted by the non-uniform z-axis slice spacing. From these slices, 2D texture features were extracted from the intratumoral and peritumoral regions on a per-pixel basis. Within the intratumoral and peritumoral regions, a total of 99 2D radiomic texture descriptors were extracted. These descriptors consisted of features that were selected to capture textural structure of intra- and peritumoral regions. In this study, we extracted 13 Haralick features from gray-level co-occurrence matrix (GLCM) that can extract textural pattern and show variation in tumor microarchitecture, heterogeneity, and local appearance of nodules. In addition, we extracted 25 Laws, 25 Laws–Laplacian, and 48 Gabor features from intra- and peritumoral regions. Laws and Laws–Laplacian are filter-based descriptors that capture combinations of five textural patterns, such as levels (L), edges (E), spots (S), waves (W), and/or ripples (R). The Gabor filter bank was used to capture texture responses at six different spatial frequencies (f = 0, 2, 4, 8, 16, or 32) within the image at eight directional orientations ($\theta = 0, \pi/8, \pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4$, and $7\pi/8$). A total of 24 shape features were also automatically extracted from the annotated nodules and investigated in the study. Shape features are used to describe the 3D geometrical composition of the segmented nodule structure including size (volume and diameter) and shape measures (sphericity, compactness, and radial distance). First-order statistics (mean, median, standard deviation, skewness, and kurtosis) for each descriptor were computed within the tumor and peritumoral region, resulting in 495 statistical features per region. Each feature was normalized to a mean of zero and standard deviation of 1 across patients before feature selection. All shape and texture features were extracted using an in-house software that was developed on MATLAB 2018 platform (MathWorks Inc., Natick, MA), an approach very similar to implementations of radiomic features like CERR and PyRadiomics (21, 22). To mitigate the effect of different acquisition parameters, only features that were stable in the context of the test–retest RIDER lung CT dataset (23) were selected for more analysis.

2.5 Statistical Analysis

2.5.1 Feature Selection

To avoid overfitting due to high complexity of features, least absolute shrinkage and selection operator (LASSO) Cox regression model was used to identify the most prognostic features to OS from highly stable features in S^{tr}. LASSO iteratively shrinks the feature coefficient estimates toward zero and involves identification of an optimal tuning parameter lambda (λ) *via* a 100 cross-validation setup. The process that is only run on the training set allows for identification of only those features with non-zero coefficients.

2.5.2 Radiomic Risk Score Generation

The RRS signature was built based on linear combination of non-zero coefficients-selected features. The association of the RRS with OS was first assessed in S^{tr} and then validated in S^{te} by using the Kaplan–Meier survival analysis. According to the rad-score threshold identified by the median of RRS, patients were classified into the high-risk or low-risk categories.

2.5.3 Prognostic Analysis

OS and PFS for univariate analysis were estimated by the Kaplan–Meier method. The multivariable survival analysis with RRS and clinicopathologic biomarker was also employed. A DCA was also used to determine the clinical utility of the RRS in predicting OS by evaluating the net benefit of high-risk patients receiving treatment at different threshold probabilities (24). Net benefit was defined as the summation of benefits minus loss results (false-positive findings) weighted by a factor related to the relative harm of not identifying a high-risk patient who might have low OS versus the relative harm of subjecting a lower-risk patient to more aggressive therapy, when the more intense therapy was not needed.

In addition, a nomogram ($N^{Rad+Clin}$) was constructed as an individualized OS prediction model in S^{tr}. Predictors, including the RRS and clinical parameters like gender, age, stage, and smoking status, were added to the prognostic nomogram. The prognostic performance of $N^{Rad+Clin}$ was estimated in S^{tr} and then evaluated in S^{te}. The consistency between the actual and predicted probabilities survival was evaluated by a calibration plot.

2.5.4 Classification

A linear discriminant analysis (LDA) classifier (M^{Rad}) was trained with the same set of features identified by LASSO. Within S^{tr}, the LDA classifier was trained over 100 iterations of threefold cross-validation. The classifier was finally locked down and then evaluated for prediction of response on S^{te}. The ability to identify response was primarily assessed by an area under the receiver operating characteristic (ROC) curve (AUC).

2.5.5 Implementation and Statistical Test

The "glmnet" package in R was used for executing the LASSO algorithm. Analysis of OS and PFS outcomes was utilized with survival methods, primarily Cox regression. The model was included indicators for the categorical low- and high-risk OS from RRS along with clinical factors. The survival probabilities of patients classified as low or high risk based on RRS was estimated and illustrated by the Kaplan–Meier curves, and relative HRs with 95% CIs were calculated using the Wald test and the G-rho rank test, in R, version 3.6.3. Survival differences were compared by the log-rank test. A multivariable Cox PH model was used to investigate the independent prognostic effect of the RRS model in comparison with clinical variables, and the likelihood ratio test was applied to confirm the independent prognostic effect of each variable.

The final selection of the model for the nomogram was conducted using a backward step-down selection process based on the Akaike information criterion (25), while internal validation was conducted through 1,000 bootstrap resamples. Harrell's concordance index (C-index), a quantitative measurement of the performance of the nomogram, was used to assess the discriminative ability of the model in survival analysis. The nomogram and calibration plots were constructed using the "rms" and "SvyNom" packages.

A heatmap dendrogram was used to display unsupervised hierarchal clustering using the radiomic texture features. A consensus clustering approach was also used to determine the number and affiliation of possible clusters within all the patient studies. Nodules belonging to different clusters may have minimal correlation, while nodules within a cluster are likely to have a high intra-class correlation.

Accuracy, precision, sensitivity, specificity, and kappa agreement between the predicted response of classifier and actual response were also computed at the optimal operating point of the ROC curve, while the operating point was defined as the threshold that maximized overall accuracy.

Differences between clinical categories were assessed using Fisher's exact test, while a two-sided Wilcoxon test was used for continuous variables. For continuous variables, the interquartile range (IQR), a measure of statistical dispersion, was also reported. IQR was calculated as the difference between the 75th and 25th percentiles.

3 RESULTS

3.1 Patient Analysis

A total of 153 patients with SCLC were included for analysis with a median age of 66 years (34–90), with 72.8% men, median OS of 9.37, and median PFS of 8.35 months. All patients with LS or ES were treated with platinum-based chemotherapy. Of these, 86% patients had extensive-stage disease, and 14% were of limited stage. No statistically significant difference was found in baseline clinical characteristics between responders and non-responders, chemotherapy agents (carboplatin and cisplatin), or time between CT acquisition and first-line chemotherapy. Chemotherapy response was achieved in 100 (66%) patients, who were labeled as responders (R), and the remaining 53 (34%) were labeled as non-responders (NR). All clinical characteristics are listed in **Table 2**.

TABLE 2 | Demographics and clinical characteristics for patients, categorized by training and test sets.

Characteristics		All patients (n = 153)	Training set (n = 77)	Test set (n = 76)	p-Value
Sex	Male	78 (52%)	43 (56%)	35 (46%)	0.26
	Female	75 (48%)	34 (44%)	41 (54%)	
Age	Median (IQR)	66 (13)	6712.5)	66 (14)	0.89
	[95% CI]	[64.97–68.16]	[64.00–68.67]	[64.62–68.98]	
Race	White	98 (82%)	48 (62%)	50 (66%)	0.73
	Black	55 (18%)	29 (38%)	26 (34%)	
Smoking	Never	6 (4%)	3 (4%)	3 (4%)	1.00
	Former/current	147 (96%)	74 (96%)	73 (96%)	
Stage	Limited stage	21 (14%)	10 (13%)	11 (14%)	0.81
	Extensive stage	132 (86%)	67 (87%)	65 (86%)	
Chemotherapy Agents	Carboplatin	64 (42%)	30 (39%)	34 (45%)	0.51
	Cisplatin	89 (58%)	47 (61%)	42 (55%)	
Median OS	Months (IQR)	9.37 (12.73)	8.27 (12.55)	10.18 (12.48)	0.12
	[95% CI]	[11.63–15.77]	[10.10–16.40]	[11.46–16.85]	
Median PFS	Months (IQR)	8.35 (11.8)	7.57 (12.2)	9.23 (11.9)	0.19
	[95% CI]	[11.2–16.84]	[9.68–17.36]	[10.37–18.68]	

3.2 Radiomic Features From Pretreatment CT Scans Were Associated With Overall Survival in Small Cell Lung Cancer

The median OS was 8.27 months (IQR = 12.55, [10.10–16.40]). A univariable Cox regression analysis in training set identified that OS was not significantly different for gender (male vs. female; HR: 0.78 [0.48–1.25]; p = 0.3; C-index = 0.52 [0.45–0.58]) or race (HR: 0.86 [0.53–1.4]; p = 0.54; C-index = 0.51 [0.44–0.57]) but was significantly different for clinical stage (LS vs. ES; HR: 1.4 [1.3–1.7]; p = 0.0002; C-index = 0.58 [0.52–0.6]). Especially, patients with brain metastasis demonstrated poorer survival, but in our dataset, this finding was not significantly associated with OS (HR: 0.52; [0.27–0.98]; p = 0.069; C-index = 0.57). **Figures 2A–C** illustrate the Kaplan–Meier curves for different clinical factors.

The radiomic score was calculated as a linear combination of the six selected features weighted by their respective coefficients. These features were identified as entropy of intratumoral Haralick feature, median of intratumoral Laws texture feature, peritumoral laws texture feature, intratumoral low-frequency Gabor feature, and peritumoral high-frequency Gabor feature. The optimum cutoff value (the median) for the RRS was found to be 0.17, and patients were stratified into high- and low-risk groups based on this value (**Figure 2D**). A univariable Cox regression analysis developed using textural features indicated that RRS was significantly associated with OS in S^{tr} (HR: 1.53 [1.1–2.2]; p = 0.021; C-index = 0.72 [0.63–0.81]) and S^{te} (HR: 1.4 [1.1–1.82]; p = 0.0127; C-index = 0.69 [0.60–0.77]). The corresponding Kaplan–Meier survival curves showed a significant difference in OS between patients with low and high RRS (p-value <0.05). The Kaplan–Meier survival curves for S^{tr} and S^{te} are shown in **Figures 2E, F**, respectively.

A multivariable Cox regression analysis identified the RRS and cancer staging (limited or extensive stage) as two major risk factors in OS for patients in S^{tr} (RRS: HR = 2.1 [1.53–2.85], p = 0.0076; clinical stage: HR = 1.66 [1.01, 2.7], p = 0.048; race: HR = 0.37 [0.12, 1.1], p = 0.071; and age: HR = 1.04 [0.99–1.09], p = 0.071; C-index = 0.75 [0.68–0.81]) and S^{te} (RRS: HR = 1.9 [1.23–2.2], p = 0.0012; clinical stage: HR = 1.61 [1.2–2.17], p = 0.041; race: HR = 0.86 [0.52–1.42], p = 0.56; and age: HR = 1.01 [0.99–1.03], p = 0.22; C-index = 0.71 [0.64–0.77]).

3.3 Integrating Clinical Parameters With Radiomic Features From Pretreatment CT Scans for Prediction of Overall Survival in Small Cell Lung Cancer

The C-index for $N^{Rad+Clin}$ to predict OS in S^{tr} was 0.68 [0.66–0.69]. In S^{te}, the C-index was 0.67 [0.63–0.68], a value that was greater than that of the conventional clinical-based model (0.54 [0.52–0.56]; p = 0.0024) and RRS alone (0.62 [0.60–0.64]; p = 0.041). The nomogram and the corresponding calibration curve are illustrated in **Figures 3A, B**, respectively. The calibration plot demonstrates an optimal consistency between $N^{Rad+Clin}$ predicted and actual observed OS in S^{te}.

Moreover, **Figure 4** shows a DCA for three models (clinical model, radiomic model, and integrated Rad+Clin model). As can be seen, the Rad+Clin model had the highest net benefit in predicting which high-risk patients should receive more aggressive treatment as compared with the low-risk patients. The Rad+Clin model yielded a greater net benefit compared with the "treat-all" or "treat-none" strategies.

3.4 Radiomic Features From Pretreatment CT Scans Were Associated With Progression-Free Survival in Small Cell Lung Cancer

The median PFS was 7.57 months (IQR = 12.2, [9.68–17.36]). A univariable Cox regression analysis in training set indicated that RRS generated for OS was also significantly associated with PFS in S^{tr} (HR = 1.89 [1.4–4.61], p = 0.047; C-index = 0.7 [0.61–0.78]) and S^{te} (HR = 1.64 [1.1–2.77], p = 0.04; C-index = 0.67 [0.60–0.74]). A multivariable Cox regression analysis identified the RRS and cancer staging as two major risk factors in PFS for patients in S^{tr} (RRS: HR = 1.8 [1.48–2.23], p = 0.0082; clinical stage: HR = 1.35 [1.08, 1.78], p = 0.033; C-index = 0.72 [0.64–0.79]) and S^{te} (RRS: HR = 1.61 [1.13, 1.96], p = 0.0046; clinical stage: HR = 1.37 [1.1, 2.01], p = 0.039; C-index = 0.70 [0.62–0.78]).

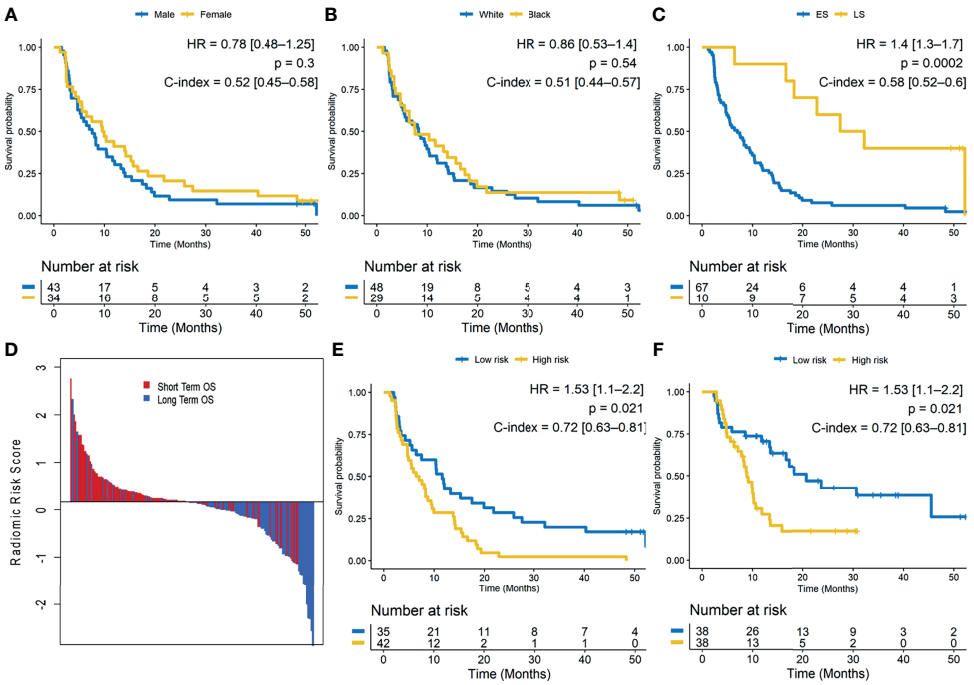


FIGURE 2 | **(A–C)** Kaplan–Meier survival curves for gender, race, and clinical stage on the training set. ES, extensive stage; LS, limited stage. **(D)** Waterfall plot of the length of overall survival (OS) based on radiomic risk score (RRS); higher risk score is associated with lower OS. **(E)** Kaplan–Meier survival curves based on the training set and **(F)** test set. A significant association of the radiomic risk score with the OS is shown in the training set and test set.

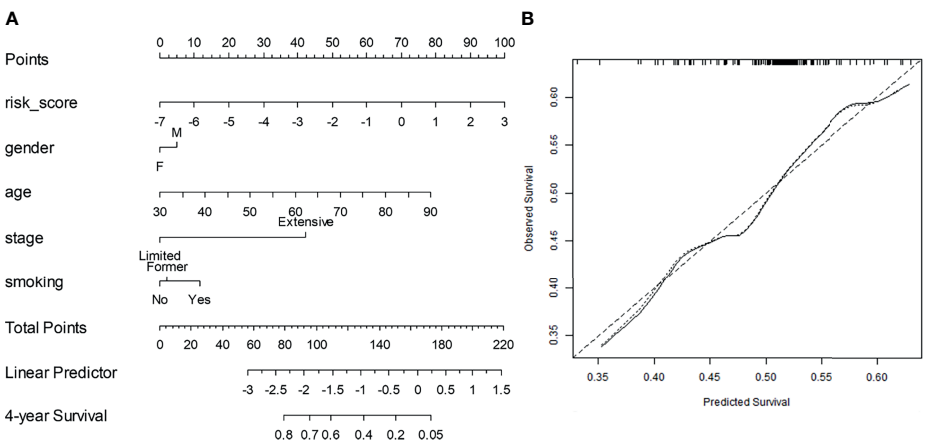


FIGURE 3 | **(A)** Integrated clinical and radiomic nomogram ($N^{Rad+Clin}$) for small cell lung cancer (SCLC) patients treated with systemic chemotherapy estimating the probability of surviving for 4 years. Instructions for reading the nomogram: locate the risk score on the risk score axis. Draw a line straight up to the Points axis to determine how many points toward the predicted probability of a 4-year overall survival (OS) that the patient receives for radiomic risk score level. Repeat this process for the other predictors, each time drawing a line straight up to the Points axis. Sum the points achieved for each predictor and locate this sum on the Total Points axis. Draw a line straight down to the 4-year Survival axis to determine the patient's probability of surviving for 4 years. Variables with the greatest discriminatory value are those with the widest point range in the nomogram. Sample data from one patient is shown (tan arrows and ovals). **(B)** Calibration curve for 4-year survival. The x-axis shows the nomogram predicted probability, while the y-axis gives the actual 4-year survival as estimated by the Kaplan–Meier method. The dotted line represents an ideal agreement between actual and predicted probabilities of 4-year survival. The solid line represents $N^{Rad+Clin}$ nomogram, and the vertical bars represent 95% CIs. Dots correspond to apparent predictive accuracy.

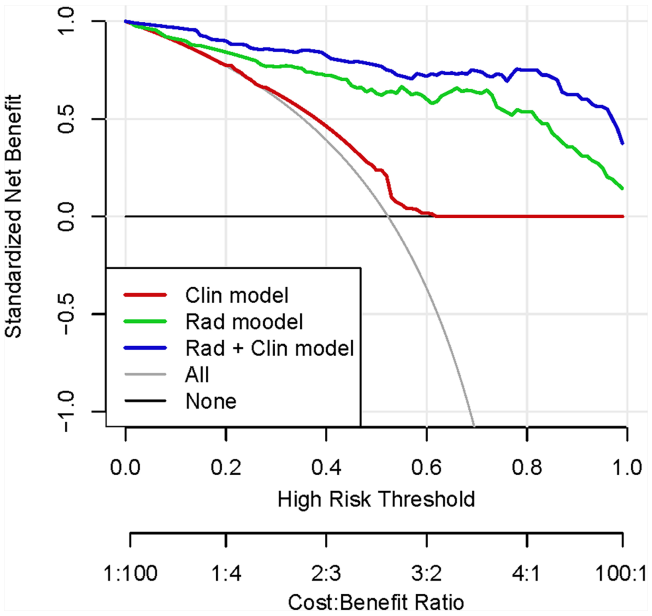


FIGURE 4 | Decision curve analysis (DCA) for each model (clinical model, radiomic model, and integrated Rad+Clin model). The integrated Rad+Clin model had the highest net benefit in predicting which high-risk patients should receive more aggressive treatment, as compared with radiomic model, a clinical model, and simple strategies such as to treat all patients or no patients. This analysis was performed across the full range of threshold probabilities at which a patient would be selected to undergo follow-up imaging.

3.5 Radiomic Features From Pretreatment CT Scans Predict Response to Chemotherapy in Small Cell Lung Cancer

Figure 5A shows distinct response associated clusters obtained *via* consensus clustering performed on a combination of features that were found to be discriminating between responders and non-responders. The two clusters had a preponderance of responders (67%) and non-responders (75%). The radiomic heatmap in **Figure 5B** shows an association between the radiomic features and chemotherapy response for all the patients included in this study.

M^{Rad} was able to predict response to chemotherapy with an AUC of 0.76 [0.75–0.79] within S^{tr} and a corresponding AUC of 0.72, an accuracy of 0.74, a precision of 0.62 ($p < 0.05$), a specificity of 0.85, a sensitivity of 0.75, and kappa agreement of 0.56 within S^{te}.

Figures 6A, B illustrate the discriminability of the intratumoral Haralick entropy and peritumoral Gabor feature for representative non-responder and responder SCLC patients before chemotherapy on the baseline CT scan. There appears to be a higher textural pattern disorder or heterogeneity within and around lesions on CT images before treatment in non-responder patients as compared with responders. This trend is also reflected in the box-and-whisker plots of the Haralick entropy and Gabor texture, illustrated in **Figure 6C**.

4 DISCUSSION

SCLC is an aggressive NE tumor of the lung, which arises from bronchial mucosa and shares many morphologic features of NE tumors (26). Chemotherapy remains the backbone of systemic treatment in SCLC. Even though most patients respond to initial treatment, relapse is inevitable, and a subset of patients are chemoresistant. Historically, SCLC was considered as a homogenous disease, and there is now preclinical evidence of inter-tumor heterogeneity with distinct molecular subtypes. Also, recent studies have demonstrated that switching of subtype within the tumor could be a reason for chemotherapeutic resistance (27–30). Currently, there are no clinically validated predictive biomarkers to select a subpopulation of patients with primary chemoresistance and/or early recurrence.

In this study, we presented novel, computer-extracted, quantitative texture features from within and around the tumor lesion from baseline CT scans that are predictive of chemoresistance and prognostic of OS, independent of clinicopathologic factors. Following additional multi-site validation, these non-invasive radiomic biomarkers can be used to predict chemoresistance upfront and orient patients to clinical trials targeting unique therapeutic vulnerabilities to improve clinical outcomes. The primary goal of this study was to determine whether a durable prediction of response to chemotherapy is possible by using radiomic texture patterns within and outside the SCLC tumor on baseline CT scans. We developed a novel RRS derived from computerized texture features from pretreatment CT that was shown to be prognostic of OS and PFS and predictive to chemotherapy response. Additionally, we presented a novel prognostic nomogram ($N^{Rad+Clin}$) that combines radiomic features with clinicopathologic parameters to predict OS.

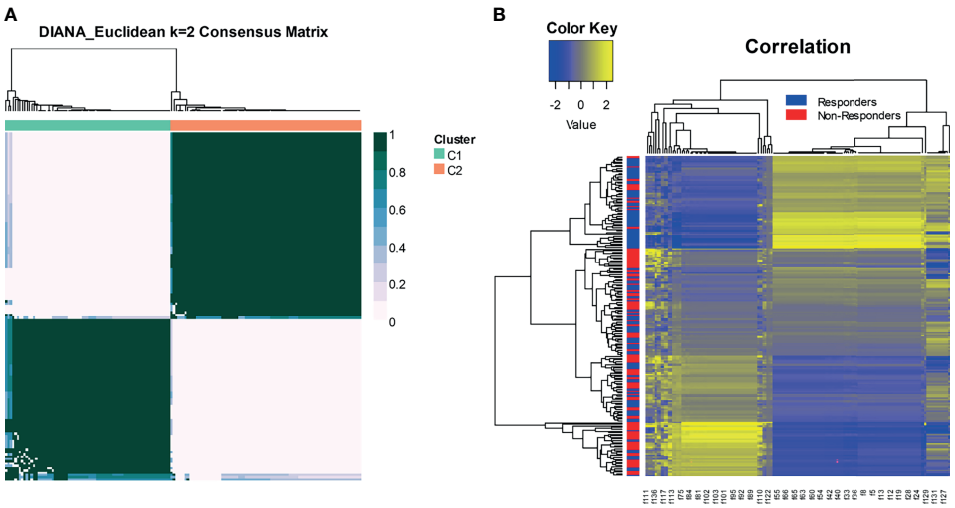


FIGURE 5 | **(A)** Consensus clustering using radiomic features. The two clusters had a preponderance of responders (67%) and non-responders (75%). **(B)** The radiomic heatmap shows an association between the radiomic features and chemotherapy response for patients in the training and test sets.

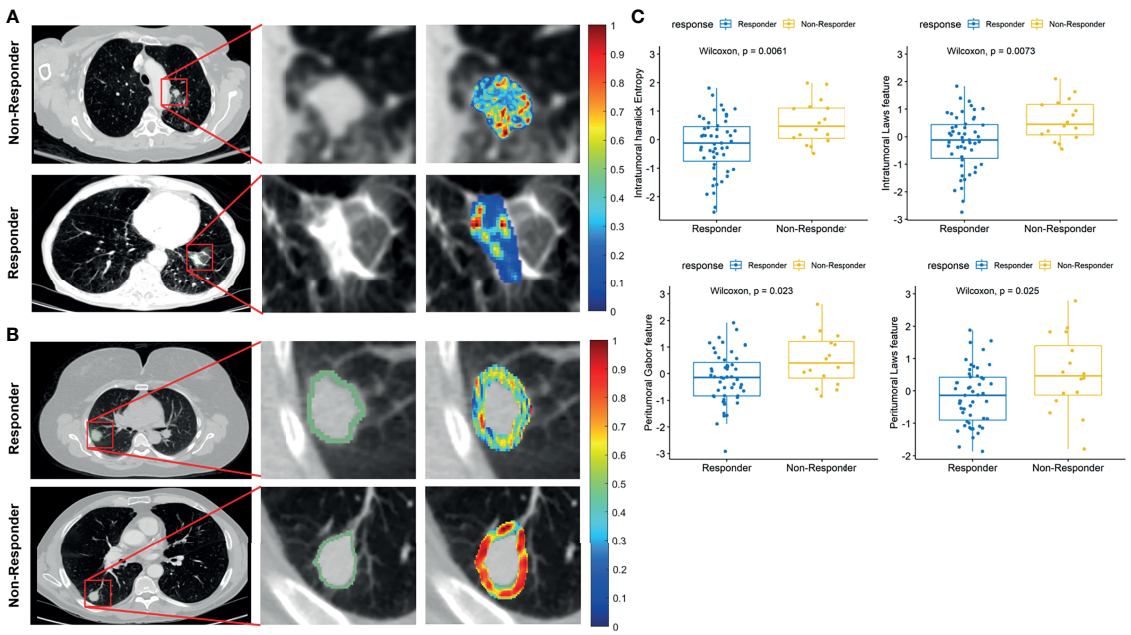


FIGURE 6 | **(A)** Segmented tumor regions and heatmap of intratumoral Haralick (entropy) feature in the pretreatment CT scans for representative non-responder and responder patients. **(B)** Segmented tumor regions and heatmap of peritumoral Gabor feature in the pretreatment CT scans for representative non-responder and responder patients. The middle column is a magnified view of tumor anatomy in both **Figures 5A, B**, and the right column is a color heatmap. **(C)** Box-and-whisker plots for four features that best distinguish chemotherapy response.

In this study, we found that the entropy of the intratumoral Haralick feature as well as the median of the intratumoral Laws texture feature had a higher expression in non-responders compared with the responders. These features can capture intratumoral heterogeneity of the tumor (15). High expression of pretreatment serum vascular endothelial growth factor (VEGF) is an important regulator of angiogenesis and vascular permeability in the cell and is known to be associated with poor response to treatment and unfavorable survival in patients with SCLC treated with chemotherapy (31). CD56 is another neural cell adhesion molecule (NCAM) expressed on the cells of tumors of NE origin including SCLC. The overexpression of CD56 is

associated with lower OS in SCLC since it inhibits tubulin polymerization and microtubule assembly, causing healthy cell death (32). In addition, increased expression of metalloproteinases (MMPs) is associated with poor prognosis. While we did not explicitly confirm this association, it may be that the overexpression of intratumoral Haralick and Laws features reflect the overexpression of VEGF and CD56 inside a tumor, which in turn are associated with poor OS and response to chemotherapy. Another reason could also be that certain high-risk radiomic features might be reflective of tumor hypoxia. Histological examination of SCLC biopsies showed that at least half of all newly diagnosed SCLC patients have tumor hypoxia (33), which promotes tumor proliferation, increases the metastatic potential, and confers resistance to therapy (34). Also, hypoxic regions of the tumor are relatively devoid of blood vessels, making it difficult for drugs to diffuse and reach to the tumor bed.

Our experiment showed that the peritumoral Laws texture feature had higher expression in non-responders compared with the responders and associated with lower OS. Our findings could provide new insight into tumor microenvironment biology, while its appearance on radiographic imaging may be explained as follows. SCLC can be classified into NE-high and NE-low tumors, based on a different immunogenic expression. NE-high is defined as a cold or "immune desert" phenotype, based on low levels of immune cell expression, whereas NE-low is defined as a "hot" or "immune oasis" phenotype associated with increased immunogenicity (35, 36). This extrapolates to better therapy response in NE-low SCLC as compared with NE-high SCLC patients (37, 38). The differences observed in the expression of peritumoral Laws texture features between responders and non-responders could be related to differential expression of immune cells around the tumor.

Our result showed that the RRS and cancer staging (limited or extensive stage) were only two major risk factors in predicting OS in both training (RRS: HR = 2.1, [1.53, 2.85], p = 0.0076; clinical stage: HR = 1.66, [1.01, 2.7], p = 0.048; race: HR = 0.37, [0.12, 1.1], p = 0.071; and age: HR = 1.04, [0.99, 1.09], p = 0.071; C-index = 0.75) and test sets (RRS: HR = 1.9, [1.23, 2.2], p = 0.0012; clinical stage: HR = 1.61, [1.2, 2.17], p = 0.041; race: HR, 0.86, 95% CI: [0.52, 1.42], p = 0.56; and age: HR, 1.01, 95% CI: [0.99, 1.03], p = 0.22; C-index = 0.71). Moreover, RRS and cancer staging were also two major risk factors in predicting PFS in training (RRS: HR = 1.8, [1.48, 2.23], p = 0.0082; clinical stage: HR = 1.35, [1.08, 1.78], p = 0.033; C-index = 0.72) and test sets (RRS: HR = 1.61, [1.13, 1.96], p = 0.0046; clinical stage: HR = 1.37, [1.1, 2.01], p = 0.039; C-index = 0.70). In addition, while radiomic features were associated with response to chemotherapy, the clinical biomarkers (age, sex, and tumor stage) were not able to predict response to chemotherapy.

Finally, we presented a nomogram that integrated RRS with clinical biomarkers ($N^{Rad+Clin}$) to further improve its prognostic accuracy. The $N^{Rad+Clin}$ exhibited the highest C-index value in both training and test cohorts as compared with the clinical or radiomic models alone. We also evaluated $N^{Rad+Clin}$ by DCA and calculated the net benefit of our model. The decision curve indicated that $N^{Rad+Clin}$ had the highest overall net benefit in predicting high-risk patients for receiving more aggressive treatment than the clinicopathologic measurements across all threshold probability values. With the capability to assess prognosis and response to therapy upfront, the oncologist can be assisted with decision making to estimate therapeutic outcomes for a given patient to reduce ineffective treatments and/or associated toxicity.

To the best of our knowledge, this study is the first to explore the relationship of radiomic features with chemotherapy response and OS in SCLC patients. The convergence of these areas provides a new radiomic model that could yield effective non-invasive prediction of treatment response without sacrificing transparency of biological rationale.

We acknowledge that our study did have its limitations. The cohort sizes for both training and test are relatively small; however, given the relatively low incidence of SCLC as compared with other lung cancer subtypes, the number of patients in the historic studies has been small as well. We also studied a single institution, which may affect the results. Additionally, a couple of recent studies have rigorously and quantitatively investigated the influence of convolution kernels, reconstruction algorithms, and slice thickness on radiomic features for characterization of lung nodules on CT. In the present study, we did not explicitly consider the influence of these parameters on the extracted texture features but randomly distributed the cases with different image acquisition parameters between training and test sets to account for variability. Since features pertaining to radiographic images reflect hallmarks of tumor biology and not drug/agent specific, we expect that the predictive aspects of these features will remain broadly similar across different chemotherapeutic agents. We hope to address in future studies the rigorous morphologic and molecular underpinning of the radiomic findings that were shown in this study.

5 CONCLUSION

In conclusion, our results suggest that radiomic texture features from baseline CT scans of SCLC patients can predict resistance to platinum-based chemotherapy. Our study highlights that these radiomic features are also associated with OS in patients with SCLC, and we presented an integrated nomogram that can estimate the survival probability based on the RRS and clinical biomarkers.

DATA AVAILABILITY STATEMENT

Data are available upon reasonable request. Access to datasets from the University Hospitals Cleveland Medical Center (used with permission for this study) should be requested directly from these institutions via their data access request forms. Subject to the institutional review boards' ethical approval, unidentified data would be made available as a test subset.

ETHICS STATEMENT

This study was conducted in full accordance with the Health Insurance Portability and Accountability Act (HIPAA)

regulations after approval from the Institutional Review Board at Case Western Reserve University (Cleveland, OH), and the IRB waived the requirements for informed consent of all patients because of the retrospective, non-interventional, and non-therapeutic nature of this study.

AUTHOR CONTRIBUTIONS

Conceptualization: PJ, MK, and AM. Methodology: PJ, MK, and AM. Software: MK. Validation: PJ, MK, and AM. Formal analysis: MK, PJ, and PF. Investigation: MK and PJ. Resources: MK. Data curation: PJ, KB, PR, AG, VV, and AD. Writing—original draft preparation: PJ, MK, and AG. Writing—review and editing: PJ, MK, AG, KB, PR, VV, PF, AD, and AM. Visualization: MK. Supervision: AM. All authors contributed to the article and approved the submitted version.

FUNDING

The research reported in this publication was supported by the National Cancer Institute under award numbers 1U24CA199374-01, R01CA202752-01A1, R01CA208236-01A1, R01CA216579-01A1, R01CA220581-01A1, 1U01CA239055-01, 1U01CA248226-01, and 1U54CA254566-01, National Heart, Lung and Blood Institute 1R01HL15127701A1, National Institute for Biomedical Imaging and Bioengineering 1R43EB028736-01, National Center for Research Resources under award number 1 C06 RR12463-01, VA Merit Review Award IBX004121A from the United States Department of Veterans Affairs Biomedical Laboratory Research and Development Service, the Office of the Assistant Secretary of Defense for Health Affairs, through the Breast Cancer Research Program (W81XWH-19-1-0668), the Prostate Cancer Research Program (W81XWH-15-1-0558, W81XWH-20-1-0851), the Lung Cancer Research Program (W81XWH-18-1-0440, W81XWH-20-1-0595), the Peer Reviewed Cancer Research Program (W81XWH-18-1-0404), the Kidney Precision Medicine Project (KPMP) Glue Grant, the Ohio Third Frontier Technology Validation Fund, the Clinical and Translational Science Collaborative of Cleveland (UL1TR0002548) from the National Center for Advancing Translational Sciences (NCATS) component of the National Institutes of Health and NIH roadmap for Medical Research, and The Wallace H. Coulter Foundation Program in the Department of Biomedical Engineering at Case Western Reserve University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the U.S. Department of Veterans Affairs, the Department of Defense, or the United States Government.

REFERENCES

1. Lemjabbar-Alaoui H, Hassan OU, Yang YW, Buchanan P. Lung Cancer: Biology and Treatment Options. *Biochim Biophys Acta* (2015) 1856(2):189–210. doi: 10.1016/j.bbcan.2015.08.002
2. Alvarado-Luna G, Morales-Espinosa D. Treatment for Small Cell Lung Cancer, Where Are We Now?-A Review. *Transl Lung Cancer Res* (2016) 5 (1):26–38. doi: 10.3978/j.issn.2218-6751.2016.01.13
3. Riaz SP, Lüchtenborg M, Coupland VH, Spicer J, Peake MD, Møller H. *Lung Cancer - Small Cell: Statistics [Webpage on the Internet]*. Alexandria, VA: American Society of Clinical Oncology (ASCO; Cancer.Net (2019). Available at: https://www.cancer.net/cancer-types/lung-cancer-small-cell/statistics.
4. Hurwitz JL, McCoy F, Scullin P, Fennell DA. New Advances in the Second-Line Treatment of Small Cell Lung Cancer. *Oncologist* (2009) 14(10):986–94. doi: 10.1634/theoncologist.2009-0026
5. Schneider BJ. Management of Recurrent Small Cell Lung Cancer. *J Natl Compr Canc Netw* (2008) 6(3):323–31. doi: 10.6004/jnccn.2008.0027
6. Tas F, Ciftci R, Kilic L, Karabulut S. Age Is a Prognostic Factor Affecting Survival in Lung Cancer Patients. *Oncol Lett* (2013) 6(5):1507–13. doi: 10.3892/ol.2013.1566
7. Fischer BM, Mortensen J, Langer SW, Loft A, Berthelsen AK, Petersen BI, et al. A Prospective Study of PET/CT in Initial Staging of Small-Cell Lung Cancer: Comparison With CT, Bone Scintigraphy and Bone Marrow Analysis. *Ann Oncol* (2007) 18(2):338–45. doi: 10.1093/annonc/mdl374
8. Sandler AB. Current Management of Small Cell Lung Cancer. *Semin Oncol* (1997) 24(4):463–76. doi: 10.1016/j.ccm.2011.07.002
9. Kalemkerian GP, Akerley W, Bogner P, Borghaei H, Chow LQ, Downey RJ, et al. National Comprehensive Cancer Network. Small Cell Lung Cancer. *J Natl Compr Canc Netw* (2013) 11(1):78–98. doi: 10.6004/jnccn.2013.0011
10. Foster NR, Mandrekar SJ, Schild SE, Nelson GD, Rowland KMJr, Deming RL, et al. Prognostic Factors Differ by Tumor Stage for Small Cell Lung Cancer: A Pooled Analysis of North Central Cancer Treatment Group Trials. *Cancer* (2009) 115(12):2721–31. doi: 10.1002/cncr.24314
11. Dowlati A, Lipka MB, McColl K, Dabir S, Behtaj M, Kresak A, et al. Clinical Correlation of Extensive-Stage Small-Cell Lung Cancer Genomics. *Ann Oncol* (2016) 27(4):642–7. doi: 10.1093/annonc/mdw005
12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169
13. Khorrami M, Prasanna P, Gupta A, Patil P, Velu PD, Thawani R, et al. Changes in CT Radiomic Features Associated With Lymphocyte Distribution Predict Overall Survival and Response to Immunotherapy in Non-Small Cell Lung Cancer. *Cancer Immunol Res* (2020) 8(1):108–19. doi: 10.1158/2326-6066.CIR-19-0476
14. Khorrami M, Khunger M, Zagouras A, Patil P, Thawani R, Bera K, et al. Combination of Peri- and Intratumoral Radiomic Features on Baseline CT Scans Predicts Response to Chemotherapy in Lung Adenocarcinoma. *Radiol Artif Intell* (2019) 1(2):e180012. doi: 10.1148/ryai.2019180012
15. Khorrami M, Jain P, Bera K, Alilou M, Thawani R, Patil P, et al. Predicting Pathologic Response to Neoadjuvant Chemoradiation in Resectable Stage III Non-Small Cell Lung Cancer Patients Using Computed Tomography Radiomic Features. *Lung Cancer* (2019) 135:1–9. doi: 10.1016/j.lungcan.2019.06.020
16. Vaidya P, Bera K, Patil PD, Gupta A, Jain P, Alilou M, et al. Novel, Non-Invasive Imaging Approach to Identify Patients With Advanced Non-Small Cell Lung Cancer at Risk of Hyperprogressive Disease With Immune Checkpoint Blockade. *J Immunother Cancer* (2020) 8(2):e001343. doi: 10.1136/jitc-2020-001343
17. Budai BK, Tóth A, Borsos P, Frank VG, Shariati S, Fejér B, et al. Three-Dimensional CT Texture Analysis of Anatomic Liver Segments can Differentiate Between Low-Grade and High-Grade Fibrosis. *BMC Med Imaging* (2020) 20(1):108. doi: 10.1186/s12880-020-00508-w
18. Khorrami M, Bera K, Thawani R, Rajiah P, Gupta A, Fu P, et al. Distinguishing Granulomas From Adenocarcinomas by Integrating Stable and Discriminating Radiomic Features on Non-Contrast Computed Tomography Scans. *Eur J Cancer* (2021) 148:146–58. doi: 10.1016/j.ejca.2021.02.008

19. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1). *Eur J Cancer* (2009) 45(2):228–47. doi: 10.1016/j.ejca
20. Komatsu T, Kunieda E, Oizumi Y, Tamai Y, Akiba T. Clinical Characteristics of Brain Metastases From Lung Cancer According to Histological Type: Pretreatment Evaluation and Survival Following Whole-Brain Radiotherapy. *Mol Clin Oncol* (2013) 1(4):692–8. doi: 10.3892/mco.2013.130
21. Apte AP, Iyer A, Crispin-Ortuzar M, Pandya R, van Dijk LV, Spezi E, et al. Technical Note: Extension of CERR for Computational Radiomics: A Comprehensive MATLAB Platform for Reproducible Radiomics Research. *Med Phys* (2018) 13:10.1002/mp.13046. doi: 10.1002/mp.13046
22. Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* (2017) 77(21):e104–7. doi: 10.1158/0008-5472.CAN-17-0339
23. Armato SG 3rd, Meyer CR, Mcnitt-Gray MF, McLennan G, Reeves AP, Croft BY, et al. The Reference Image Database to Evaluate Response to Therapy in Lung Cancer (Rider) Project: A Resource for the Development of Change-Analysis Software. *Clin Pharmacol Ther* (2008) 84:448–56. doi: 10.1038/clpt.2008.161
24. Vickers AJ, Elkin EB. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Med Decis Making* (2006) 26(6):565–74. doi: 10.1177/0272989X06295361
25. Li W, Nyholt DR. Marker Selection by Akaike Information Criterion and Bayesian Information Criterion. *Genet Epidemiol* (2001) 21(Suppl 1):S272–7. doi: 10.1002/gepi.2001.21.s1.s272
26. Chong S, Lee KS, Chung MJ, Han J, Kwon OJ, Kim TS. Neuroendocrine Tumors of the Lung: Clinical, Pathologic, and Imaging Findings. *Radiographics* (2006) 26(1):41–57; discussion 57–8. doi: 10.1148/rg.261055057
27. Stewart CA, Gay CM, Xi Y, Sivajothi S, Sivakamasundari V, Fujimoto J, et al. Single-Cell Analyses Reveal Increased Intratumoral Heterogeneity After the Onset of Therapy Resistance in Small-Cell Lung Cancer. *Nat Cancer* (2020) 1:423–36. doi: 10.1038/s43018-019-0020-z
28. Simpson KL, Stoney R, Frese KK, Simms N, Rowe W, Pearce SP, et al. A Biobank of Small Cell Lung Cancer CDX Models Elucidates Inter- and Intratumoral Phenotypic Heterogeneity. *Nat Cancer* (2020) 1:437–51. doi: 10.1038/s43018-020-0046-2
29. Ireland AS, Micinski AM, Kastner DW, Guo B, Wait SJ, Spainhower KB, et al. MYC Drives Temporal Evolution of Small Cell Lung Cancer Subtypes by Reprogramming Neuroendocrine Fate. *Cancer Cell* (2020) 38(1):60–78.e12. doi: 10.1016/j.ccell.2020.05.001
30. Gay CM, Stewart CA, Park EM, Diao L, Groves SM, Heeke S, et al. Patterns of Transcription Factor Programs and Immune Pathway Activation Define Four Major Subtypes of SCLC With Distinct Therapeutic Vulnerabilities. *Cancer Cell* (2021) 39(3):346–60.e7. doi: 10.1016/j.ccell.2020.12.014
31. Salven P, Ruotsalainen T, Mattson K, Joensuu H. High Pre-Treatment Serum Level of Vascular Endothelial Growth Factor (VEGF) Is Associated With Poor Outcome in Small-Cell Lung Cancer. *Int J Cancer* (1998) 79(2):144–6. doi: 10.1002/(sici)1097-0215(19980417)79:2<144::aid-ijc8>3.0.co;2-t
32. Karim SM, Zekri J. Chemotherapy for Small Cell Lung Cancer: A Comprehensive Review. *Oncol Rev* (2012) 6(1):e4. doi: 10.4081/oncol.2012.e4
33. Ioannou M, Papamichali R, Kouvaras E, Mylonis I, Vageli D, Kerenidou T, et al. Hypoxia Inducible Factor-1 Alpha and Vascular Endothelial Growth Factor in Biopsies of Small Cell Lung Carcinoma. *Lung* (2009) 187(5):321–9. doi: 10.1007/s00408-009-9169-z
34. Bryant JL, Meredith SL, Williams KJ, White A. Targeting Hypoxia in the Treatment of Small Cell Lung Cancer. *Lung Cancer* (2014) 86(2):126–32. doi: 10.1016/j.lungcan.2014.08.003
35. Rudin CM, Poirier JT, Byers LA, Dive C, Dowlati A, George J, et al. Molecular Subtypes of Small Cell Lung Cancer: A Synthesis of Human and Mouse Model Data. *Nat Rev Cancer* (2019) 19(5):289–97. doi: 10.1038/s41568-019-0133-9
36. McColl K, Wildey G, Sakre N, Lipka MB, Behtaj M, Kresak A, et al. Reciprocal Expression of INSM1 and YAP1 Defines Subgroups in Small Cell Lung Cancer. *Oncotarget* (2017) 8:73745–56. doi: 10.18632/oncotarget.20572
37. Schwendenwein A, Megyesfalvi Z, Barany N, Valko Z, Bugyik E, Lang C, et al. Molecular Profiles of Small Cell Lung Cancer Subtypes: Therapeutic Implications. *Mol Ther Oncolytics* (2021) 20:470–83. doi: 10.1016/j.omto.2021.02.004
38. Saito M, Shiraishi K, Goto A, Suzuki H, Kohno T, Kono K. Development of Targeted Therapy and Immunotherapy for Treatment of Small Cell Lung Cancer. *Jpn J Clin Oncol* (2018) 48(7):603–8. doi: 10.1093/jjco/hyy068

Conflict of Interest: AM is an equity holder in Elucid Bioimaging and in Inspirata, Inc. In addition, he has served as a scientific advisory board member for Inspirata, Inc., AstraZeneca, Bristol Meyers-Squibb, and Merck. Currently, he serves on the advisory board of Aiforia Inc. He also has sponsored research agreements with Philips, AstraZeneca, Boehringer-Ingelheim, and Bristol Meyers-Squibb. His technology has been licensed to Elucid Bioimaging. He is also involved in a NIH U24 grant with PathCore Inc. and three different R01 grants with Inspirata, Inc. AD has received consultancy fees for advisory committees from BMS, AZ, Bayer, and Jazz Pharmaceuticals. AG has received research support from General Electric Healthcare.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Jain, Khorrami, Gupta, Rajiah, Bera, Viswanathan, Fu, Dowlati and Madabhushi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 28 October 2021
doi: 10.3389/fonc.2021.759053

^{18}F-PSMA-1007 PET/CT Performance on Risk Stratification Discrimination and Distant Metastases Prediction in Newly Diagnosed Prostate Cancer

*Zhuonan Wang[1], Anqi Zheng[1], Yunxuan Li[1], Weixuan Dong[1], Xiang Liu[1], Wang Yuan[1], Fan Gao[2] and Xiaoyi Duan[1]**

[1] PET/CT Unit, Department of Medical Imaging, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China, [2] Clinical Research Center, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China

OPEN ACCESS

Edited by:
Kyung Hyun Sung, UCLA Health System, United States

Reviewed by:
Orazio Schillaci, University of Rome Tor Vergata, Italy
Min Yang, Jiangsu Institute of Nuclear Medicine, China

***Correspondence:**
Xiaoyi Duan
duanxyi@yeah.net

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *15 August 2021*
Accepted: *12 October 2021*
Published: *28 October 2021*

Citation:
Wang Z, Zheng A, Li Y, Dong W, Liu X, Yuan W, Gao F and Duan X (2021) 18F-PSMA-1007 PET/CT Performance on Risk Stratification Discrimination and Distant Metastases Prediction in Newly Diagnosed Prostate Cancer. Front. Oncol. 11:759053. doi: 10.3389/fonc.2021.759053

Objective: To evaluate the prediction performance of ^{18}F-PSMA-1007 PET/CT and clinicopathologic characteristics on prostate cancer (PCa) risk stratification and distant metastatic prediction.

Materials and Methods: A retrospective analysis was performed on 101 consecutively patients with biopsy or radical prostatectomy proved PCa who underwent ^{18}F-PSMA-1007 PET/CT. The semi-quantitative analysis provided minimum, maximum and mean standardized uptake (SUVmin, SUVmax and SUVmean) of PCa. Association between clinicopathologic characteristics (total prostate-specific antigen, tPSA and Gleason Score, GS) and PET/CT indexes were analyzed. The diagnostic performance of distant metastatic on PET/CT parameters, tPSA and GS was evaluated using logistic regression analyses. A path analysis was conducted to evaluate the mediating effect of tPSA level on the relation between semi-quantitative parameters of primary tumors and metastatic lesions.

Results: The PET/CT parameters were all higher in high risk stratification subgroups (tPSA>20 ng/mL, GS ≥ 8, and tPSA>20 ng/mL and/or GS ≥ 8, respectively) with high sensitivity (86.89%, 90.16% and 83.61%, respectively). The SUVmax, tPSA and GS could effectively predict distant metastatic with high sensitivity of SUVmax (90.50%) compared with tPSA (57.14%) and GS (55.61%). With a cutoff value of 29.01ng/mL for tPSA, the detection rate of distant metastasis between low and high prediction tPSA group had statistical differences (50.00% vs. 76.60%, respectively; P = 0.006) which was not found on guideline tPSA level (P>0.05). 6/15 (40%) patients tPSA between 20ng/mL to 29.01ng/mL without distant metastases may change the risk stratification. Finally, tPSA had a partial mediating effect on SUVmax of primary tumors and metastases lesions.

Conclusion: The ^{18}F-PSMA-1007 PET/CT SUVmax has a higher sensitivity and can be an "imaging biomarker" for primary PCa risk stratification. The prediction tPSA level (29.01 ng/mL) is more conducive to the assessment of distant metastasis and avoid unnecessary biopsy.

Keywords: primary prostate cancer, ^{18}F-PSMA-1007 PET/CT, risk stratification, distant metastasis, tPSA, SUVmax

INTRODUCTION

Prostate cancer (PCa) is the highest malignant male tumor and one of the leading causes of mortality among men worldwide (1, 2). The biological behaviors of PCa malignancy are largely heterogeneous, directly impacting prognostic grouping, and treatment options (3). In addition, assessments of the distant metastatic status for PCa patients have recently received increasing attention due to the heightening mortality rate (4, 5). Therefore, the precise systemic staging of primary PCa risk stratification before treatment plays a crucial role in designing the management strategy for the individualized treatment option. According to both American Urological Association (AUA) and the European Association of Urology (EAU) guidelines, patients with total prostate-specific antigen (tPSA) > 20 ng/mL and/or Gleason Score ≥ 8 are high-risk, the probability of distant metastasis and mortality will increase significantly and may not suitable for active surveillance programs, radical prostatectomy or radiotherapy treatment (6–8). However, tPSA is organ-specific but not tumor-specific, the biological behaviors of prostate malignancy are largely heterogeneous, and the specificity of the ability of tPSA to reflect distant metastasis remains debatable (6). Using tPSA as the only indicator for risk stratification discrimination and distant metastases prediction may causing in large numbers of unnecessary prostate biopsies (9–12). Also, elderly patients with severe comorbidities or undergoing anticoagulation therapy may not be the optimal candidates for biopsies and may cause adverse effects and higher costs (10). In these cases, it is urgent to find objective and accurate imaging biomarker for risk stratification classification with noninvasive approach based on imaging analysis.

The prostate-specific membrane antigen (PSMA) is a type II transmembrane glycoprotein that is primary expressed in prostatic tissues, and its expression correlated with the degree of malignancy and further increases in metastatic (13, 14). The ability of PSMA to easily penetrate tissues and diffuse with solid tumor lesions can reflect the statuses of metastasis (3, 15, 16). Prior studies show that PSMA PET/CT is superior to conventional imaging methods for lymph node metastatic detection, and that the pre-treatment tPSA level and Gleason Score are associated with the PSMA uptake in primary PCa (17, 18). Furthermore, the Maximum Standardized Uptake Value (SUVmax) is the most commonly used semi-quantitative parameter in PET/CT and prior studies has already used to assess the degree of malignancy of PCa and predict extended pelvic lymph node metastases in intermediate to high-risk PCa patients by ^{68}Ga-PSMA-11 or ^{68}Ga-PSMA-617 (19). ^{18}F-PSMA-1007 is advantaged by its higher spatial resolution images and non-urinary excretion that reduces urinary clearance, this approach bears a great potential to facilitate the detection of primary PCa and metastatic lesions (20, 21). However, to our knowledge, no prior studies have employed ^{18}F-PSMA-1007 PET/CT to evaluate the diagnostic performance in risk stratification and distant metastases prediction in primary PCa.

The present study aims to retrospective investigated the role of ^{18}F-PSMA-1007 PET/CT semi-quantitative parameters correlation among newly diagnosed PCa imaging, tPSA levels and Gleason Score, and to evaluate the prediction performance of ^{18}F-PSMA-1007 PET/CT and clinicopathologic characteristics on PCa risk stratification and distant metastatic prediction.

MATERIALS AND METHODS

Patients

The study has been approved by the institutional review board (No. 2019LSYZD-J1-H) and was conducted in accordance with the Declaration of Helsinki. All subjects signed an informed consent form. We performed retrospective analysis for 101 patients with primary PCa confirmed by biopsy or radical prostatectomy between September 2020 and May 2021. All participants included in the data analysis were evaluated by ^{18}F-PSMA-1007 PET/CT and had tPSA value measured within 4 weeks prior to the ^{18}F-PSMA-1007 PET/CT imaging. Diagnosis of PCa proven through histological examination served as reference for the PET imaging analyses (18). Patients were excluded from analysis if they 1) had received local or systemic treatment, 2) lacked histological examination proven diagnosis or tPSA value, 3) had incomplete imaging data. The flowchart of patient enrollment is provided in **Figure 1**.

^{18}F-PSMA-1007 Acquisition and Imaging Analysis

All ^{18}F-PSMA-1007 PET/CT data was acquired on a PET/CT scanner (Gemini 64TF, Philips, Netherlands) at a single location. Radiolabeling was performed using a fully automated radiopharmaceutical synthesis device based on a modular concept (MINItrace, GE Healthcare, USA). Over 99% radiochemical purification yield ^{18}F-PSMA was obtained and examined by both radio-thin layer chromatography (TLC) and high-performance liquid chromatography (HPLC) analysis. Patients received intravenous injection of ^{18}F-PSMA-1007 (3.7 MBq/kg body weight), and completed PET and CT scans 90 minutes after the injection. Low-dose CT scans from head to the proximal thighs (pitch 0.8 mm, 60 mA, 140 kV [peak], tube single turn rotation time 1.0 s and 5-mm slice thickness) for PET attenuation were acquired (pitch 0.8 mm, automatic mA, 140 kV [peak] and 512 × 512 matrix). Whole-body PET scans were performed in three-dimensional mode (emission time: 90 s per bed position, scanned at a total of 7-10 beds).

All ^{18}F-PSMA-1007 PET/CT images were analyzed using Fusion Viewer software in the Extended Brilliance Workstation (EBW, Philips, Netherlands). Two experienced nuclear medicine specialists jointly interpreted all ^{18}F-PSMA-1007 PET/CT scans, and performed comprehensive analysis of available clinical data. Consensuses were achieved through discussion when conclusions between the two specialists were discordant. The PET indexes (including SUVmin, SUVmax and SUVmean) of the primary PCa was calculated automatically with a manually adapted isocontour threshold centered on lesions with focally increased uptake corresponding to the tumor site verified by TRUS biopsy or radical prostatectomy (17). The PET indexes

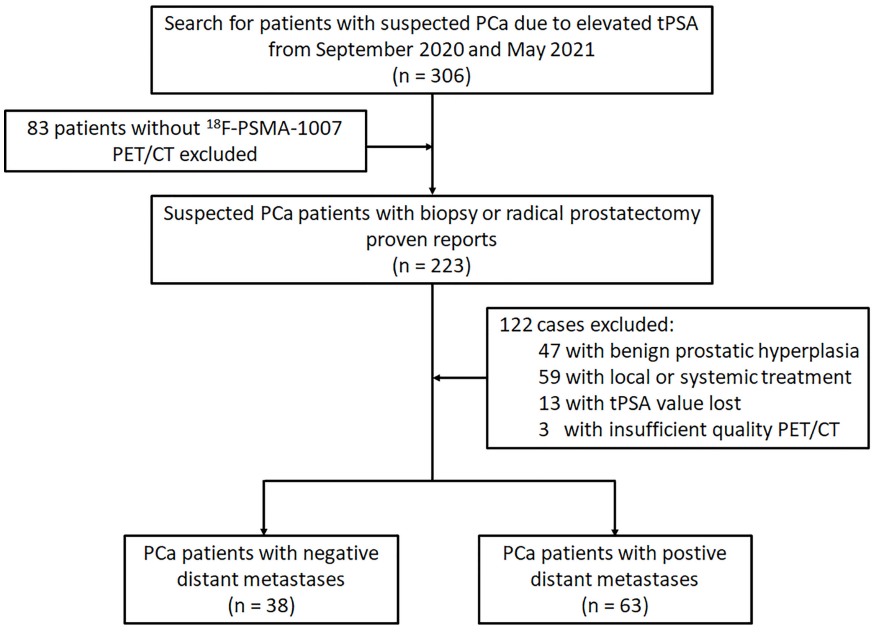


FIGURE 1 | Flowchart of the PCa patient's cohort.

values were also calculated for metastases lesions. The positive lesion was defined by an uptake higher than the local background and not associated with physiologic uptake per the guideline of the Society of Nuclear Medicine and Molecular Imaging and the European Association of Nuclear Medicine (22, 23). The identified metastases were also consistent with PCa lesions pathologic tracer accumulation (6, 22, 23). The PET/CT distant metastasis positive lesions were also composite validated by other imaging approaches (bone scan and MR) as a reference, and the patients were followed for tPSA measurements, imaging follow-up (PET/CT, bone scan and MR), disease management as the metastatic definition reference (24, 25). PET/CT scan findings was classified as (a) primary tumor, and (b) distance metastasis (abdominal lymph nodes, bone and internal organs) (**Figure 2**).

Statistical Analysis

Descriptive statistics was used to display patient data as median, mean, standard deviation range or percentages, where applicable. Correlation among PET/CT indexes (including SUVmin, SUVmax and SUVmean) and variables were evaluated with Spearman's rank correlation coefficient. Wilcoxon- Mann-Whitney U test was used to test the subgroups (tPSA > 20 ng/mL vs tPSA ≤ 20 ng/mL; Gleason Score ≥ 8 vs Gleason Score<8; High-risk vs Low-Intermediate risk) PET indexes differences. The risk stratification discrimination of PET/CT indexes and distant metastases status prediction combined PET/CT indexes with clinicopathologic characteristics were all assessed using receiver operating characteristic (ROC) curve analysis. The PET/CT parameters and clinicopathologic characteristics were used to construct logistic regression prediction model of metastasis findings on ^{18}F-PSMA-1007. The metastasis diagnostic efficiency of the model was evaluated by ROC curve and the best threshold of tPSA and PET/CT indexes performance was based on Youden index. Statistical significance of the association between positive/negative metastasis findings on ^{18}F-PSMA-1007 and clinicopathologic characteristics was assessed with chi-squared test. The path analyses to examine the potential mediating role of tPSA level on PET/CT index across different metastatic statuses. A significance level of $\alpha = 0.05$ (two-tailed) was applied. Statistical analyses were performed using IBM SPSS Statistics version 13.0, GraphPad Prism software, version 8.4 and MedCalc version 19.0.

RESULTS

Demographic information and clinical characteristics of the participants were summarized in **Table 1**. A total of 208 lesions were identified using ^{18}F-PSMA-1007 PET/CT, including 101 primary prostate tumors and 107 distant metastases lesions. The median SUVmax of all primary PCa lesions was 26.00 (range: 5.95-101.89), which was significantly higher than that of the metastasis's lesions (median SUVmax: 16.90, range: 5.44-150.24). Comparing the subgroup with tPSA > 20 ng/mL and that with tPSA ≤ 20 ng/mL, Gleason Score ≥ 8 and that with Gleason Score < 8, and high risk and that with low-intermediate risk, the PET/CT semi-quantitative parameters (SUVmin, SUVmax and SUVmean) of the first group were all higher than that of the second group in **Table 2**. The PET/CT indexes were all significantly correlated with the tPSA (r_s = 0.405, 0.380, 0.418, respectively, $P < 0.001$), Gleason Score (r_s = 0.407, 0.339, 0.387, respectively, $P < .001$) and risk stratification (r_s = 0.432, 0.354, 0.430, respectively, $P < .001$).

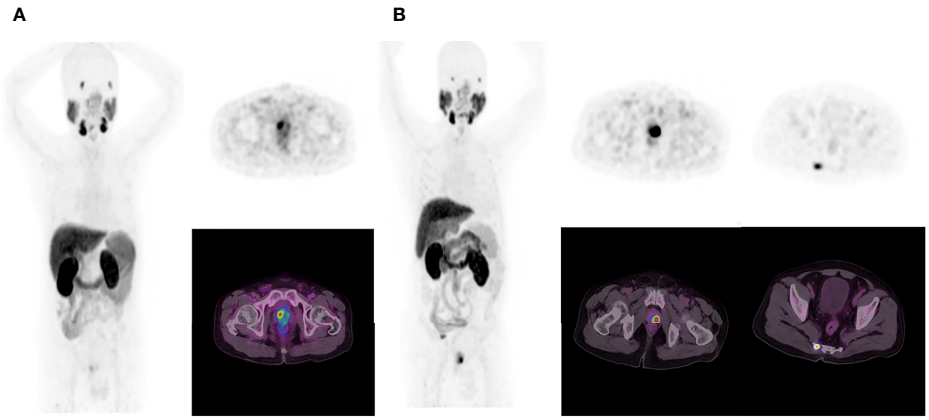


FIGURE 2 | In the first patient **(A)**, ^{18}F-PSMA-1007 PET/CT results show primary prostate cancer on whole-body maximum intensity projection (MIP) image (left). Axial PET image (right up) and axial fused image (right bottom) show primary tumor in the right prostate cancer lobe. In the second patient **(B)**, ^{18}F-PSMA-1007 PET/CT results show primary prostate cancer and tailbone metastasis on MIP image (left).

To be specific, risk stratification model for PCa was constructed based on semi-quantitative parameters. Area under the ROC curve (AUC) was analyzed with 0.692 (95% CI: 0.593 - 0.780) with sensitivity 86.89% and specificity 55.00% for the SUVmin model (Cut-off value 6.56), 0.684 (95% CI: 0.584 - 0.773) with 90.16% and 47.50% for the SUVmax model (Cut-off value 31.19) and 0.706 (95% CI: 0.607 - 0.792) with 83.61% and 55.00% for the SUVmean model (Cut-off value 10.22) (**Figure 3A**). This study also tests the potential value of ^{18}F-PSMA-1007 PET/CT in predicting the risk of PCa metastasis. The AUC, sensitivity and specificity of SUVmin, SUVmax, SUVmean, tPSA and Gleason Score were measured respectively. The AUC results were 0.602 (95% CI: 0.500 - 0.698) with sensitivity 85.71% and specificity 39.47% for the SUVmin model (Cut-off value 3.35, P >.05); AUC 0.645 (95% CI: 0.543 - 0.738) with sensitivity 90.50% and specificity 34.22% for the SUVmax model (Cut-off value 13.76, P <.05), AUC 0.619 (95% CI: 0.517 - 0.714) with sensitivity 93.71% and specificity 31.63% for the SUVmean model (Cut-off value 13.76, P <.05), AUC 0.656 (95% CI: 0.555-0.748) sensitivity with 57.14% and specificity 73.68% for the tPSA model (Cut-off value 29.01, P <.05) and AUC 0.709 (95% CI: 0.610 - 0.795) with sensitivity 55.61% and specificity 81.64% for the Gleason Score model (Cut-off value 8, P <.05) (**Figure 3B**).

To further evaluate the diagnostic strength of PET/CT indexes (SUVmax and SUVmean), tPSA level and Gleason Score for metastasis, SUVmax, SUVmean, tPSA (tPSA low: ≤20 ng/mL and tPSA high: > 20 ng/mL) and Gleason Score (low-intermediate: <8 and high: ≥ 8) were entered as independent variables in a logistics regression. The metastasis status was entered as a binary outcome variable. As shown in **Table 3**, the logistic regression found the SUVmax, tPSA level (high or low) and Gleason Score (low-intermediate risk or high-risk) were independent predictors for metastasis status. We also applied multivariate logistic regression for metastasis diagnosis. The SUVmax and Gleason Score (low-intermediate risk or high-risk) were stable for metastasis status prediction (OR 1.081, P = .040; OR 2.602, P = .042, respectively). The tPSA level (high or low) had no significant difference in multivariate logistic regression results. Details can be found in **Supplementary Materials**.

We further analyzed the correlation between tPSA level (tPSA low and tPSA high), Gleason Score (low-intermediate and high), SUVmax and positive distant metastasis findings on ^{18}F-PSMA-1007 PET/CT. The detection rate of distant metastasis was lower in the low-PSA compared with that in the high PSA group with no significance difference (52.63% vs. 69.84%, respectively; P >.05; **Figure 4A**). The detection rate of distant metastasis was lower in the low-intermediate Gleason Score group than in the high Gleason Score group (44.73% vs. 73.02%, respectively; P = .004; **Figure 4B**). We next sought to determine the tPSA level for optimal predicting the metastasis findings based on prior ROC curve. Basing on the cutoff value of 29.01 ng/mL for the prediction

TABLE 1 | Demographic and clinical characteristics of the 101 study participants.

Characteristic	Value
Age (years)	
Median (range)	72 (50-90)
Mean ± SD	71.39 ± 8.72
tPSA (ng/mL)	
Median (range)	24.97 (0.17-2139)
Mean ± SD	123.57 ± 296.50
Non-metastatic Patients (%)	38 (37.6%)
Metastatic Patients (%)	63 (62.4%)
Gleason Score	3 (3.0%)
6	35 (34.7%)
7	21 (20.8%)
8	42 (41.5%)
9	
Risk Group	
Low-Intermediate risk	38 (37.6%)
High risk	63 (62.4%)
Number and percentage of metastatic malignant lesions	
Extrapelvic Lymph node metastases	67/107 (62.6%)
Bone metastases	40/107 (37.4%)

TABLE 2 | ^{18}F-PSMA-1007 PET/CT parameters of different tPSA, Gleason Score and risk stratification subgroups.

Categorical variable	PSA ≤ 20 (n = 40)	PSA > 20 (n = 61)	Sig	GS < 8 (n =38)	GS ≥ 8 (n = 63)	Sig	Low-Intermediate risk (n = 22)	High-risk (n = 79)	Sig
SUVmin									
Mean ± SD (range)	6.85 ± 5.84 (1.41-23.8)	10.68 ± 7.27 (1.89-37.95)	*P = 0.01*	6.77 ± 5.53 (1.41-21.18)	10.61 ± 7.37 (1.75-37.95)	*P = 0.01*	4.92 ± 4.77 (1.41-21.12)	32.37 ± 19.77 (1.75-37.95)	*P = 0.002*
SUVmax									
Mean ± SD (range)	22.35 ± 16.30 (5.95-57.83)	33.84 ± 20.23 (7.06-101.89)	*P = 0.006*	22.80 ± 14.99 (5.95-57.83)	33.21 ± 20.96 (7.06-101.89)	*P = 0.013*	18.22 ± 14.13 (5.95–57.83)	15.88 ± 10.49 (7.06-101.89)	*P = 0.004*
SUVmean									
Mean ± SD (range)	10.75 ± 8.80 (2.67-36.19)	16.42 ± 10.77 (3.72-57.85)	*P = 0.01*	10.69 ± 8.15 (2.67-31.75)	16.27 ± 11.04 (3.72-57.85)	*P = 0.012*	8.05 ± 7.28 (2.67-31.75)	10.34 ± 7.04 (3.72-57.85)	*P = 0.003*

tPSA level, we categorized patients into a low prediction tPSA (tPSA ≤ 29.01 ng/mL) group and a high prediction tPSA group (tPSA>29.01 ng/mL). The detection rate of distant metastasis was lower in the low prediction tPSA compared with that in the high prediction tPSA group with statistical significance difference (50.00% vs. 76.60%, respectively; P = .006; **Figure 4C**). When applied the optimal prediction value of SUVmax (13.76), the high prediction SUVmax had higher detection rate for distant metastasis (47.39% vs.73.02%, respectively; P <.05; **Figure 4D**). In order to better prove the benefit from predicting PSA levels for risk stratification, we retrospectively analyzed primary PCa patients with tPSA between 20ng/mL to 29.01ng/mL. The ^{18}F-PSMA-1007 PET/CT results of PCa patients found that 6 out of 15 (40%) did not have distant metastases.

To better identify the association between prediction tPSA value and SUVmax, we divided PCa patients into six subgroups based on their tPSA level (i.e., tPSA ≤ 29.01 ng/mL, tPSA>29.01 ng/mL) and metastasis staging (i.e., primary prostate tissue without metastasis, primary prostate tissue with metastasis, and primary PCa metastasis focis). ANOVA analysis found that SUVmax was significantly different among the high prediction tPSA subgroups (P = .001) but not within the low prediction tPSA subgroups (P >.05). Post-hoc analysis showed that, among the high prediction tPSA subgroups, only the primary prostate tissue with metastasis group and the primary PCa metastasis foci group demonstrated significantly different SUVmax level (P <.001) (**Figure 5**).

To evaluate the mediation role of tPSA level on SUVmax, we constructed a path analysis model by including SUVmax of primary PCa lesions as the predictor, tPSA level as the mediator, and SUVmax of metastasis foci as the outcome. A positive correlation was found between the primary PCa SUVmax and metastasis foci SUVmax (P <.001). tPSA was found to have a partial mediating effect between the SUVmax values of primary tumors and metastases lesions (P < .05) (**Figure 6**).

DISCUSSION

The present study analyzed the application value of ^{18}F-PSMA-1007 PET/CT to detect risk stratification discrimination and distant metastases prediction in primary PCa. The SUVmax can be used as an "imaging biomarker" for distant metastasis risk prediction in primary PCa. Higher tPSA levels may be more

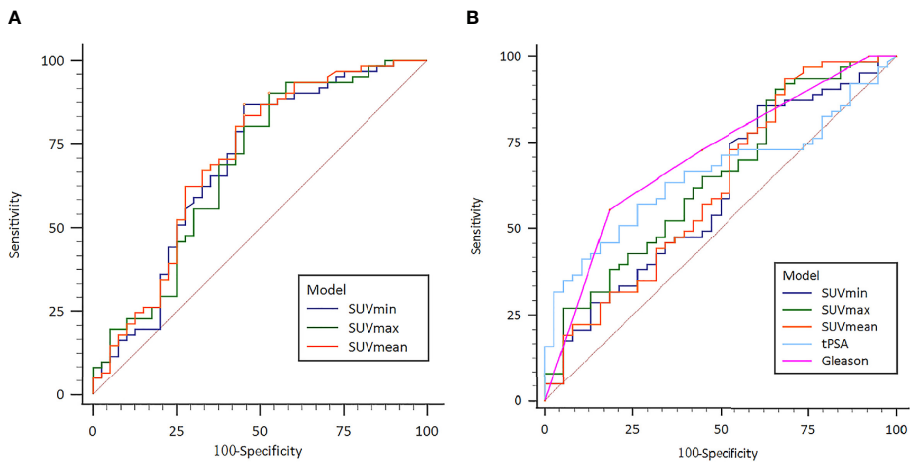


FIGURE 3 | **(A)** Receiver operating characteristic (ROC) curve of the ^{18}F-PSMA-1007 PET/CT indexes for prostate cancer risk stratification; **(B)** Receiver operating characteristic (ROC) curve of the ^{18}F- PSMA-1007 PET/CT prostate cancer metastasis risk prediction.

TABLE 3 | Logistic analyses of factors predicting prostate cancer metastasis.

Categorical variable	OR	95% CI	P
SUVmax	1.032	1.005-1.059	*0.018*
Gleason Score (high vs low-intermediate)	3.120	1.179-8.257	*0.019*
SUVmean	1.040	0.994-1.088	0.087
tPSA (high vs low)	2.389	1.043-5.472	*0.038*

tPSA, total PSA; OR, odds ratio; CI, confidence interval. SUVmax and SUVmean here represent the PET/CT indexes of the primary tumor.

likely to benefit from PET/CT detection of distant metastases. The SUVmax can effectively identify PCa heterogeneity of the primary tumor and metastasis lesions. Reference to the prediction tPSA level (29.01 ng/mL) can at least be a partial but important contribution to risk stratification and avoiding unnecessary invasive examinations. The tPSA has a partial mediating effect between the primary tumor and metastases lesions.

PSMA overexpression in primary PCa was correlated with advanced tumor malignant status with higher tPSA level and Gleason Score (15). Prior studies have shown the SUVmax of ^{68}Ga-PSMA PET/CT associated with tPSA and Gleason Score (3, 17, 26). This study investigated the use of semi-quantitative parameters to determine the risk stratification, similarly correlation was also found between ^{18}F-PSMA -1007 PET/CT parameters with both tPSA and Gleason Score. For PCa risk stratification, the SUVmin, SUVmax and SUVmean can accurately identify high tPSA level (tPSA > 20), high Gleason Score (≥ 8) and high-risk (tPSA > 20 or/and Gleason Score ≥ 8) primary PCa patients The ROC curve analysis indicated that these three semi-quantitative parameters have high sensitivity (86.89%, 90.16% and 83.61%, respectively) to screen out all PCa patients for high risk prediction and may satisficed clinical needs.

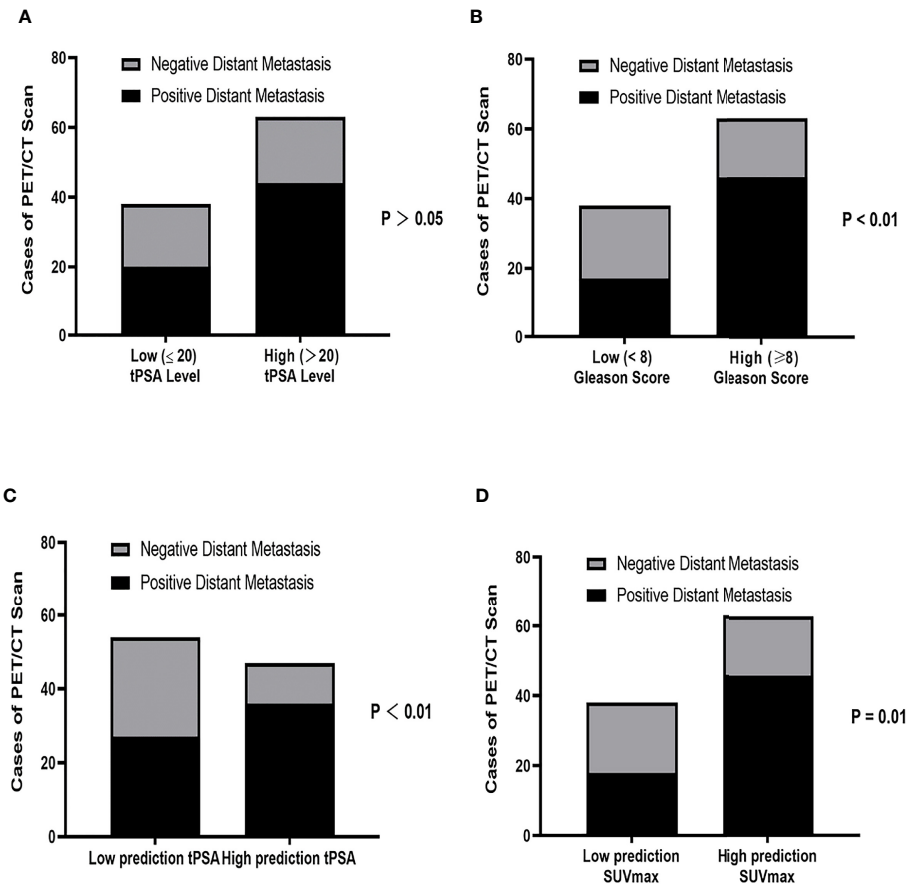


FIGURE 4 | Correlation between **(A)** tPSA level (tPSA high: >20 ng/mL vs tPSA low: ≤20 ng/mL), **(B)** Gleason Score (low-intermediate: <8 vs high: ≥8), **(C)** Prediction tPSA level (tPSA<29.01 ng/mL vs tPSA ≥ 29.01 ng/mL) and **(D)** Prediction SUVmax (SUVmax ≤ 13.76 vs SUVmax>13.76) for positive distant metastasis findings on 18F PSMA-1007 PET/CT. tPSA, total PSA; PCa, prostate cancer.

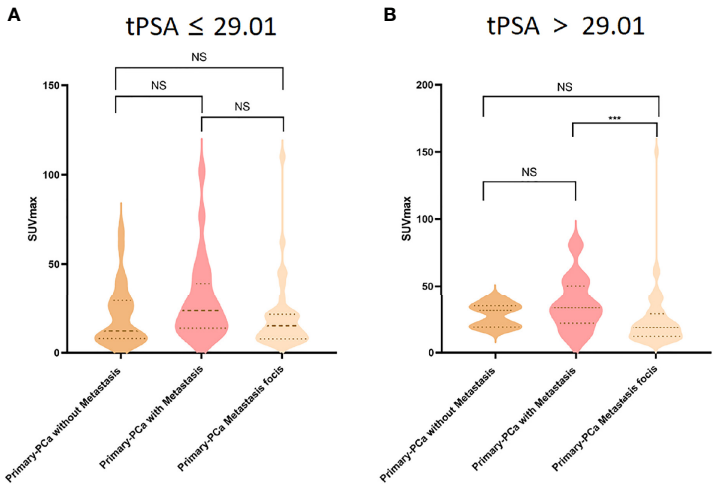


FIGURE 5 | Comparison of ^{18}F-PSMA-1007 SUVmax uptake in primary prostate tissue without metastasis (dark yellow violin box), primary prostate tissue with metastasis (pink violin box) and primary prostate cancer metastasis tissues (light yellow violin box) in tPSA ≤ 29.01 **(A)** and tPSA>29.01 **(B)** tPSA, total PSA; PCa, prostate cancer. ***P < 0.05, NS, No statistical difference.

PCa metastasis risk prediction included both intra-pelvic and distant metastases, however, the treatment options of different location metastases may be quite different (7, 8). Series studies have demonstrated that PSMA PET/CT is more efficient than traditional imaging methods for detecting distant metastases in primary PCa patients and reflect the malignancy and staging PCa, preventing patients from undergoing repeated inspections and invasive biopsy (27–32). Our study explored the prediction value of the distant metastatic risk of primary PCa with ^{18}F-PSMA -1007 PET/CT. The distant metastatic risk prediction model constructed by SUVmax, tPSA and Gleason Score can be used as independent factors of distant metastatic assessment. In the process of visual assessment, there is heterogeneity and false positive rate of lesions in distant location outside the prostate. When pathological results are not available, it may affect the choice of treatment. According to our results, SUVmax with an optimal value 13.76 has a higher sensitivity 90.50% compared with both tPSA (57.14%) and Gleason Score (55.61%), the positive detection rate will increase significantly, which may provide a reference for the diagnosis of distant metastasis. We also analyzed the correlation between distant metastases status and high-risk factors (tPSA, Gleason Score and SUVmax). Similar with previous results, the detection rate of distant metastasis increases with primary tumor malignancy and the tPSA level (33, 34). ^{18}F-PSMA -1007 PET/CT positive distant metastasis PCa patients had a higher Gleason Score and SUVmax with statistical difference compared with PET/CT negative patients.

Prior studies using a tPSA cutoff value of 30 ng/ml in groups of men showed diagnose performance for primary PCa ranging

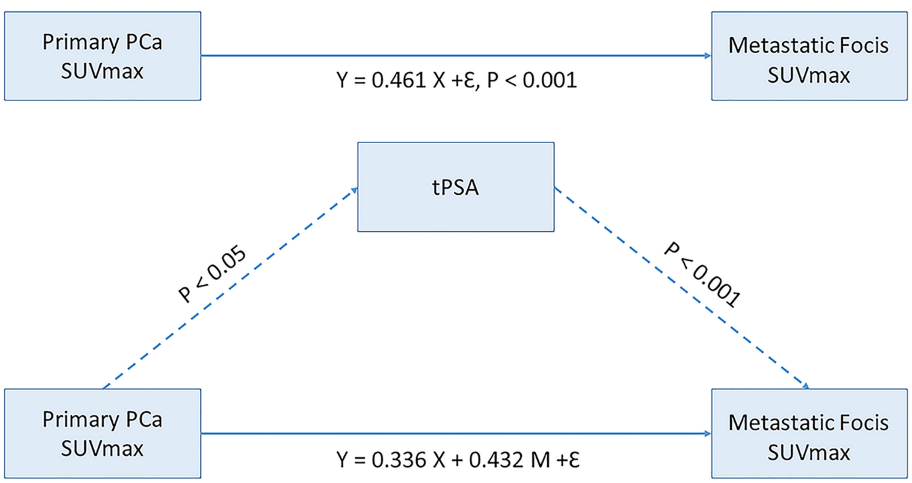


FIGURE 6 | Mediating effect model between primary prostate cancer SUVmax and Metastatic focis. tPSA, total PSA; PCa, prostate cancer.

from 90% to 95.7% (10). Among those who were not diagnosed with malignant PCa through biopsy, bone scan still detected positive metastases. The current study further found that the detection rate of distant metastasis was lower in the low prediction tPSA compared with that in the high prediction tPSA group, and higher detection rate compared with former tPSA (73.68% vs 69.84%). Furthermore, patients with tPSA between 20ng/mL to 29.01ng/mL, the ^{18}F-PSMA-1007 PET/CT results found that 6 out of 15 (40%) did not have distant metastases. Based on the original guidelines may lead to unnecessary biopsy in some patients and the prediction tPSA level an at least be a partial but important contribution to risk stratification and avoiding unnecessary biopsy.

To understand how tPSA level may help differentiate the primary tumor from the distant metastasis, we grouped the PCa patients based on their prediction tPSA level. The SUVmax difference between primary tumors and metastatic lesions in metastatic PCa patients with a high prediction tPSA level (> 29.01ng/mL). Primary PCa can be a heterogeneous multifocal tumor, and may lead to metastatic lesions with varying characteristics (35–40). This finding may further reflect the specificity of the source of distant metastases and prostate characteristics at different tPSA levels. Considering the potential value of tPSA and SUVmax and the differential performance of these two indicators in the primary tumors and metastatic lesions, we recommended that when the patient's tPSA is higher than the prediction level, the patients might be benefit from ^{18}F-PSMA -1007 PET/CT scan for distant metastasis detection.

Our study was limited by the retrospective data collection and a relatively small sample size. Future work in this area should consider combining PET/CT imaging findings with pathological results of distant metastatic lesions to improve overall diagnostic accuracy. In addition, benign hyperplasia and inflammation may interfere with the SUVmax measurements of PCa. Methods to improve the accuracy of SUVmax measurement may help increase detection specificity. We also plan to implement the analytic approach for sites of metastases in larger datasets currently being collected in the lab with the hope of the association between sites of metastases and tPSA level.

In summary, ^{18}F-PSMA-1007 PET/CT is of good application value in the risk stratification and distant metastatic of primary PCa. When tPSA is higher than the prediction level, the probability of distant metastasis will increase. The newly tPSA prediction level may change the risk stratification and patient's treatment options.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the institutional review board of The First Affiliated Hospital of Xi'an Jiaotong University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ZW draft the manuscript, contributed to the conception and design, analysis and interpretation of data. AZ contributed to the design and analysis of data. YL contributed to the analysis and interpretation of data. WD contributed to the conception and design, and acquisition of data. XL contributed to the conception and design, and acquisition of data. WY contributed to the acquisition of data and analysis of data. FG contributed to the conception and design, analysis and interpretation of data. XD contributed to revise of the manuscript critically for important intellectual content. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the 2020 Natural Science Basic Research Program of Shaanxi Province, China (2020JZ-38) and 2020 Clinical Research Project of The First Affiliated Hospital of Xi'an Jiaotong University (XJTU1AF-CRF-2020-008).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.759053/full#supplementary-material

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. *CA Cancer J Clin* (2017) 67(1):7–30. doi: 10.3322/caac.21387
2. Global Burden of Disease Cancer C, Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol* (2017) 3(4):524–48. doi: 10.1001/jamaoncol.2016.5688
3. Liu C, Liu T, Zhang N, Liu Y, Li N, Du P, et al. (68)Ga-PSMA-617 PET/CT: A Promising New Technique for Predicting Risk Stratification and Metastatic Risk of Prostate Cancer Patients. *Eur J Nucl Med Mol Imaging* (2018) 45 (11):1852–61. doi: 10.1007/s00259-018-4037-9
4. Cooperberg MR, Broering JM, Carroll PR. Time Trends and Local Variation in Primary Treatment of Localized Prostate Cancer. *J Clin Oncol* (2010) 28 (7):1117–23. doi: 10.1200/JCO.2009.26.0133
5. Heidenreich A, Bastian PJ, Bellmunt J, Bolla M, Joniau S, van der Kwast T, et al. EAU Guidelines on Prostate Cancer. Part II: Treatment of Advanced, Relapsing, and Castration-Resistant Prostate Cancer. *Eur Urol* (2014) 65 (2):467–79. doi: 10.1016/j.eururo.2013.11.002
6. Zhang J, Shao S, Wu P, Liu D, Yang B, Han D, et al. Diagnostic Performance of (68)Ga-PSMA PET/CT in the Detection of Prostate Cancer Prior to Initial Biopsy: Comparison With Cancer-Predicting Nomograms.

Eur J Nucl Med Mol Imaging (2019) 46(4):908–20. doi: 10.1007/s00259-018-4255-1
7. Sanda MG, Cadeddu JA, Kirkby E, Chen RC, Crispino T, Fontanarosa J, et al. Clinically Localized Prostate Cancer: AUA/ASTRO/SUO Guideline. Part I: Risk Stratification, Shared Decision Making, and Care Options. *J Urol* (2018) 199(3):683–90. doi: 10.1016/j.juro.2017.11.095
8. Lam TBL, MacLennan S, Willemse PM, Mason MD, Plass K, Shepherd R, et al. EAU-EANM-ESTRO-ESUR-SIOG Prostate Cancer Guideline Panel Consensus Statements for Deferred Treatment With Curative Intent for Localised Prostate Cancer From an International Collaborative Study (DETECTIVE Study). *Eur Urol* (2019) 76(6):790–813. doi: 10.1016/j.eururo.2019.09.020
9. Semjonow A, Brandt B, Oberpenning F, Roth S, Hertle L. Discordance of Assay Methods Creates Pitfalls for the Interpretation of Prostate-Specific Antigen Values. *Prostate Suppl* (1996) 7:3–16. doi: 10.1002/(SICI)1097-0045(1996)7+<3::AID-PROS1>3.0.CO;2-T
10. Gerstenbluth RE, Seftel AD, Hampel N, Oefelein MG, Resnick MI. The Accuracy of the Increased Prostate Specific Antigen Level (Greater Than or Equal to 20 Ng./Ml.) in Predicting Prostate Cancer: Is Biopsy Always Required? *J Urol* (2002) 168(5):1990–3. doi: 10.1097/01.ju.0000033330.06269.6d
11. Li Y, Han D, Wu P, Ren J, Ma S, Zhang J, et al. Comparison of (68)Ga-PSMA-617 PET/CT With mpMRI for the Detection of PCa in Patients With a PSA Level of 4-20 Ng/Ml Before the Initial Biopsy. *Sci Rep* (2020) 10(1):10963. doi: 10.1038/s41598-020-67385-9
12. Roobol MJ, Steyerberg EW, Kranse R, Wolters T, van den Bergh RC, Bangma CH, et al. A Risk-Based Strategy Improves Prostate-Specific Antigen-Driven Detection of Prostate Cancer. *Eur Urol* (2010) 57(1):79–85. doi: 10.1016/j.eururo.2009.08.025
13. Silver DA, Pellicer I, Fair WR, Heston WD, Cordon-Cardo C. Prostate-Specific Membrane Antigen Expression in Normal and Malignant Human Tissues.. *Clin Cancer Res* (1997) 3(1):81–5.
14. Evans MJ, Smith-Jones PM, Wongvipat J, Navarro V, Kim S, Bander NH, et al. Noninvasive Measurement of Androgen Receptor Signaling With a Positron-Emitting Radiopharmaceutical That Targets Prostate-Specific Membrane Antigen. *Proc Natl Acad Sci USA* (2011) 108(23):9578–82. doi: 10.1073/pnas.1106383108
15. Sweat SD, Pacelli A, Murphy GP, Bostwick DG. Prostate-Specific Membrane Antigen Expression Is Greatest in Prostate Adenocarcinoma and Lymph Node Metastases. *Urology* (1998) 52(4):637–40. doi: 10.1016/S0090-4295(98)00278-7
16. Marchal C, Redondo M, Padilla M, Caballero J, Rodrigo I, Garcia J, et al. Expression of Prostate Specific Membrane Antigen (PSMA) in Prostatic Adenocarcinoma and Prostatic Intraepithelial Neoplasia. *Histol Histopathol* (2004) 19(3):715–8. doi: 10.14670/HH-19.715
17. Uprimny C, Kroiss AS, Decristoforo C, Fritz J, von Guggenberg E, Kendler D, et al. (68)Ga-PSMA-11 PET/CT in Primary Staging of Prostate Cancer: PSA and Gleason Score Predict the Intensity of Tracer Accumulation in the Primary Tumour. *Eur J Nucl Med Mol Imaging* (2017) 44(6):941–9. doi: 10.1007/s00259-017-3631-6
18. Uprimny C, Kroiss AS, Fritz J, Decristoforo C, Kendler D, von Guggenberg E, et al. Early PET Imaging With [68]Ga-PSMA-11 Increases the Detection Rate of Local Recurrence in Prostate Cancer Patients With Biochemical Recurrence. *Eur J Nucl Med Mol Imaging* (2017) 44(10):1647–55. doi: 10.1007/s00259-017-3743-z
19. Ferraro DA, Muehlematter UJ, Garcia Schuler HI, Rupp NJ, Huellner M, Messerli M, et al. (68)Ga-PSMA-11 PET has the Potential to Improve Patient Selection for Extended Pelvic Lymph Node Dissection in Intermediate to High-Risk Prostate Cancer. *Eur J Nucl Med Mol Imaging* (2020) 47(1):147–59. doi: 10.1007/s00259-019-04511-4
20. Giesel FL, Hadaschik B, Cardinale J, Radtke J, Vinsensia M, Lehnert W, et al. F-18 Labelled PSMA-1007: Biodistribution, Radiation Dosimetry and Histopathological Validation of Tumor Lesions in Prostate Cancer Patients. *Eur J Nucl Med Mol Imaging* (2017) 44(4):678–88. doi: 10.1007/s00259-016-3573-4
21. Kuten J, Fahoum I, Savin Z, Shamni O, Gitstein G, Hershkovitz D, et al. Head-To-Head Comparison of (68)Ga-PSMA-11 With (18)F-PSMA-1007 PET/CT in Staging Prostate Cancer Using Histopathology and Immunohistochemical Analysis as a Reference Standard. *J Nucl Med* (2020) 61(4):527–32. doi: 10.2967/jnumed.119.234187
22. Rauscher I, Maurer T, Fendler WP, Sommer WH, Schwaiger M, Eiber M. (68)Ga-PSMA Ligand PET/CT in Patients With Prostate Cancer: How We Review and Report. *Cancer Imaging* (2016) 16(1):14. doi: 10.1186/s40644-016-0072-6
23. Fendler WP, Eiber M, Beheshti M, Bomanji J, Ceci F, Cho S, et al. (68)Ga-PSMA PET/CT: Joint EANM and SNMMI Procedure Guideline for Prostate Cancer Imaging: Version 1.0. *Eur J Nucl Med Mol Imaging* (2017) 44(6):1014–24. doi: 10.1007/s00259-017-3670-z
24. Chen R, Wang Y, Shi Y, Zhu Y, Xu L, Huang G, et al. Diagnostic Value of (18)F-FDG PET/CT in Patients With Biochemical Recurrent Prostate Cancer and Negative (68)Ga-PSMA PET/CT. *Eur J Nucl Med Mol Imaging* (2021) 48(9):2970-2977. doi: 10.1007/s00259-021-05221-6
25. Jiao J, Kang F, Zhang J, Quan Z, Wen W, Zhao X, et al. Establishment and Prospective Validation of an SUVmax Cutoff Value to Discriminate Clinically Significant Prostate Cancer From Benign Prostate Diseases in Patients With Suspected Prostate Cancer by (68)Ga-PSMA PET/CT: A Real-World Study. *Theranostics* (2021) 11(17):8396–411. doi: 10.7150/thno.58140
26. Koerber SA, Utzinger MT, Kratochwil C, Kesch C, Haefner MF, Katayama S, et al. (68)Ga-PSMA-11 PET/CT in Newly Diagnosed Carcinoma of the Prostate: Correlation of Intraprostatic PSMA Uptake With Several Clinical Parameters. *J Nucl Med* (2017) 58(12):1943–8. doi: 10.2967/jnumed.117.190314
27. Mottaghy FM, Heinzel A, Verburg FA. Molecular Imaging Using PSMA PET/CT Versus Multiparametric MRI for Initial Staging of Prostate Cancer: Comparing Apples With Oranges? *Eur J Nucl Med Mol Imaging* (2016) 43(8):1397–9. doi: 10.1007/s00259-016-3389-2
28. Maurer T, Gschwend JE, Rauscher I, Souvatzoglou M, Haller B, Weirich G, et al. Diagnostic Efficacy of (68)Gallium-PSMA Positron Emission Tomography Compared to Conventional Imaging for Lymph Node Staging of 130 Consecutive Patients With Intermediate to High Risk Prostate Cancer. *J Urol* (2016) 195(5):1436–43. doi: 10.1016/j.juro.2015.12.025
29. Pyka T, Okamoto S, Dahlbender M, Tauber R, Retz M, Heck M, et al. Comparison of Bone Scintigraphy and (68)Ga-PSMA PET for Skeletal Staging in Prostate Cancer. *Eur J Nucl Med Mol Imaging* (2016) 43(12):2114–21. doi: 10.1007/s00259-016-3435-0
30. Sachpekidis C, Baumer P, Kopka K, Hadaschik BA, Hohenfellner M, Kopp-Schneider A, et al. (68)Ga-PSMA PET/CT in the Evaluation of Bone Metastases in Prostate Cancer. *Eur J Nucl Med Mol Imaging* (2018) 45(6):904–12. doi: 10.1007/s00259-018-3936-0
31. Roach PJ, Francis R, Emmett L, Hsiao E, Kneebone A, Hruby G, et al. The Impact of (68)Ga-PSMA PET/CT on Management Intent in Prostate Cancer: Results of an Australian Prospective Multicenter Study. *J Nucl Med* (2018) 59(1):82–8. doi: 10.2967/jnumed.117.197160
32. Afshar-Oromieh A, Babich JW, Kratochwil C, Giesel FL, Eisenhut M, Kopka K, et al. The Rise of PSMA Ligands for Diagnosis and Therapy of Prostate Cancer. *J Nucl Med* (2016) 57(Suppl 3):79S–89S. doi: 10.2967/jnumed.115.170720
33. Simsek DH, Sanli Y, Engin MN, Erdem S, Sanli O. Detection of Metastases in Newly Diagnosed Prostate Cancer by Using 68Ga-PSMA PET/CT and Its Relationship With Modified D'Amico Risk Classification. *Eur J Nucl Med Mol Imaging* (2020) 48(5):1639–49. doi: 10.1007/s00259-020-04995-5
34. Hofman MS, Lawrentschuk N, Francis RJ, Tang C, Vela I, Thomas P, et al. Prostate-Specific Membrane Antigen PET-CT in Patients With High-Risk Prostate Cancer Before Curative-Intent Surgery or Radiotherapy (proPSMA): A Prospective, Randomised, Multicentre Study. *Lancet* (2020) 395(10231):1208–16. doi: 10.1016/S0140-6736(20)30314-7
35. Neal DEJr., Clejan S, Sarma D, Moon TD. Prostate Specific Antigen and Prostatitis. I. Effect of Prostatitis on Serum PSA in the Human and Nonhuman Primate. *Prostate* (1992) 20(2):105–11. doi: 10.1002/pros.2990200205
36. Moser PL, Brunner A, Horninger W, Bartsch G, Mikuz G. Correlation Between Inflammatory Cells (T and B Lymphocytes, Macrophages) in Prostate Biopsies and Elevated PSA Levels in a PSA Screening Population. *Urology* (2002) 59(1):68–72. doi: 10.1016/S0090-4295(01)01493-5
37. Schatteman PH, Hoekx L, Wyndaele JJ, Jeuris W, Van Marck E. Inflammation in Prostate Biopsies of Men Without Prostatic Malignancy or Clinical Prostatitis: Correlation With Total Serum PSA and PSA Density. *Eur Urol* (2000) 37(4):404–12. doi: 10.1159/000020161
38. Berglund E, Maaskola J, Schultz N, Friedrich S, Marklund M, Bergenstrahle J, et al. Spatial Maps of Prostate Cancer Transcriptomes Reveal an Unexplored

Landscape of Heterogeneity. *Nat Commun* (2018) 9(1):2419. doi: 10.1038/s41467-018-04724-5
39. Tu SM, Lin SH, Logothetis CJ. Stem-Cell Origin of Metastasis and Heterogeneity in Solid Tumours. *Lancet Oncol* (2002) 3(8):508–13. doi: 10.1016/S1470-2045(02)00820-3
40. Hupe MC, Philippi C, Roth D, Kumpers C, Ribbat-Idel J, Becker F, et al. Expression of Prostate-Specific Membrane Antigen (PSMA) on Biopsies Is an Independent Risk Stratifier of Prostate Cancer Patients at Time of Initial Diagnosis. *Front Oncol* (2018) 8:623. doi: 10.3389/fonc.2018.00623

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Wang, Zheng, Li, Dong, Liu, Yuan, Gao and Duan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 05 November 2021
doi: 10.3389/fonc.2021.747305

Visualization of microRNA-21 Dynamics in Neuroblastoma Using Magnetic Resonance Imaging Based on a microRNA-21-Responsive Reporter Gene

Guangcheng Bao[1,2,3,4†], Jun Sun[2,3,4,5†], Helin Zheng[1†], Jingxin Hou[6], Jie Huang[1,2,3,4], Jie Wei[1,2,3,4], Yuanqiao Fu[1,2,3,4], Jiawen Qiu[1,2,3,4], Xuefeng Zou[7], Bin Xiang[8] and Jinhua Cai[1,2,3,4]*

[1] Department of Radiology, Children's Hospital of Chongqing Medical University, Chongqing, China, [2] Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing, China, [3] Key Laboratory of Pediatrics in Chongqing, Chongqing, China, [4] Chongqing International Science and Technology Cooperation Center for Child Development and Disorders, Chongqing, China, [5] Department of Radiology, Chongqing University Central Hospital, Chongqing, China, [6] Department of Radiology, Second Affiliated Hospital of Chongqing Medical University, Chongqing, China, [7] Guizhou Provincial Key Laboratory of Computational Nano-Material Science, Guizhou Education University, Guiyang, China, [8] Chemistry and Chemical Engineering, Chongqing University, Chongqing, China

OPEN ACCESS

Edited by:
Guang Yang,
Imperial College London,
United Kingdom

Reviewed by:
Zhipeng Liu,
Purdue University, United States
Arunakumar Gangaplara,
National Institute of Allergy and Infectious Diseases (NIH),
United States

***Correspondence:**
Jinhua Cai
cai_jinhua@126.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 26 July 2021
Accepted: 25 October 2021
Published: 05 November 2021

Citation:
Bao G, Sun J, Zheng H, Hou J, Huang J, Wei J, Fu Y, Qiu J, Zou X, Xiang B and Cai J (2021) Visualization of microRNA-21 Dynamics in Neuroblastoma Using Magnetic Resonance Imaging Based on a microRNA-21-Responsive Reporter Gene. Front. Oncol. 11:747305. doi: 10.3389/fonc.2021.747305

Background: MicroRNAs (miRs) have been shown to be closely associated with the occurrence and development of tumors and to have potential as diagnostic and therapeutic targets. The detection of miRs by noninvasive imaging technology is crucial for deeply understanding their biological functions. Our aim was to develop a novel miR-21-responsive gene reporter system for magnetic resonance imaging (MRI) visualization of the miR-21 dynamics in neuroblastoma.

Methods: The reporter gene ferritin heavy chain (FTH1) was modified by the addition of 3 copies of the sequence completely complementary to miR-21 (3xC_miR-21) to its 3'-untranslated region (3' UTR) and transduced into SK-N-SH cells to obtain SK-N-SH/FTH1-3xC_miR-21 cells. Then, the antagomiR-21 was delivered into cells by graphene oxide functionalized with polyethylene glycol and dendrimer. Before and after antagomiR-21 delivery, FTH1 expression, MRI contrast and intracellular iron uptake were assayed *in vitro* and *in vivo*.

Results: In the SK-N-SH/FTH1-3xC_miR-21 cells, FTH1 expression was in an "off" state due to the combination of intratumoral miR-21 with the 3' UTR of the reporter gene. AntagomiR-21 delivered into the cells bound to miR-21 and thereby released it from the 3' UTR of the reporter gene, thus "switching on" FTH1 expression in a dose-dependent manner. This phenomenon resulted in intracellular iron accumulation and allowed MRI detection *in vitro* and *in vivo*.

Conclusion: MRI based on the miR-21-responsive gene reporter may be a potential method for visualization of the endogenous miR-21 activity in neuroblastoma and its response to gene therapy.

Keywords: microRNA-21, reporter gene, ferritin heavy chain, magnetic resonance imaging, neuroblastoma

INTRODUCTION

MicroRNAs (miRs) are endogenous single-stranded small-molecule RNAs that are approximately 20-25 nucleotides long, do not encode proteins, and have a variety of important regulatory functions in cells (1). miRs can target and bind the 3'-untranslated regions (3' UTRs) of specific messenger RNAs (mRNAs) and induce their translational repression or degradation, thereby decreasing the expression of proteins encoded by the targeted mRNAs. This novel mode of post-transcriptional gene regulation plays an important role in extensive biological processes, such as cell growth, differentiation, apoptosis, tumorigenicity and chemoresistance (2, 3).

Since the first report of the role of miRs in oncogenesis (4), many miRs involved in cancer development have been identified. Among these, some function as onco-miRs by promoting tumor proliferation and angiogenesis, while some are associated with tumor inhibition (5–7). The different functions of these miRs may be related to the type and stage of cancer development. Currently, the expression profiles of miRs and their potential targets have been widely studied in breast (8), lung (9), prostate (10), colon (11), gastric (12) and ovarian (13) cancer. Among various cancer-associated miRs, miR-21 has attracted substantial attention as a potential diagnostic and therapeutic cancer target since it has been reported to be the only carcinogenic miR that is overexpressed in several types of solid tumors, including neuroblastoma, glioblastoma, breast cancer, colorectal cancer and prostate cancer (4, 8, 10, 11, 14). On the other hand, a chemically modified antisense RNA oligonucleotide has been used to target and repress the activity of miR-21, thereby inhibiting tumor development (15, 16).

Considering the potential of miR to serve as a diagnostic and therapeutic biomarker in cancer, investigations of the biological properties and expression patterns of miRs are crucial. The conventional methods for detecting endogenous miRs include northern blotting (17), quantitative reverse-transcriptase polymerase chain reaction (qRT-PCR) (9), small RNA sequencing (18) and microarray (19). However, these methods require cell lysis and are not suitable for detecting miR expression in living cells. Therefore, the development of noninvasive methods is imperative for elucidating the expression and regulation of miRs *in vivo*.

At present, several noninvasive imaging techniques, such as optical imaging, radionuclide imaging and magnetic resonance imaging (MRI), have been applied to detect biological processes *in vivo* at the cellular or molecular level (20, 21). Of these imaging methods, MRI is thought to be ideal due to its advantages, including deep tissue penetration, high tissue resolution, repeatable examination and nonrequirement for radiation. Molecular MRI usually relies on a reporter gene, which can be integrated into the cellular genome and allows for the long-term monitoring of biological events. Among various MRI reporter genes, ferritin heavy chain 1 (FTH1) has attracted substantial attention because its expression can induce the effective cellular uptake of endogenous iron, which leads to sensitive MRI signal contrast (22–28). In our previous studies, we successfully monitored mesenchymal stem cells with FTH1-based MRI *in vitro* and *in vivo* (24, 29). Considering the advantages of FTH1 as an endogenous reporter gene, it is worth investigating the feasibility of FTH1-based MRI to visualize miR activities, and this method could serve as a noninvasive imaging modality for miR detection.

In this study, we designed a novel miR-21-responsive FTH1 expression system in which the reporter gene FTH1 was modified by the addition of 3 copies of the sequence completely complementary to miR-21 (3xC_miR-21) to its 3' UTR. When the reporter gene system was transduced into neuroblastoma cells, in which miR-21 functions as an oncogenic biomarker, its expression was expected to be switched "off" due to the combination of intratumoral miR-21 with the 3' UTR of the reporter gene. An exogenous antisense oligonucleotide (antagomiR-21) was then delivered into the cells and expected to bind with miR-21 in a competing manner, thereby releasing miR-21 from the FTH1 3' UTR and switching "on" the reporter gene expression (**Figure 1**). With the reporter gene system, we aimed to provide a noninvasive and efficient imaging modality for visualizing the activity of miR-21 and its response to gene therapy.

Abbreviations: CCK-8, Cell Counting Kit-8; FTH1, Ferritin heavy chain 1; DMEM, Dulbecco's modified Eagle's medium; FAC, Ferric ammonium citrate; FBS, Fetal bovine serum; FOV, Field of view; GO, Graphene oxide; LV, Lentiviral vector; miRs, MicroRNAs; MRI, Magnetic resonance imaging; mRNAs, Messenger RNAs; PBS, Phosphate-buffered saline; PCR, Polymerase chain reaction; PEG, Polyethylene glycol; qRT-PCR, Quantitative reverse-transcriptase polymerase chain reaction; 3' UTRs, 3'-untranslated regions; 3xC_miR-21, 3 copies of the sequence completely complementary to miR-21; T2WI, T2-weighted imaging; TE, Time of echo; TEM, Transmission electron microscopy; TR, Time of repetition; WB, Western blotting.

MATERIALS AND METHODS

Construction of the miR-21-Responsive FTH1 Expression Vector

The cDNA of human FTH1 (Gene ID: 2495) was obtained *via* polymerase chain reaction (PCR) amplification using the following primers: forward, AACCGTCAGATCGCACCG GTGCCACCATGACGACCGCGTCCACCTC and reverse,

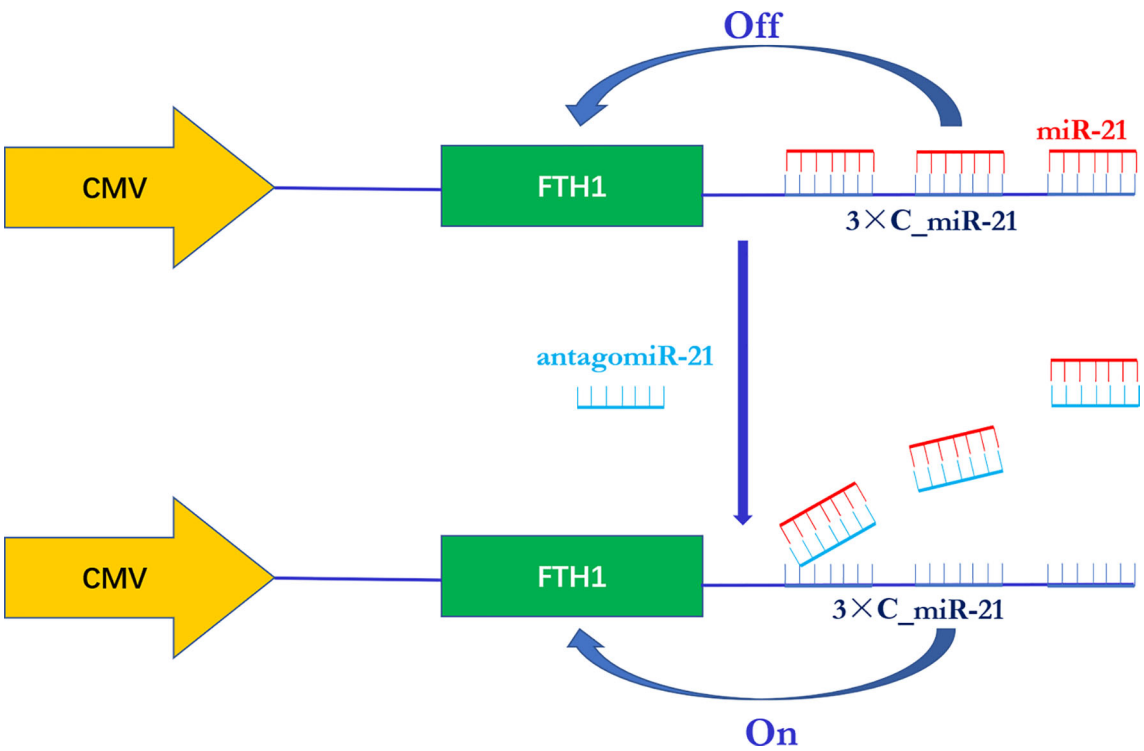


FIGURE 1 | Schematic illustration of the miR-21-responsive FTH1 gene reporter system. The reporter gene FTH1 is modified by the addition of 3 copies of the sequence completely complementary to miR-21 (3xC_miR-21) to its 3'-untranslated region (3' UTR). In the presence of miR-21, FTH1 expression is switched "off" due to the combination of miR-21 with the 3' UTR of the reporter gene. Once an exogenous antisense oligonucleotide (antagomiR-21) is administered, it binds to miR-21 in a competitive manner and releases miR-21 from the 3' UTR of FTH1, thereby switching "on" reporter gene expression.

TCCTTGTAGTCCATGAATTCGCTTTCATTATCACT GTCTC. The synthesized FTH1 cDNA was subcloned into the EcoR I and Sma I restriction endonuclease sites of the plasmid pLVX-mCMV-ZsGreen1-Puro to obtain the recombinant plasmid vector pLVX-mCMV-ZsGreen1-Puro-FTH1 (pLV-FTH1). Then, the DNA oligonucleotides containing 3xC_miR-21 (sense 5' GTAGCTTATCAGACTGATGTTGATA GTAGTAGCTTATCAGACTGATGTTGATAGTA GTAGCTTATCAGACTGATGTTGA 3'; antisense 5' TCAACATCAGTCTGATAAGCTACTACTATCAAC ATCAGTCTGATAAGCTACTACTATCAACATCAGT CTGATAAGCTAC 3') were synthesized and ligated into pLV-FTH1 to yield the final lentiviral vector plasmid containing 3xC_miR-21 (pLV-FTH1-3xC_miR-21). After verification by PCR analysis and DNA sequencing, the plasmid pLV-FTH1-3xC_miR-21 was cotransfected together with psPAX2 and pMD2.G into 293 T packaging cells (Invitrogen, Carlsbad, CA, USA) to produce the lentivirus containing FTH1-3xC_miR-21.

Cell Culture and Establishment of Stable Transgenic Cells

The wild-type human neuroblastoma cell line (SK-N-SH/WT) was provided by the Chongqing Institute of Pediatric Medicine Research and cultured in Dulbecco's modified Eagle's medium (DMEM; Gibco, Grand Island, NY, USA) containing 10% fetal bovine serum (FBS; Gibco, Grand Island, NY, USA) and 0.5% streptomycin at 37°C in a moist environment containing 5% CO2. According to the lentivirus transfection protocol, cells at 30-40% confluence were transfected with the FTH1-3xC_miR-21 lentivirus and screened with 10 μg/ml puromycin. Finally, the surviving clones were isolated and named SK-N-SH/FTH1-3xC_miR-21 cells, with SK-N-SH/WT cells and SK-N-SH/FTH1 (without 3xC_miR-21) cells serving as controls.

Assessment of Cell Growth

To investigate the effect of gene transduction on cell viability, cell growth was assayed by Cell Counting Kit-8 (CCK-8) colorimetry (Beyotime, Nanjing, Jiangsu, China). SK-N-SH/WT, SK-N-SH/FTH1 and SK-N-SH/FTH1-3×C_miR-21 cells were cultured in 96-well plates for 72 h, and the relative changes in the three cell lines were determined.

AntagomiR-21 Transfection Into Cells

The antisense RNA oligonucleotide, antagomiR-21, was provided by GenePharma, Shanghai, China. The sequence of antagomiR-21 was as follows: 5′-GUCAACAUCAGUCUG AUA AGCUA-3.

For improvement of the transfection efficiency of antagomiR-21, graphene oxide (GO) modified with polyethylene glycol

(PEG) and polyamidoamine dendrimer (GO-PEG-dendrimer) was used as gene delivery carrier. GO was synthesized by previously reported methods (30–35), and the detailed protocol for synthesizing GO-PEG-dendrimer and its characterization are shown in the **Supplementary Material**.

The efficiency of the GO-PEG-dendrimer to deliver antagomiR-21 could be affected by the GO/PEG ratio. Too little GO could reduce its capacity of combining with antagomir-21 and lead to a low transducing efficiency, while too little PEG could decrease the biocompatibility which also result a low transducing efficiency. To obtain the optimal structure of GO-PEG-dendrimer for efficiently delivering anagomiR-21 into SK-N-SH cells, we compared the transducing efficiency of the GO-PEG-dendrimer complex at various GO/PEG ratios. The GO-PEG-dendrimer at GO/PEG molar ratios of 5:1, 1:5 and 1:1 was mixed with 30 pmol of anagomiR-21 modified by the fluorescence dye Cy3 (anagomiR-21-Cy3) to prepare GO-PEG-dendrimer/anagomiR-21-Cy3 complexes. Then, 2 μL of siRNA-Mate (GenePharma, Shanghai, China), a commonly used small-molecule RNA transfection agent, was mixed with 30 pmol of antagomiR-21-Cy3 to obtain the siRNA-Mate/antagomiR-21-Cy3 complex as a control. The obtained complexes were incubated with SK-N-SH cells for 6 h, washed 3 times, fixed with 4% formaldehyde and stained with 4',6-diamidino-2-phenylindole (DAPI). Fluorescence images were obtained by a fluorescence microscope (Nikon A1R, Tokyo, Japan).

A CCK-8 assay was used to compare the cytotoxicity of the GO-PEG-dendrimer complexes and siRNA-Mate. SK-N-SH cells were divided into two groups, cultured in 96-well plates at a density of 6×10^3 cells per well, and cocultured with the GO-PEG-dendrimer (GO/PEG ratio 1:1) or siRNA-Mate at equal concentrations for 72 h. Then, the absorbance was measured every 12 h after the addition of CCK-8 reagent.

Western Blot and qRT-PCR Assays

FTH1 expression in SK-N-SH/WT, SK-N-SH/FTH1 and SK-N-SH/FTH1-3×C_miR-21 cells was detected by qRT-PCR. Briefly, total RNA was extracted using a TaKaRa MiniBEST Universal RNA Extraction Kit (TaKaRa®, Kyoto, Japan) and then reverse transcribed into cDNA using HiScript®II Q RT SuperMix (Vazyme, Nanjing, China) for qPCR, which was carried out using AceQ® qPCR SYBR Green Master Mix (Vazyme) on an AB Step One plus Real Time PCR System (Applied Biosystems AB, Waltham Mass, USA). GAPDH expression served as the quantitative internal control for FTH1, and each experiment was performed in triplicate. The primers for GAPDH and FTH1 detection were as follows: GAPDH-F AGAAGGCTGGG GCTCATTTG and GAPDH-R AGGGGCCATCC ACAGTCTTC; FTH1-F GGAATTCATGACGACCGC GTCCAC and FTH1-R CCCCGGGAGCTTTCATTAT CACTGTCTCCC.

Simultaneously, proteins were extracted from the three cell lines and prepared for Western blot (WB) analysis. Briefly, the cells were cleaved in cleavage buffer (Sigma, St. Louis, Missouri, USA) containing protease inhibitors, phosphatase inhibitors and 100 mM benzene sulfonyl fluoride, and the total protein concentration in the sample was determined by the BCA method. Thirty micrograms of protein from each sample was then added to a 12% sodium dodecyl sulfate-polyacrylamide gel and transferred onto a polyvinylidene fluoride (PVDF) membrane (Bio-Rad, California, USA). For detection of FTH1, the membrane was blocked with 5% bovine serum albumin (BSA; Beyotime) for 1 h and then incubated with a primary antibody that specifically recognized FTH1 (rabbit anti-FTH1, 1:1000; Abcam, Cambridge, England) or GAPDH (rabbit anti-FTH1, Abcam) overnight at 4°C. The membrane was then washed with Tris-buffered saline containing Tween-20 and incubated with a secondary antibody (goat anti-rabbit IgG, 1:5000; Sigma) for 2 h. The protein bands were observed using an enhanced chemiluminescence kit (Sigma). The relative expression of FTH1 in SK-N-SH/FTH1-3C×_miR-21 cells was normalized to that of GAPDH, and semiquantitative analysis was carried out based on the band strength.

For determination of the correlation between FTH1 expression and the amount of antagomiR-21, SK-N-SH/FTH1-3×C_miR-21 cells transfected with antagomiR-21 at different concentrations (0, 10, 10, 20, 30, 40, 40, 50, or 60 nmol/L) were cultured for 24 h, and cellular proteins were collected and prepared for WB analysis.

MRI of Cell Pellets

To assess the effect of FTH1 expression on cellular iron transfer, we cultured SK-N-SH/FTH1-3C×_miR-21 cells in medium containing 500 μmol/L ferric ammonium citrate (FAC; Sigma) and then transfected them with antagomiR-21 at the same concentration gradient as that used for the WB experiment. After 24 h, all the treated cells were washed thoroughly with phosphate-buffered saline (PBS) to remove the free FAC, digested with ethylenediaminetetraacetic acid (EDTA), suspended in PBS, and then transferred into a 0.6-mL Eppendorf tube to prepare the cell phantom for MRI *in vitro*. The cellular phantom was imaged using a 7.0 T MR scanner (Bruker, Karlsruhe, Germany). The spin echo (SE) T2 weighted imaging (T2WI) parameters were as follows: time of repetition (TR), 2,500 ms; time of echo (TE), 35 ms; field of view (FOV), 120 mm×120 mm; slice thickness, 1 mm; slice interval, 0.1 mm. The parameters for the multiecho sequence were as follows: TR, 2,000 ms; TE, 8~200 ms with a step size of 8 ms (25-point T2 mapping); other parameters, including the FOV, matrix and slice thickness, matched those used for T2WI imaging. T2 maps were obtained by image postprocessing, and the R2 value was measured.

To evaluate FTH1 expression and its effect on iron transfer in groups of cells treated with or without antagomir-21, SK-N-SH/WT, SK-N-SH/FTH1 and SK-N-SH/FTH1-3×C_miR-21 cells were cultured with antagomir-21 at the optimal concentration and 500 μmol/L FAC for 24 h. Before and after antagomiR-21 treatment, the three cell lines were subjected to MRI examination.

Intracellular Iron Detection and Quantification

Prussian blue staining and transmission electron microscopy (TEM) were performed to detect intracellular iron accumulation. SK-N-SH/WT, SK-N-SH/FTH1 and SK-N-SH/FTH1-3×C_miR-21 cells transfected with or without antagomir-21 were cultured together with 500 μmol/L FAC for 24 h. Then, the cells were harvested and subjected to Prussian blue staining and TEM examination according to our previous protocols (24). The intracellular accumulation of iron was observed under a light microscope (Nikon, Tokyo, Japan) or an H-7500 transmission electron microscope (Hitachi, Tokyo, Japan). For quantification of the intracellular iron content, 1×10^6 cells per group were prepared according to the previous protocols (36). The iron concentration was measured using an atomic absorption spectrophotometer (Huaguang HG-960 2A, Shenyang, China). Each sample was measured 3 times. The concentration values are presented in units of pg/cell.

In Vivo Experiments

The experimental animals were purchased from the Department of Medical Experimental Animals of Chongqing Medical University and raised at the Experimental Animal Center of our hospital. All animal protocols used in this study were reviewed and approved by the Animal Care and Use Committee of Chongqing Medical University, and the experimental procedures were performed in accordance with the National Institutes of Health guidelines. All efforts were made to minimize animal suffering.

The nude mice were divided into the following 3 groups: SK-N-SH/WT, SK-N-SH/FTH1 and SK-N-SH/FTH1-3C×_miR-21. A total of 2×10^6 cells were inoculated subcutaneously into the right flanks of nude mice in each of the three groups to establish the transplanted tumor model of neuroblastoma. To enhance iron recruitment into cell xenografts, the animals were treated with FAC at a dose of 5 mg/L *via* their drinking water. When the tumors reached approximately 300 mm^3 in size, the GO-PEG dendrimer/antagomiR-21 complex was administered *via* caudal vein injection at an antagomiR-21 dose of 30 nmol/g according to the manufacturer's instructions. Before and 24 h after antagomiR-21 treatment, the animals were anesthetized and subjected to MRI examination. A 7.0 T MRI scanner was applied, and MR images were obtained by using the T2WI and multiecho sequences. The T2WI scanning parameters were as follows: TR, 2,500 ms; TE, 35 ms. The multiecho parameters were as follows: TR, 2,000 ms; TE, 8~200 ms with a step size of 8 ms (25-point T2 mapping); matrix, 380×300; FOV, 160 mm; and slice thickness, 1.2 mm. The R2 value was measured on T2 maps obtained by image postprocessing.

Immediately after MRI, the animals were sacrificed, and the masses were removed for histological examination. Prussian blue staining was also performed to detect iron accumulation in the tumors. In addition, the intratumoral iron content was quantified according to the protocols used for intracellular iron quantification and are presented in units of mg/g.

Statistical Analysis

All data are expressed as the mean ± standard deviation. Statistical Package for the Social Sciences version 13.0 (SPSS Inc., Chicago, IL, USA) was used for the statistical analyses. One-way analysis of variance and the least significant difference method were used to compare differences among the groups. *P* values less than 0.05 were considered statistically significant.

RESULTS

The miR-21-Responsive FTH1 Reporter Gene System and Its Effect on Cell Viability

The miR-21-responsive FTH1 gene reporter system, namely, FTH1-3×C_miR-21, which contained 3 copies of a sequence completely complementary to miR-21 at the 3' UTR of the FTH1 gene, was successfully constructed. This reporter gene system was transduced into SK-N-SH cells, and SK-N-SH/FTH1-3×C_miR-21 cells were established.

To verify the activity of the reporter gene system, we evaluated FTH1 expression in the SK-N-SH/WT, SK-N-SH/FTH1 and SK-N-SH/FTH1-3×C_miR-21 cells by qRT-PCR and WB. The qRT-PCR results showed that the mRNA expression levels of the FTH1 gene in the SK-N-SH/FTH1 and SK-N-SH/FTH1-3×C_miR-21 groups were similar and both were significantly higher than that in SK-N-SH/WT group ($P < 0.0001$). The WB results showed that the protein level of FTH1 was increased significantly ($P < 0.0001$) in the SK-N-SH/FTH1 group but not altered significantly in the SK-N-SH/FTH1-3×C_miR-21 group compared with the SK-N-SH/WT group (**Figures 2A, B**). These results indicated that miR-21 inhibited FTH1 expression at the post-transcriptional level.

To assess the safety of the reporter gene system, cell viability was evaluated. Morphologically, no changes were observed among the SK-N-SH/WT, SK-N-SH/FTH1 and SK-N-SH/FTH1-3×C_miR-21 cells (**Figure 2C**). The CCK-8 assay results also showed no significant differences in cell proliferation among the three groups (**Figure 2D**).

Gene Transfection Efficiency and Cytotoxicity of the GO-PEG Dendrimer

Immunofluorescence staining and fluorescence intensity quantification showed that the GO-PEG dendrimer with various GO/PEG ratios differentially introduced antagomir-21 into cells. Among the three structures with GO/PEG ratios of 1:5, 1:1 and 5:1, the GO-PEG dendrimer with a GO/PEG ratio of 1:1 was the most efficient and was more efficient than the commonly used transfection reagent siRNA-Mate (**Figures 3A, B**).

To study the cytotoxicity of the GO-PEG-dendrimer, we evaluated the survival rate of SK-N-SH cells. As shown in **Figure 3C**, the cytotoxicity of the GO-PEG-dendrimer was much lower than that of siRNA-Mate, indicating that GO-PEG-dendrimer is safe as a gene delivery vector.

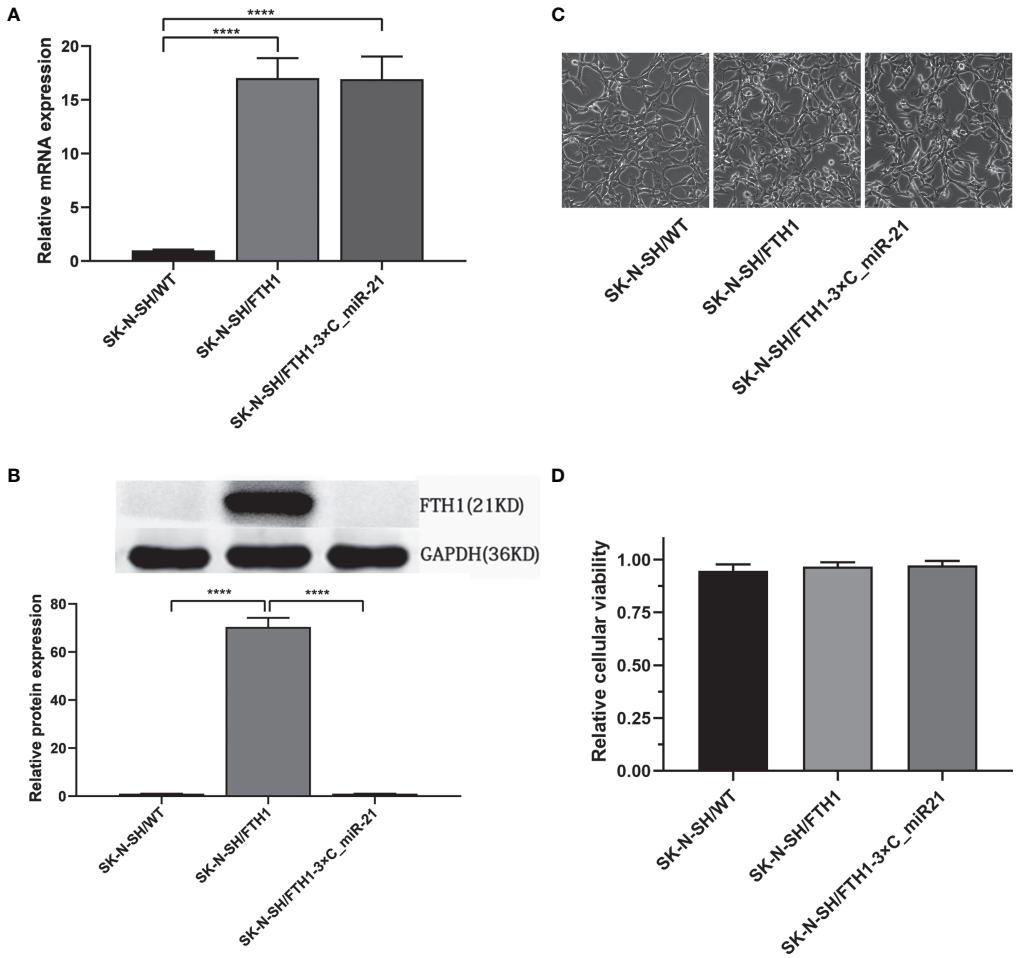


FIGURE 2 | FTH1 expression and its effect on viability in the three groups of cells. The qRT-PCR results **(A)** showed that the mRNA expression levels of the FTH1 gene in the SK-N-SH/FTH1 and SK-N-SH/FTH1-3 × C_miR-21 groups was similar, and both groups had significantly higher levels than the SK-N-SH/WT group. The WB results **(B)** showed that compared with the SK-N-SH/WT group, the SK-N-SH/FTH1 group showed significantly increased FTH1 expression at the protein level, while the SK-N-SH/FTH1-3×C_miR-21 group did not show alterations in FTH1 expression. No morphological changes were observed among the SK-N-SH/WT, SK-N-SH/FTH1 and SK-N-SH/FTH1-3×C_miR-21 cells **(C)**. The CCK-8 results revealed no difference in cell viability among the three groups **(D)**. Three independent experiments were performed. **** indicates $P < 0.0001$.

Dose-Dependent FTH1 Expression Regulated by AntagomiR-21

To investigate the ability of antagomiR-21 to knock down miR-21 and reverse FTH1 expression in SK-N-SH/FTH1-3×C_miR-21 cells, we carried out WB to evaluate FTH1 expression in the cells treated with different concentrations of antagomir-21. AntagomiR-21 delivery into cells reversed FTH1 expression in a dose-dependent manner. In the absence of antagomiR-21, no obvious FTH1 expression was detected in the SK-N-SH/FTH1-3×C_miR-21 cells in which miR-21 suppressed the reporter gene by binding with the completely complimentary sequence of 3×C_miR-21. After transfection with antagomir-21, FTH1 expression increased gradually as the antagomir-21 concentration increased, peaking at 40 nmol/L antagomir-21, and then slightly decreased as the antagomiR-21 continuously increased (**Figures 4A, B**). The cellular MRI T2WI and R2 measurements of cells treated with different concentrations of antagomiR-21 were consistent with the WB results (**Figures 4C–E**).

Cellular MRI Contrast Produced by the Regulatory Expression of FTH1

The expression of FTH1 may increase the ability of cells to internalize and store iron ions, thereby decreasing the MRI signal. To assess the feasibility of using MRI to detect miR-21 activity regulated by antagomiR-21, we first determined the optimal concentration (40 nmol/L) at which antagomiR-21 induced the most obvious MRI signal change. Then, MRI was performed in 3 groups of cells treated with the optimal concentration of antagomiR-21: SK-N-SH/WT, SK-N-SH/FTH1 and SK-N-SH/FTH1-3×C_miR-21 cells. A significant signal reduction was observed in only the SK-N-SH/FTH1-

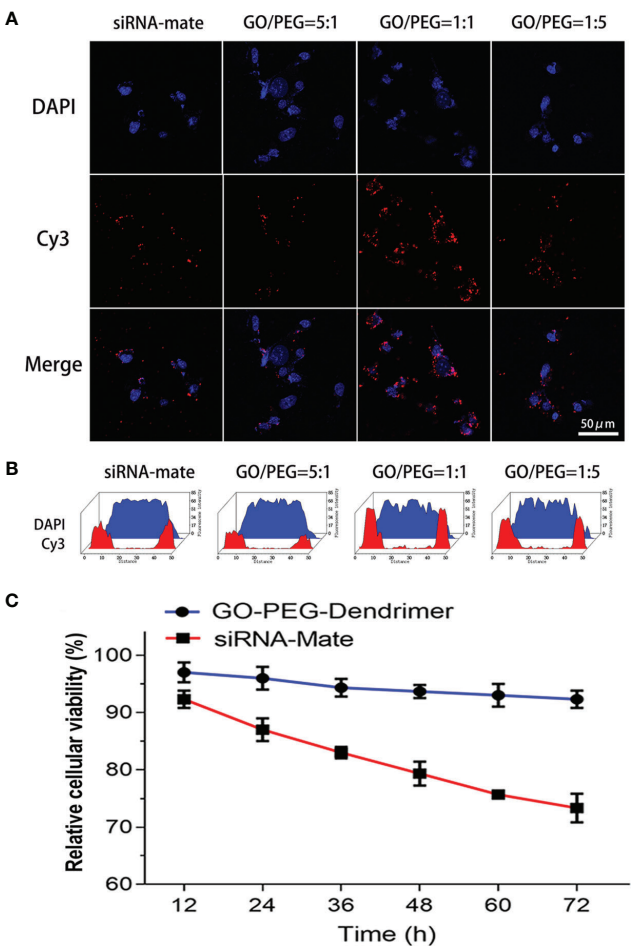


FIGURE 3 | Transfection efficiency and cytotoxicity of the GO-PEG-dendrimer. Both the immunofluorescence staining **(A)** and fluorescence intensity quantification **(B)** results show that among the three GO-PEG-dendrimer structures with different GO/PEG ratios, the structure with a GO/PEG ratio of 1:1 had the highest transfection efficiency and was more efficient than the commonly used transfection reagent siRNA-Mate. The other two structures with GO/PEG ratios of 5:1 and 1:5 had transfection efficiencies that were similar to that of the siRNA-Mate. The CCK-8 assay **(C)** showed that the GO-PEG-dendrimer was much less cytotoxic than siRNA-Mate during the 72 h of coculture with SK-N-SH cells.

3×C_miR-21 cells after antagomiR-21 transfection compared with the signal prior to antagomiR-21 treatment ($P<0.05$). There were no changes in the MRI signals change before and after antagomiR-21 treatment in the other two groups of cells (**Figures 5A, B**).

The Prussian blue staining and TEM results were consistent with the MRI findings. Before transfection of antagomir-21, large amounts of iron particles were detected in the cytoplasms of only SK-N-SH/FTH1 cells, while no iron accumulation was observed in SK-N-SH/WT or SK-N-SH/FTH1-3×C_miR-21 cells. After antagomiR-21 treatment, SK-N-SH/FTH1-3×C_miR-21 cells exhibited obvious iron accumulation compared with that prior to antagomiR-21 treatment. In the other two groups of cells, however, there was no change in iron accumulation before and after antagomiR-21 treatment (**Figures 5C, D**). Quantification of the intracellular iron content (**Figure 5E**) showed that the iron content in the SK-N-SH/FTH1-3×C_miR-21 cells was 0.96 ± 0.17 pg/cell after antagomiR-21 treatment, significantly higher than that before antagomiR-21 treatment (0.12 ± 0.02 pg/cell) ($P<0.05$). There was no difference of iron content before and after antagomiR-21 treatment in the other two groups.

The above results suggest that miR-21 in SK-N-SH/FTH1-3×C_miR-21 cells can be inhibited by antagomiR-21 and reboot the expression of FTH1, thereby resulting in intracellular iron accumulation. This process can be detected by MRI signal changes.

In Vivo MRI of AntigomiR-21 Delivery

Animal tumor models of three SK-N-SH cell lines were successfully established. Before administration of the GO-PEG-dendrimer/antagomiR-21 complex, the MRI signal was obviously decreased and the R2 value was significantly increased in the SK-N-SH/FTH1 group ($P<0.05$) compared with the SK-N-SH/WT group, while those in the SK-N-SH/FTH1-3×C_miR-21 group did not differ from those in the SK-N-SH/WT group. After treatment with the complex, the SK-N-SH/FTH1-3×C_miR-21 group showed a significantly decreased MRI signal and increased R2 value compared with those prior to treatment ($P<0.05$). In the other two groups, the MRI signals and R2 values did not differ before and after antagomiR-21 treatment (**Figures 6A, B**).

Further intratumoral iron detection and quantification matched the MRI results. Intratumoral iron content was significantly increased after antagomiR-21 treatment in only the SK-N-SH/FTH1-3×C_miR-21 group compared with that prior to treatment (**Figures 6C, D**).

DISCUSSION

Increasing evidence has demonstrated that miRs are involved in tumorigenesis and tumor progression *in vitro* and *in vivo*, which provides a new pathway for exploring the potential of miRs as biomarkers for tumor diagnosis and/or therapy (2, 3, 15, 37). In this study, we designed a novel gene reporter system which contained three copies of the sequence completely complimentary to miR-21 in the 3' UTR of FTH1 for the MRI visualization of miR-21 in neuroblastoma. The results showed that the miR-21-responsive reporter gene was sensitively regulated by miR-21 in neuroblastoma cells. Compared with the regular FTH1 reporter gene, the miR-21-responsive reporter gene was "switched off" by the combination of intratumoral miR-21 and the 3' UTR of FTH1. AntagomiR-21 induced exogenously "switched on" FTH1 expression in the cells by binding to miR-21 and thereby promoting its release from the FTH1 3' UTR. FTH1 expression induced a sufficient amount of intracellular iron accumulation for visualization by MRI *in vitro* and *in vivo*. In addition, the reporter gene expression reversed by antagomiR-21 was dose-dependent, suggesting that miR-21 activity can be

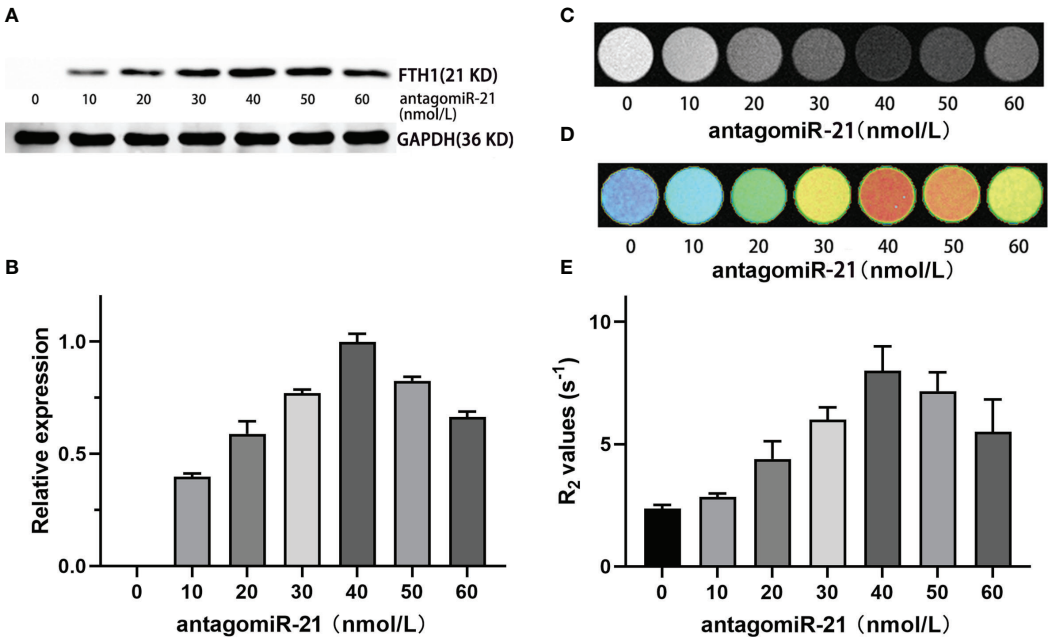


FIGURE 4 | Regulation of FTH1 expression by antagomiR-21. WB **(A)** and gray-scale **(B)** results show that the delivery of antagomiR-21 into cells reversed FTH1 expression in a dose-dependent manner. In the absence of antagomiR-21, no obvious FTH1 expression was detected in SK-N-SH/FTH1-3 × PT cells in which miR-21 suppressed the reporter gene by binding with the completely complimentary sequence of 3×C_miR-21. After transfection with antagomir-21, the expression of FTH1 increased gradually as the antagomir-21 concentration increased, peaking at 40 nmol/L, and then slightly decreased as the antagomiR-21 concentration continuously increased. The cellular MRI T2WI **(C)**, T2 map **(D)** and R2 measurements **(E)** in cells treated with different concentrations of antagomiR-21 were consistent with the WB results.

quantified by MRI. To the best of our knowledge, this is the first report of using an MRI-based gene reporter system to study miR-21 activity in tumors. Our research provides a potential noninvasive imaging method for evaluating miR-21 as a diagnostic or therapeutic target.

Recently, optical imaging techniques based on luciferase or bioluminescent reporter genes have been applied in the miR research field and been shown to serve as noninvasive methods for detecting and quantifying miRs *in vivo*. Ko et al. (38) developed a reporter gene system, CMV/Gluc/3×PT-miR-124, and transduced it into stem cells, and the results showed miR-124a-dependent Gluc reporter expression during neural differentiation. Kim et al. (39) utilized a CMV/Gluc-3xPT_miR221 reporter system to image miR-221 in papillary thyroid carcinoma and showed that Gluc activity was regulated according to the miR-221 levels *in vitro* and *in vivo*. These optical imaging reporter gene systems have greatly expanded our understanding of miR activities *in vivo*. However, optical imaging modalities can be used only for studies in small animals or surface tissues due to their poor tissue penetration, which has limited their clinical application. In the present study, we utilized MRI to detect miR-21 activity and achieved results consistent with those of optical imaging studies. In addition, compared with optical imaging, MRI can better penetrate deep tissue, which may make it more applicable in the miR imaging research field and more feasible for clinical translation.

Employing antisense oligonucleotides to antagonize the functions of target miRs *in vivo* has been proposed as a potential gene therapy for various diseases (40–42). This therapeutic strategy, however, has largely been impeded due to the lack of effective gene delivery carriers. In this study, we constructed a polyamidoamine dendrimer and PEG-functionalized GO conjugate for the delivery of antagomiR-21 into SK-N-SH cells. The efficiency of the GO-PEG-dendrimer to deliver antagomiR-21 was affected by the GO/PEG ratio. Among the three GO/PEG ratio groups, the GO-PEG dendrimer with a GO/PEG ratio of 1:1 delivered antagomiR-21 more efficiently than the commonly used transfection agent siRNA-Mate. The other two structures, which had GO/PEG ratios of 5:1 and 1:5, had delivery efficiencies that were similar to that of siRNA-Mate. We speculate that the decreased transfection efficiency of the GO-PEG-dendrimer with a GO/PEG ratio of 5:1 was due to the decreased biocompatibility of GO-PEG dendrimers. The lower transfection efficiency of the GO-PEG-dendrimer with a GO/PEG ratio of 1:5 may be explained by the reduced capacity of the GO to combine with antagomir-21. Based on the optimal GO/PEG ratio, we further compared the cytotoxicities of GO-PEG-dendrimer and siRNA-Mate. The cell viability of the SK-N-SH cells transfected with the GO-PEG-dendrimer was higher than that of the cells transfected with siRNA-Mate during 72 h of continuous culture, suggesting that the GO-PEG-dendrimer is a safe gene delivery vector. In addition, the *in vitro* and *in vivo*

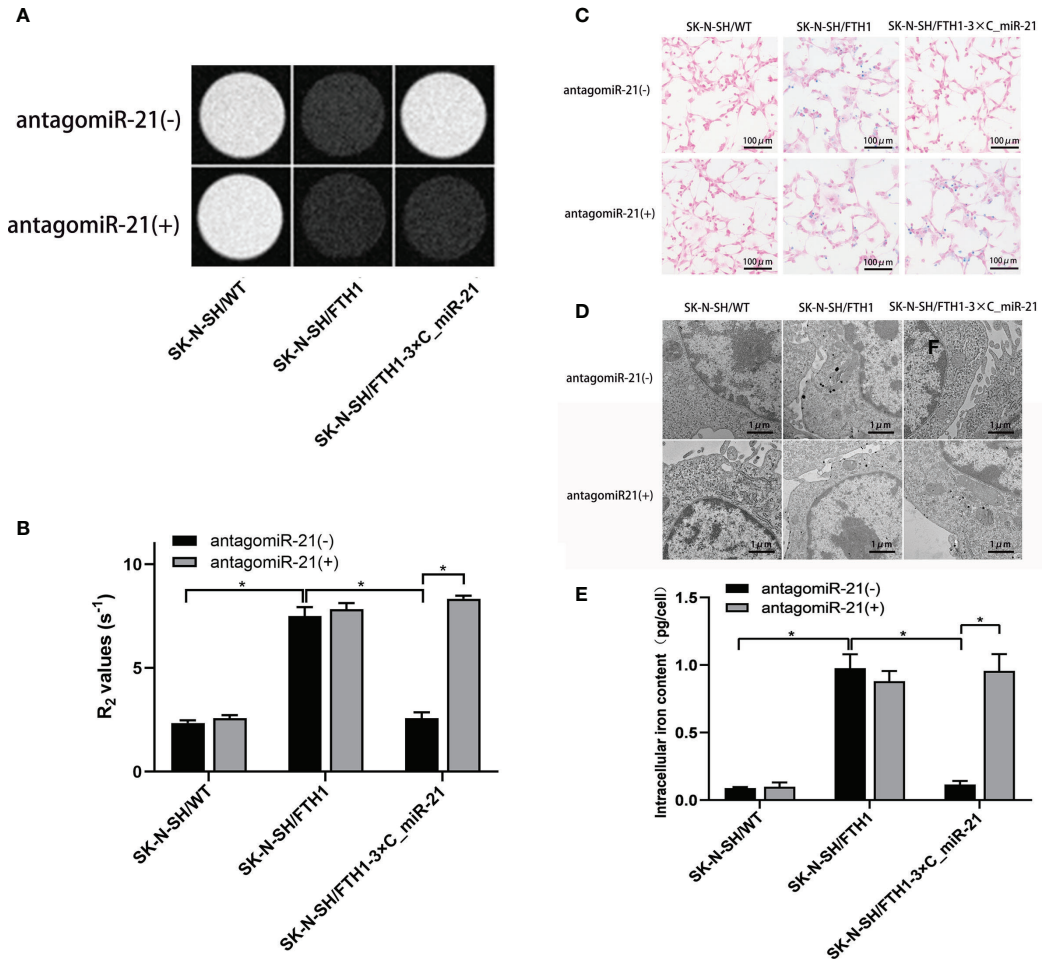


FIGURE 5 | Cellular MRI and intracellular iron detection. MRI imaging shows a significant T2WI signal reduction **(A)** and R2 value increase **(B)** in the SK-N-SH/FTH1-3×C_miR-21 cells after antagomiR-21 transfection compared with that prior to antagomiR-21 treatment. In the SK-N-SH/WT and SK-N-SH/FTH1 cells, however, there were no differences in the MRI signals or R2 values before and after antagomiR-21 treatment. The Prussian blue staining **(C)** and TEM **(D)** results revealed that before antagomiR-21 transfection, large amounts of iron particles were detected in the cytoplasm of only SK-N-SH/FTH1 cells, while iron accumulation was not observed in SK-N-SH/WT or SK-N-SH/FTH1-3×C_miR-21 cells. After antagomiR-21 treatment, SK-N-SH/FTH1-3×C_miR-21 cells exhibited obvious iron accumulation compared with that prior to antagomiR-21 treatment. In the other two groups of cells, however, there was almost no change in the iron accumulation before and after antagomiR-21 treatment. Further iron content quantification **(E)** shows the same results as observed in Prussian blue staining and TEM. Three independent experiments were performed. * indicates $P < 0.05$.

experiments revealed that the antagomiR-21 delivered by the GO-PEG-dendrimer definitely induced FTH1 expression by antagonizing miR-21, thereby indirectly confirming the GO-PEG-dendrimer efficiency as a gene delivery vector.

Although the miR-21-responsive FTH1 gene reporter system was proven to have potential for the MRI assessment of miR-21 function in tumor cells, there are several drawbacks in this study. First, the reporter gene system was transduced into the tumor cells in advance to evaluate the activity of miR-21, and this method is limited to experimental studies and is not feasible for clinical applications. To use this gene reporter system to detect intracellular miR-21 for diagnostic purposes in humans, the system should be delivered into tumor cells by an efficient gene carrier to examine the results in real time. Second, to increase reporter gene expression-induced intracellular iron uptake to a level that is sufficient to generate MRI contrast *in vivo*, exogenous iron was administered to the animal at a certain concentration. At present, it could be difficult to produce an obvious MRI signal change depending merely on endogenous iron recruitment. With the development of MRI equipment and technology, we believe that the endogenous iron concentration would provide enough the intracellular iron accumulation for MRI contrast production, which are needed to explore in further studies. Third, the MRI examination was performed only once after antagomir-21 delivery *in vivo* in consideration of animal welfare and fast tumor necrotization. A more suitable tumor xenograft model should be established to allow MRI examinations at more time points, which may enable a longitudinal MRI monitoring study of miR functions.

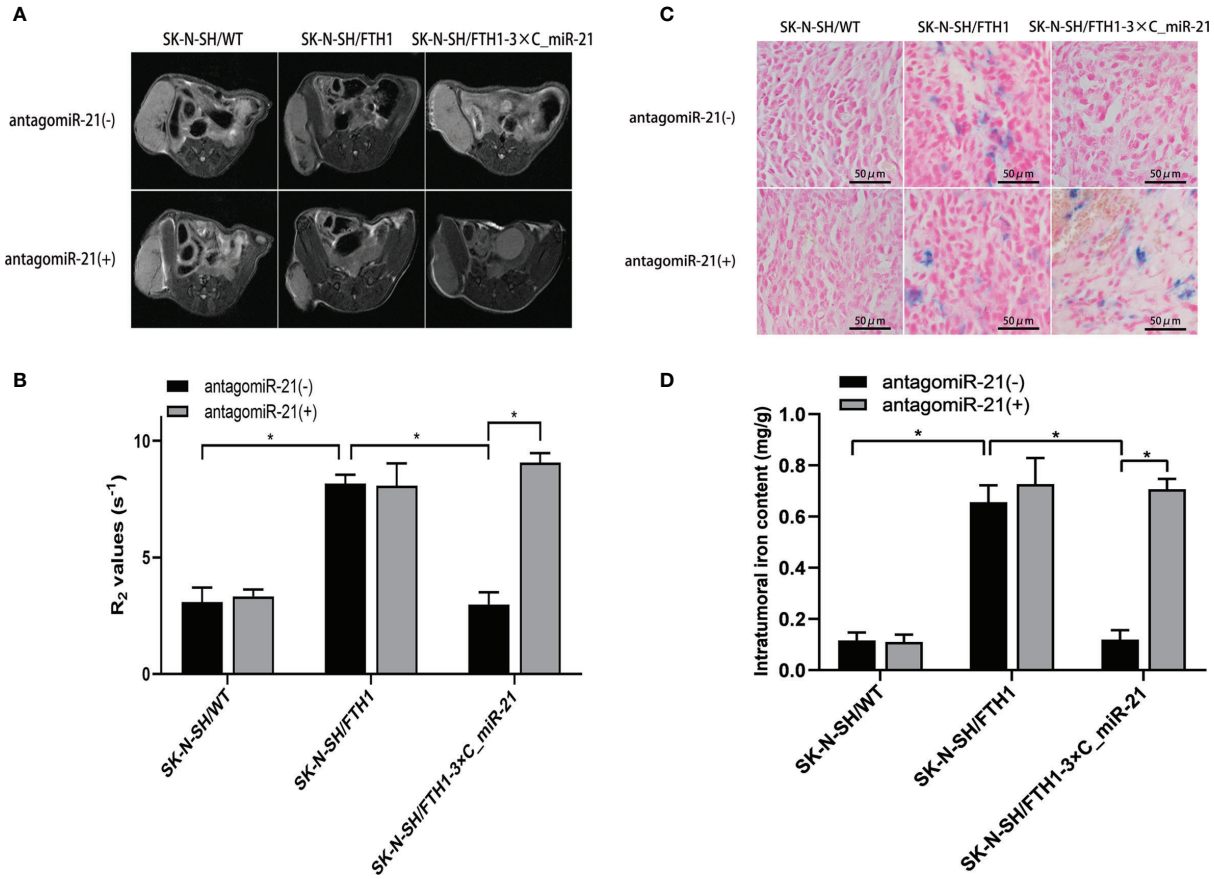


FIGURE 6 | MRI and iron detection in xenografts before and after antagomiR-21 administration. Before antagomiR-21 administration, the MRI signal in the SK-N-SH/FTH1 group was obviously decreased and the R2 value was significantly increased compared with those in the SK-N-SH/WT group, but these values in the SK-N-SH/FTH1-3×C_miR-21 group did not differ from those in the SK-N-SH/WT group. After treatment with the antagomiR-21, the SK-N-SH/FTH1-3×C_miR-21 group showed a significantly decreased MRI signal and increased R2 value compared with those prior to treatment. In the other two groups, the MRI signals and R2 values did not differ before and after antagomiR-21 treatment **(A, B)**. The Prussian blue staining **(C)** and intratumoral iron content quantification **(D)** results match the MRI results. Intratumoral iron accumulation was significantly increased after antagomiR-21 treatment in only the SK-N-SH/FTH1-3×C_miR-21 group. Three independent experiments were performed. * indicates $P < 0.05$.

CONCLUSION

In summary, we developed a novel miR-21-responsive gene reporter system for the MRI visualization of miR-21 dynamics in neuroblastoma. In the presence of active miR-21 in tumor cells, the reporter gene expression was in an "off" state due to the specific combination of miR-21 with 3 copies of completely complimentary sequences (3xC_miR-21) at the 3' UTR of the reporter gene. The exogenous delivery of antagomiR-21 into the cells by the GO-PEG-dendrimer switched "on" the reporter gene expression because miR-21 bound with antagomiR-21 in a competitive manner and was thus released from the FTH1 3' UTR. Reporter gene expression resulted in intracellular iron accumulation, thereby allowing MRI visualization. This study provides a noninvasive and feasible imaging method for the detection of miR-21, which is regarded as a potential diagnostic and/or therapeutic target in tumors.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The animal study was reviewed and approved by Animal Care and Use Committee of Chongqing Medical University.

AUTHOR CONTRIBUTIONS

GB and JC supervised the research. All authors were involved in the design of the study or the acquisition and analysis of the data.

GB, JS, JHou, HZ, YF, XZ, and BX performed related experiments. GB, JS, HZ, JQ, JHua, and YF analyzed the data and prepared the figures. GB, JS, and JC wrote the main manuscript. GB, JS, and HZ contributed equally to this work. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by a grant from the National Natural Science Foundation of China (no. 81771892).

ACKNOWLEDGMENTS

We thank the Key Laboratory of Pediatrics in Chongqing for providing parts of the experimental apparatus. We also thank American Journal Experts (AJE) for editing the English language in our manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.747305/full#supplementary-material

REFERENCES

1. Lu TX, Rothenberg ME. MicroRNA. *J Allergy Clin Immunol* (2018) 141:1202–7. doi: 10.1016/j.jaci.2017.08.034
2. Liu Z, Zhang XH, Callejas-Díaz B, Mullol J. MicroRNA in United Airway Diseases. *Int J Mol Sci* (2016) 17:716. doi: 10.3390/ijms17050716
3. Ha M, Kim VN. Regulation of microRNA Biogenesis. *Nat Rev Mol Cell Biol* (2014) 15:509–24. doi: 10.1038/nrm3838
4. Bandrés E, Cubedo E, Agirre X, Malumbres R, Zárate R, Ramirez N, et al. Identification by Real-Time PCR of 13 Mature microRNAs Differentially Expressed in Colorectal Cancer and Non-Tumoral Tissues. *Mol Cancer* (2006) 5:29. doi: 10.1186/1476-4598-5-29
5. Bommer GT, Gerin I, Feng Y, Kaczorowski AJ, Kuick R, Love RE, et al. P53-Mediated Activation of miRNA34 Candidate Tumor-Suppressor Genes. *Curr Biol* (2007) 17:1298–307. doi: 10.1016/j.cub.2007.06.068
6. Si ML, Zhu S, Wu H, Lu Z, Wu F, Mo YY. miR-21-Mediated Tumor Growth. *Oncogene* (2007) 26:2799–803. doi: 10.1038/sj.onc.1210083
7. Asangani IA, Rasheed SA, Nikolova DA, Leupold JH, Colburn NH, Post S, et al. MicroRNA-21 (miR-21) Post-Transcriptionally Downregulates Tumor Suppressor Pdcd4 and Stimulates Invasion, Intravasation and Metastasis in Colorectal Cancer. *Oncogene* (2008) 27:2128–36. doi: 10.1038/sj.onc.1210856
8. Du J, Yang S, An D, Hu F, Yuan W, Zhai C, et al. BMP-6 Inhibits microRNA-21 Expression in Breast Cancer Through Repressing Deltaef1 and AP-1. *Cell Res* (2009) 19:487–96. doi: 10.1038/cr.2009.34
9. Markou A, Tsaroucha EG, Kaklamanis L, Fotinou M, Georgoulias V, Lianidou ES. Prognostic Value of Mature microRNA-21 and microRNA-205 Overexpression in Non-Small Cell Lung Cancer by Quantitative Real-Time RT-PCR. *Clin Chem* (2008) 54:1696–704. doi: 10.1373/clinchem.2007.101741
10. Li T, Li D, Sha J, Sun P, Huang Y. MicroRNA-21 Directly Targets MARCKS and Promotes Apoptosis Resistance and Invasion in Prostate Cancer Cells. *Biochem Biophys Res Commun* (2009) 383:280–5. doi: 10.1016/j.bbrc.2009.03.077
11. Gao JS, Zhang Y, Tang X, Tucker LD, Tarwater PM, Quesenberry PJ, et al. The Evi1, microRNA-143, K-Ras Axis in Colon Cancer. *FEBS Lett* (2011) 585:693–9. doi: 10.1016/j.febslet.2011.01.033
12. Yao Q, Cao Z, Tu C, Zhao Y, Liu H, Zhang S. MicroRNA-146a Acts as a Metastasis Suppressor in Gastric Cancer by Targeting WASF2. *Cancer Lett* (2013) 335:219–24. doi: 10.1016/j.canlet.2013.02.031
13. Cheng W, Liu T, Wan X, Gao Y, Wang H. MicroRNA-199a Targets CD44 to Suppress the Tumorigenicity and Multidrug Resistance of Ovarian Cancer-Initiating Cells. *FEBS J* (2012) 279:2047–59. doi: 10.1111/j.1742-4658.2012.08589.x
14. Chen Y, Tsai YH, Fang Y, Tseng SH. Micro-RNA-21 Regulates the Sensitivity to Cisplatin in Human Neuroblastoma Cells. *J Pediatr Surg* (2012) 47:1797–805. doi: 10.1016/j.jpedsurg.2012.05.013
15. Cheng AM, Byrom MW, Shelton J, Ford LP. Antisense Inhibition of Human miRNAs and Indications for an Involvement of miRNA in Cell Growth and Apoptosis. *Nucleic Acids Res* (2005) 33:1290–7. doi: 10.1093/nar/gki200
16. Sun X, Chen Y, Yu H, Machan JT, Alladin A, Ramirez J, et al. Anti-miRNA Oligonucleotide Therapy for Chondrosarcoma. *Mol Cancer Ther* (2019) 18:2021–9. doi: 10.1158/1535-7163.MCT-18-1020
17. Pall GS, Codony-Servat C, Byrne J, Ritchie L, Hamilton A. Carbodiimide-Mediated Cross-Linking of RNA to Nylon Membranes Improves the Detection of siRNA, miRNA and piRNA by Northern Blot. *Nucleic Acids Res* (2007) 35:e60. doi: 10.1093/nar/gkm112
18. Zhang L, Wei P, Shen X, Zhang Y, Xu B, Zhou J, et al. MicroRNA Expression Profile in Penile Cancer Revealed by Next-Generation Small RNA Sequencing. *PLoS One* (2015) 10:e0131336. doi: 10.1371/journal.pone.0131336
19. Thomson JM, Parker J, Perou CM, Hammond SM. A Custom Microarray Platform for Analysis of microRNA Gene Expression. *Nat Methods* (2004) 1:47–53. doi: 10.1038/nmeth704
20. Zhou Z, Lu ZR. Molecular Imaging of the Tumor Microenvironment. *Adv Drug Deliv Rev* (2017) 113:24–48. doi: 10.1016/j.addr.2016.07.012
21. Mezzanotte L, Van 't Root M, Karatas H, Goun EA, Löwik C. *In Vivo* Molecular Bioluminescence Imaging: New Tools and Applications. *Trends Biotechnol* (2017) 35:640–52. doi: 10.1016/j.tibtech.2017.03.012
22. Lu CY, Ji JS, Zhu XL, Tang PF, Zhang Q, Zhang NN, et al. T2-Weighted Magnetic Resonance Imaging of Hepatic Tumor Guided by SPIO-Loaded Nanostructured Lipid Carriers and Ferritin Reporter Genes. *ACS Appl Mater Interfaces* (2017) 9:35548–61. doi: 10.1021/acsami.7b09879
23. Li D, Li Y. The Interaction Between Ferroptosis and Lipid Metabolism in Cancer. *Signal Transduct Target Ther* (2020) 5:108. doi: 10.1038/s41392-020-00216-5
24. He X, Cai J, Li H, Liu B, Qin Y, Zhong Y, et al. *In Vivo* Magnetic Resonance Imaging of Xenografted Tumors Using FTH1 Reporter Gene Expression Controlled by a Tet-on Switch. *Oncotarget* (2016) 7:78591–604. doi: 10.18632/oncotarget.12519
25. Iordanova B, Ahrens ET. *In Vivo* Magnetic Resonance Imaging of Ferritin-Based Reporter Visualizes Native Neuroblast Migration. *Neuroimage* (2012) 59:1004–12. doi: 10.1016/j.neuroimage.2011.08.068
26. Ziv K, Meir G, Harmelin A, Shimoni E, Klein E, Neeman M. Ferritin as a Reporter Gene for MRI: Chronic Liver Over Expression of H-Ferritin During Dietary Iron Supplementation and Aging. *NMR BioMed* (2010) 23:523–31. doi: 10.1002/nbm.1491
27. Velde GV, Rangarajan JR, Toelen J, Dresselaers T, Ibrahimi A, Krylychkina O, et al. Evaluation of the Specificity and Sensitivity of Ferritin as an MRI Reporter Gene in the Mouse Brain Using Lentiviral and Adeno-Associated Viral Vectors. *Gene Ther* (2011) 18:594–605. doi: 10.1038/gt.2011.2
28. Cohen B, Ziv K, Plaks V, Harmelin A, Neeman M. Ferritin Nanoparticles as Magnetic Resonance Reporter Gene. *Wiley Interdiscip Rev Nanomed Nanobiotechnol* (2009) 1:181–8. doi: 10.1002/wnan.11
29. Mu T, Qin Y, Liu B, He X, Liao Y, Sun J, et al. *In Vitro* Neural Differentiation of Bone Marrow Mesenchymal Stem Cells Carrying the FTH1 Reporter Gene and Detection With MRI. *BioMed Res Int* (2018) 2018:1978602. doi: 10.1155/2018/1978602
30. Zhang M, Liu X, Huang J, Wang L, Shen H, Luo Y, et al. Ultrasmall Graphene Oxide Based T(1) MRI Contrast Agent for *In Vitro* and *In Vivo* Labeling of

Human Mesenchymal Stem Cells. *Nanomedicine* (2018) 14:2475–83. doi: 10.1016/j.nano.2017.03.019
31. Gu Y, Guo Y, Wang C, Xu J, Wu J, Kirk TB, et al. A Polyamidoamne Dendrimer Functionalized Graphene Oxide for DOX and MMP-9 shRNA Plasmid Co-Delivery. *Mater Sci Eng C Mater Biol Appl* (2017) 70:572–85. doi: 10.1016/j.msec.2016.09.035
32. Wang F, Zhang B, Zhou L, Shi Y, Li Z, Xia Y, et al. Imaging Dendrimer-Grafted Graphene Oxide Mediated anti-miR-21 Delivery With an Activatable Luciferase Reporter. *ACS Appl Mater Interfaces* (2016) 8:9014–21. doi: 10.1021/acsami.6b02662
33. Ribeiro BFM, Souza MM, Fernandes DS, Do Carmo DR, Machado-Santelli GM. Graphene Oxide-Based Nanomaterial Interaction With Human Breast Cancer Cells. *J BioMed Mater Res A* (2020) 108:863–70. doi: 10.1002/jbm.a.36864
34. Qu Y, Sun F, He F, Yu C, Lv J, Zhang Q, et al. Glycyrrhetinic Acid-Modified Graphene Oxide Mediated siRNA Delivery for Enhanced Liver-Cancer Targeting Therapy. *Eur J Pharm Sci* (2019) 139:105036. doi: 10.1016/j.ejps.2019.105036
35. Alibolandi M, Mohammadi M, Taghdisi SM, Ramezani M, Abnous K. Fabrication of Aptamer Decorated Dextran Coated Nano-Graphene Oxide for Targeted Drug Delivery. *Carbohydr Polym* (2017) 155:218–29. doi: 10.1016/j.carbpol.2016.08.046
36. Sun J, Huang J, Bao G, Zheng H, Wang C, Wei J, et al. MRI Detection of the Malignant Transformation of Stem Cells Through Reporter Gene Expression Driven by a Tumor-Specific Promoter. *Stem Cell Res Ther* (2021) 12:284. doi: 10.1186/s13287-021-02359-w
37. Kim S, Hwang DW, Lee DS. A Study of microRNAs *in Silico* and *In Vivo*: Bioimaging of microRNA Biogenesis and Regulation. *FEBS J* (2009) 276:2165–74. doi: 10.1111/j.1742-4658.2009.06935.x
38. Ko HY, Hwang DW, Lee DS, Kim S. A Reporter Gene Imaging System for Monitoring microRNA Biogenesis. *Nat Protoc* (2009) 4:1663–9. doi: 10.1038/nprot.2009.119
39. Kim HJ, Chung JK, Hwang DW, Lee DS, Kim S. *In Vivo* Imaging of miR-221 Biogenesis in Papillary Thyroid Carcinoma. *Mol Imaging Biol* (2009) 11:71–8. doi: 10.1007/s11307-008-0188-6
40. Song H, Oh B, Choi M, Oh J, Lee M. Delivery of anti-microRNA-21 Antisense-Oligodeoxynucleotide Using Amphiphilic Peptides for Glioblastoma Gene Therapy. *J Drug Target* (2015) 23:360–70. doi: 10.3109/1061186X.2014.1000336
41. Costa PM, Pedroso de Lima MC. MicroRNAs as Molecular Targets for Cancer Therapy: On the Modulation of microRNA Expression. *Pharmaceuticals (Basel)* (2013) 6:1195–220. doi: 10.3390/ph6101195
42. Liu XQ, Song WJ, Sun TM, Zhang PZ, Wang J. Targeted Delivery of Antisense Inhibitor of miRNA for Antiangiogenesis Therapy Using cRGD-Functionalized Nanoparticles. *Mol Pharm* (2011) 8:250–9. doi: 10.1021/mp100315q

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Bao, Sun, Zheng, Hou, Huang, Wei, Fu, Qiu, Zou, Xiang and Cai. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 17 November 2021
doi: 10.3389/fonc.2021.778728

Multimodal Imaging Technology Effectively Monitors HER2 Expression in Tumors Using Trastuzumab-Coupled Organic Nanoparticles in Patient-Derived Xenograft Mice Models

Li Wen[1,2†], Lei Xia[2†], Xiaoyi Guo[2], Hai-Feng Huang[3], Feng Wang[2], Xian-teng Yang[1,3], Zhi Yang[1,2*] and Hua Zhu[1,2*]*

[1] Medical College, Guizhou University, Guiyang, China, [2] Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), NMPA Key Laboratory for Research and Evaluation of Radiopharmaceuticals (National Medical Products Administration), Department of Nuclear Medicine, Peking University Cancer Hospital & Institute, Beijing, China, [3] Department of Orthopaedics, Guizhou Provincial People's Hospital, Guiyang, China

Trastuzumab is a monoclonal antibody targeting human epidermal growth factor 2 (HER2), which has been successfully used in the treatment of patients with breast cancer and gastric cancer; however, problems concerning its cardiotoxicity, drug resistance, and unpredictable efficacy still remain. Herein, we constructed novel organic dopamine–melanin nanoparticles (dMNs) as a carrier and then surface-loaded them with trastuzumab to construct a multifunctional nanoprobe named Her-PEG-dMNPs. We used micro-PET/CT and PET/MRI multimodality imaging to evaluate the retention effect of the nanoprobe in HER2 expression in gastric cancer patient-derived xenograft (PDX) mice models after labeling of the radionuclides ^{64}Cu or ^{124}I and MRI contrast agent Mn^{2+}. The nanoprobes can specifically target the HER2-expressing SKOV-3 cells *in vitro* (3.61 ± 0.74 vs. 1.24 ± 0.43 for 2 h, $P = 0.002$). *In vivo*, micro-PET/CT and PET/MRI showed that the ^{124}I-labeled nanoprobe had greater contrast and retention effect in PDX models than unloaded dMNPs as carrier (1.63 ± 0.07 vs. 0.90 ± 0.04 at 24 h, $P = 0.002$), a similarity found in ^{64}Cu-labeled Her-PEG-dMNPs. Because ^{124}I has a longer half-life and matches the pharmacokinetics of the nanoparticles, we focused on the further evaluation of ^{124}I-Her-PEG-dMNPs. Furthermore, immunohistochemistry staining confirmed the overexpression of HER2 in the animal model. This study developed and validated novel HER2-specific multimodality imaging nanoprobes for quantifying HER2 expression in mice. Through the strong retention effect of the tumor site, it can be used for the promotion of monoclonal antibody treatment effect and process monitoring.

Keywords: trastuzumab, dMNPs, PET/MRI imaging, HER2, micro-PET/CT imaging

OPEN ACCESS

Edited by:
Kyung Hyun Sung,
UCLA Health System, United States

Reviewed by:
Zheng Han,
Johns Hopkins University, United States
Min Yang,
Jiangsu Institute of Nuclear Medicine, China

***Correspondence:**
Xianteng Yang
yangxian.teng@163.com
Zhi Yang
pekyz@163.com
Hua Zhu
zhuhuananjing@163.com

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *17 September 2021*
Accepted: *20 October 2021*
Published: *17 November 2021*

Citation:
Wen L, Xia L, Guo X, Huang H-F, Wang F, Yang X-t, Yang Z and Zhu H (2021) Multimodal Imaging Technology Effectively Monitors HER2 Expression in Tumors Using Trastuzumab-Coupled Organic Nanoparticles in Patient-Derived Xenograft Mice Models. Front. Oncol. 11:778728. doi: 10.3389/fonc.2021.778728

INTRODUCTION

Human epidermal growth factor was discovered in 1962, while structurally related receptors were discovered in 1978 (1). As for the HER family, it is mainly composed of four receptors, HER1, HER2, HER3, and HER4, which can interact with each other in regulating different biological behaviors such as cell proliferation, differentiation, and survival. Human epidermal growth factor 2 (HER2) is a transmembrane glycoprotein, also known as CERBB-2 and ERbb2, which promotes malignant biological behaviors such as cell proliferation and migration through the formation of homologous or heterologous dimers with family members (2). Amplification of the HER gene and overexpression of its products were found for the first time in breast cancer (3, 4). Positive HER2 expression was found in 15%–20% of breast cancers, which had a high degree of malignancy and poor prognosis (4). HER2 is an ideal molecular target for cancer, highly expressed in a variety of cancer cells, such as colorectal cancer, ovarian cancer, prostate cancer, bladder cancer, and lung cancer, especially gastric and esophageal cancer (3–6).

According to "Cancer Statistics, 2020", there were approximately 19.3 million new cancer cases and 10.0 million cancer deaths in the world (7). Gastric cancer has been one of the leading causes of common cancer and cancer-related deaths. Trastuzumab (Herceptin), a recombinant humanized IgG monoclonal antibody, binds to the extracellular domain IV of HER2 for antitumor activity and was approved by the US Food and Drug Administration as first-line treatment of advanced HER2-positive gastric cancer (8). HER2 was overexpressed in 7%–34% of gastric cancer patients and has been demonstrated to increase overall survival in locally advanced or metastatic gastric cancer (9, 10). At present, many researchers have reported ^{68}Ga, ^{18}F, ^{89}Zr, ^{64}Cu, and other positron nuclide-labeled PET molecular probes targeting HER2, for example, monoclonal antibody fragments, affibody, and nanobody (11–16). Preliminary clinical trials have been conducted studying ^{64}Cu/^{89}Zr-trastuzumab, ^{89}Zr-pertuzumab, ^{68}Ga-nanobody, and so on (14, 17–19). However, some probes still have problems such as drug resistance and toxic side effects (17, 20). In addition, intratumoral heterogeneity and small biopsy tumor samples may not represent HER2 expression across the tumor or metastatic tumor lesions (6, 21). Therefore, a more accurate, non-invasive HER2 status assessment is needed to help guide physicians to tailor treatment for each patient, monitor HER2-targeted therapy responses, and identify those patients who develop resistance.

Multimodal tumor molecular imaging is a hot spot in medical imaging research. By constructing specific molecular probes for tumor cell surface receptors and combining the functions of different imaging contrast agents to provide multiple imaging effects, collaborative imaging that cannot be matched by single imaging can be obtained, which has a good clinical application prospect in the early diagnosis of diseases. In view of the problems existing in the application of trastuzumab, nanomedicine based on molecular imaging can provide a direction for solving such concerns. The hydrodynamic properties of antibodies will be changed by combining monoclonal antibodies with high-quality nanoparticles, which can improve the tumor uptake and retention rate and reduce heart toxicity (22, 23). On the other hand, the nanoprobe can load highly sensitive imaging molecules to achieve multimodality imaging function, which can be used to monitor the treatment process in real time and adjust the individualized treatment plan. This is also a major trend of precision medicine in the future (23). Nanoparticle-based probes enable non-invasive imaging of proteolytic activity for cancer diagnosis. However, there are also some toxic problems in the application of nanoparticles. The exogenous inorganic nanoparticles not only provide a strong contrast, but also have high sensitization and high toxicity and are difficult to remove *in vivo*, which cannot be ignored (24–26). The emergence of organic nanoparticles based on endogenous biological materials increases the application scope and safety of nanotechnology and has been widely studied in the field of molecular imaging (26, 27).

Melanin is a ubiquitous, amorphous, and irregular biopolymer, which is widely distributed in almost all living organisms and has many unique functions (28, 29). Dopamine–melanin nanoparticles (dMNPs) consist of naturally producing dopamine with good biocompatibility and biodegradability (30). It retains many characteristics of melanin and can directly complex with various metal ions without chelating agents, which significantly simplifies the preparation process and reduces the heterogeneity of the nanoparticles. dMNPs have been demonstrated to be used for tumor multimodality imaging, such as photoacoustic imaging (PAI), positron emission tomography (PET), and magnetic resonance imaging (MRI) (27, 31). This study reports the characterization, quality control, and micro-PET/CT and PET/MRI multimodality imaging evaluation of a novel nanoprobe—(^{124}I/^{64}Cu, Mn)-Her-PEG-dMNPs—and applied this probe in gastric cancer patient-derived xenograft (PDX) models. The results indicate that the novel nanoprobe Mn-Her-PEG-dMNP has a good retention effect on tumors and has great potential in the treatment of gastric cancer.

MATERIALS AND METHODS

Materials

Dopamine hydrochloride, $MnCl_2$, and ammonia aqueous solution were obtained from Sigma-Aldrich (St. Louis, MO, USA). Amine-mPEG5000 (mPEG5000–NH_2, 5,000 Da) was acquired from Aladdin (Shanghai, China). Herceptin was obtained from Roche Pharmaceuticals (Shanghai, China) Ltd. 2-Mercaptoethanol was purchased from Sinopharm Chemical Reagent Co. Ltd. For the cell experiments, Dulbecco's modified Eagle's medium (DMEM) and fetal bovine serum (FBS) were purchased from Gibco (USA); penicillin–streptomycin mixture double and anti-0.25% trypsin (including EDTA) were purchased from Invitrogen (USA). In the experiment, high nuclear pure simple ^{124}I was prepared by ^{124}Te solid target of HM-20 cyclotron at the Nuclear Medicine Department of Peking University Cancer Hospital. The nuclide was purified by ^{124}I special purification device and collected in 0.02 M NaOH

solution (radioactive concentration 148 MBq/ml). The positron nuclide ^{64}Cu was prepared by bombarding a ^{64}Ni solid target with a medical cyclotron at the Nuclear Medicine Department of Peking University Cancer Hospital with 3.7 MBq/μl in 0.01 M HCl solution (32).

Synthesis and Modification of dMNPs

Dopamine–melanin nanoparticles were prepared according to the methods reported in previous literature (30). Briefly, ammonia aqueous solution (0.3 ml, NH_4OH, 28%–30%) was mixed with ethanol (4 ml) and deionized water (9 ml) under mild stirring at 40°C for 30 min. Fifty milligrams of dopamine hydrochloride was dissolved in water (1 ml), then injected into the above mixture solution, and stirred for 48 h. The color of this solution changes from light yellow to dark brown. The bright black solution was then centrifuged with a centrifugal filter (Amicon centrifugal filter device, MWCO = 50 kDa) and washed several times to remove the unpolymerized dopamine–melanin molecules. The reaction solution was taken and diluted to 5 ml ultrapure water by ultrasonication for 5 min, and the reaction system pH = 9 was adjusted by NH_4OH (25%–28%). Then, the solution was slowly dripped into the reaction bottle containing 25 mg diamino polyethylene glycol (NH_2–PEG_{5000}–NH_2) and stirred at room temperature for 12 h. The unreacted NH_2–PEG_{5000}–NH_2 was removed by a 50-kDa ultrafiltration centrifuge to obtain NH_2–PEG-dMNPs.

Preparation of the Her-PEG-dMNPs and Mn-Her-PEG-dMNPs

The amino-sulfhydryl cross-linking agent Sulfo-SMCC was used as a linker in the combination of Herceptin-SH and NH_2–PEG-dMNPs. Five milligrams of NH_2–PEG-dMNPs (0.5×10^{-4} mM) was fully dispersed in 1 ml phosphate buffered solution (PBS), and excess amino-sulfhydryl cross-linking agent Sulfo-SMCC (5 μM) was stirred at room temperature for 2 h. Then, the unreacted Sulfo-SMCC was separated by the PD-10 column. The reaction solution was concentrated to 0.6 ml by ultrafiltration and slowly added to Herceptin-SH (0.33×10^{-4} mM). After reacting for 2 h at room temperature with stirring, the unreacted sulfhydryl–Herceptin was removed using a PD-10 column (GE Healthcare, USA) with an eluent of PBS (pH = 7.4) and lyophilized to obtain a purified Her-PEG-dMNP. Two milligrams of Her-PEG-dMNPs and $MnCl_2$ (10×10^{-3} mM) were fully dispersed in 1 ml deionized water and stirred at room temperature for 2 h. The Mn-Her-PEG-dMNPs were obtained by centrifugation with an ultrafiltration centrifuge tube (MWCO = 10 kDa).

Characterization

The morphology of Her-PEG-dMNPs was characterized by using a JEM-2100F transmission electron microscope with an accelerating voltage of 200 kV, and the sample was made by using a film of copper mesh. Dynamic light scattering (DLS) was used to detect the hydrodynamic diameter of the dMNPs, PEG-dMNPs, and Her-PEG-dMNPs with a laser light-scattering spectrometer (SLS-5022F, Germany) ^{1}H-NMR spectra of dMNPs and PEG-dMNPs were detected by a 600-MHz NMR (Bruker) spectrometer with D_2O as the solvent.

Cytotoxicity Assay

The viability and proliferation of HeLa and HepG2 cells were evaluated by a standard 3-(4,5-dimethylthiazol2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay. Typically, HeLa and HepG2 cells were incubated in the culture medium at 37°C in an atmosphere of 5% CO_2 and 95% air for 24 h. Then, the medium was removed. The cells were incubated in cultured medium containing Her-PEG-dMNPs and Mn-Her-PEG-dMNPs with different concentrations of 0.125, 0.25, 0.5, 1, and 2 mg/ml for 24 h and washed with the medium twice. The MTT solution (20 μl, 0.5 mg/ml) was added to each well, and the cells were incubated for 4 h without light at 37°C. After removal of the supernatant, the residues were lysed with 150 μl of dimethyl sulfoxide (DMSO). Finally, the light absorption (OD) values were measured at 490 nm with a microplate reader. The untreated cells were processed analogously, and their OD values were used as the reference to calculate 100% cellular viability.

Radiosynthesis and Quality Control of the Radiotracer

^{64}Cu was diluted with hydrochloric acid (0.01 M, 300 μl), then a small amount of sodium acetate (0.1 mol/L) was added to adjust the pH to 5.5, and then Her-PEG-dMNP (2 mg/ml) was added and incubated at 40°C for 1 h. When the reaction was over, saturated EDTA solution (0.05 M) was added and mixed with the reaction liquid. ^{64}Cu-Her-PEG-dMNP was purified using PD-10 column in PBS (pH = 7.4). ^{124}I can label Her-PEG-dMNPs *in situ* using N-bromosuccinimide as the oxidant. The ^{124}I buffer was diluted with PB solution (0.1 M, 500 μl) before being added to the Mn-Her-PEG-MNPs. Then, 20 μl N-bromosuccinimide (NBS, 5.6 mM) was added to the above solution for 1 min. The radiolabeled reaction was then ended and the ^{124}I-Her-PEG-dMNP was achieved *via* the PD-10 column in PBS buffer, and radio-TLC was used to detect the radiolabeling rate of the ^{124}I-Her-PEG-dMNPs with acetone as the developer. The stability of the ^{64}Cu-HER-PEG-dMNPs and ^{124}I-Her-PEG-dMNPs *in vitro* was measured by incubations in PBS (pH = 7.4) and 5% HSA solution. Radio-TLC was used to analyze the radiochemical purity of the ^{124}I/^{64}Cu-Her-PEG-dMNPs at different time points.

In-Vitro Cell Uptake

One day before the study, human ovarian carcinoma SKOV-3 cells were seeded in 24-well plates at an average of 1×10^5 cells per well and incubated overnight in 1 ml of DMEM medium. Prior to the study, the cells were washed twice with 1 ml of cold PBS, and after the culture solution was aspirated in each well, the cells were re-incubated with 500 μl of serum-free medium for 2 h at 37°C. Then, four wells in every dish were divided into one group. A total of 37 kBq of ^{124}I/^{64}Cu-Her-PEG-dMNPs in 500 μl serum-free medium was added to each well, and the cells were cultured at 37°C with 5% CO_2. To assess the specific targeting ability of ^{124}I/^{64}Cu-Her-PEG-dMNPs for targeting HER2, three groups of SKOV-3 cells (n = 4) were co-incubated with ^{124}I/^{64}Cu-HER-PEG-dMNPs and 20 μg of Her-PEG-dMNPs. After co-incubation with the radiotracer for 10, 30, 60, and 120 min, the cell culture media of each group was sequentially

aspirated, and the cells were rinsed three times with 1 ml of cold PBS and lysed with 500 μl of 0.5 M NaOH. Cell lysates in each well were then collected and counted by a γ-counter.

Subcutaneous Tumor Models

All animal procedures were performed in accordance with the Guidelines for the Care and Use of Laboratory Animals of Peking University (Beijing, China), and experiments were approved by the Animal Ethics Committee of the National Regulation of China for the Care and Use of Laboratory Animals and in compliance with the guidelines established by Peking University Cancer Hospital Animal Care and Use Committee. The PDX models were constructed as in reference (11). Fresh tumor tissue mass derived from patients was inoculated into the right groin area of NOD-SCID mice, and the tumor tissue was removed when the tumor grew to 1 cm in diameter. The tissue blocks were cut into 0.05 cm * 0.05 cm * 0.05 cm for subculture.

Pharmacokinetics Studies

Healthy KM mice (18–20 g) were used to conduct the pharmacokinetics study. Mice were sealed off 3 days in advance with 0.5% potassium iodide (KI), and ^{124}I-Her-PEG-dMNP (3.7 MBq) was intravenously injected into each mouse (n = 5) at set intervals (1, 3, 5, 10, 15, 30, 60, and 120 min and 4, 8, 24, 48, and 72 h). Then, ^{64}Cu-Her-PEG-dMNP also was intravenously injected into each mouse (n = 5) at set intervals (1, 3, 5, 10, 15, 30, 60, and 120 min and 4, 8, 24, and 48 h) and into the blood through the orbital vein. The blood samples were weighed before counting using a γ-counter. The results were calculated as the percentage of injected dose per gram (%ID/g, mean ± SD), and the parameters of the pharmacokinetics were analyzed using the GraphPad Prism software.

Micro-PET/CT Imaging Studies in PDX NOD-SCID Mice

Micro-PET/CT imaging studies were conducted by using PDX model NOD-SCID mice with a micro-PET/CT system (Super Nova PET/CT, PINGSENG, Shanghai, China). Tumor-bearing NOD-SCID mice were injected *in situ* with ^{124}I-Her-PEG-dMNPs and ^{124}I-PEG-dMNPs (1.11 MBq). Anesthesia for the animals was performed using a PINGSENG isoflurane gas anesthesia system. Axial and coronal PET/CT images of the tumor were collected at 2, 24, 48, and 72 h. The mice were anesthetized with isoflurane before imaging, and anesthesia was maintained during imaging (volume coefficient 1%). Micro-PET images were acquired for 15 min and reconstructed with attenuation correction based on the CT data (CT–AC reconstruction). Regions of interest (ROIs) were drawn on the CT images and further mapped on PET. The experiments were repeated three times at each time point.

PET/MRI of the Mn-Her-PEG-dMNPs

For *in-vivo* MRI, PDX model NOD-SCID mice (n = 3) were first anesthetized by intraperitoneal injection of 4% chloral hydrate solution (50 mg/kg), when the tumors reached approximately 0.8 cm, followed by *in-situ* injection of Mn-Her-PEG-dMNPs (0.125 mM, 30 μl/mouse). Axial images of tumor locations were collected at 2, 24, and 72 h, using a clinical 1.5-T magnetic resonance scanner (Siemens) with animal microcoils. The MRI sequence was as follows: TR, 531 ms; TE, 9.1 ms; flip angle, 30°; FOV, 160 × 100 mm^2; matrix, 256 × 256; and slice thickness, 3 mm. The analysis of the PET/MRI signal in the tumor images was conducted with the 3D Slicer software.

Immunohistochemistry of PDX Tumor Tissues

Mice were sacrificed after the last imaging point, and then the tumor tissue was immediately removed and fixed in formalin. For immunohistochemical staining, sections of PDX tumor tissue were placed in a beaker containing citrate antigenic repair solution. Intermittent heating was performed three times with microwaves. Then, the tissue was treated with 10% goat serum at 37°C for 30 min. After being rinsed three times with PBS, the samples were incubated overnight with anti-HER2 antibody (1:300) at 4°C, and then sections were incubated with biotinylated goat anti-rabbit (HRP) antibody (1:1,000) at 37°C for 30 min. The sections were then stained with DAB substrate kit and hematoxylin and differentiated with 0.25% hydrochloric alcohol. Sections were dehydrated and sealed with neutral gum and photographed for analysis.

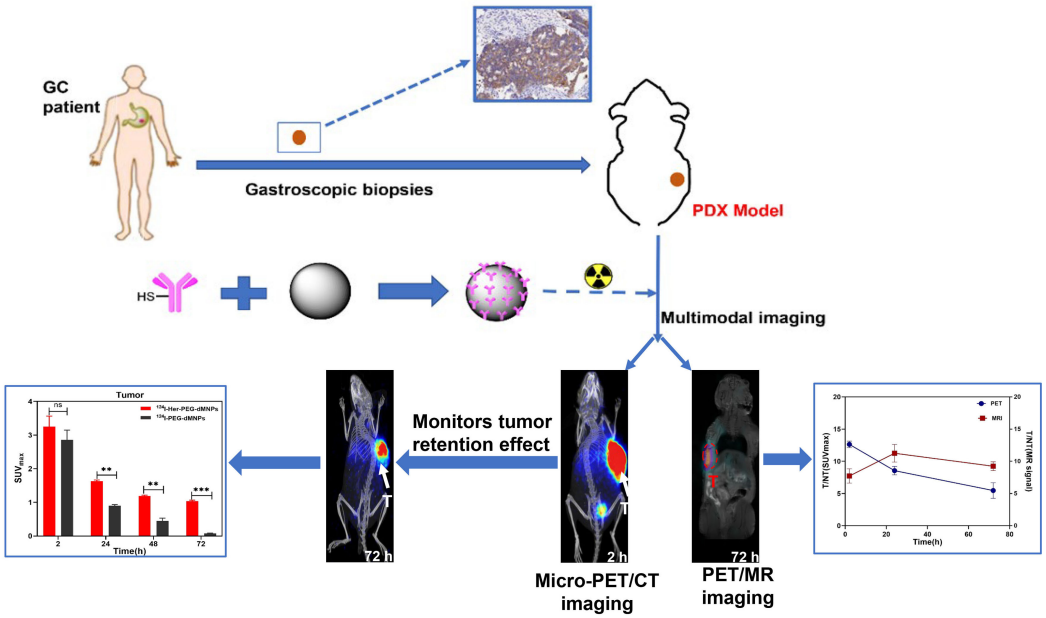
RESULTS

Synthesis and Characterization of the HER-PEG-dMNPs

Scheme 1 shows the reaction flowchart of the Her-PEG-dMNP imaging probe. Typical scanning electron microscopy (SEM) and transmission electron microscopy (TEM) imaging showed that the synthesized dMNPs and Her-PEG-DMNPs had clear structures and uniform and regular shapes and sizes (**Figures 1A, B**). As shown in **Figures 1C, D** and **Figure S1**, the hydrodynamic particle size of dMNPs is about 62.25 ± 6.12 nm, and PEG-dMNP was 101.28 ± 13.63 nm, which was slightly higher than before. To construct a targeted probe for tumors with high HER2 expression, trastuzumab with an active sulfhydryl group was bound to the surface of the nanoparticles by electrophilic substitution, and the size of Her-PEG-dMNPs increased to 148.16 ± 20.04 nm. The ^{1}H-NMR spectra of PEG-dMNPs in D_2O showed an emerging peak at 3.5 ppm, which represented the –OCH_2CH_2O– group of the PEG molecule (**Figures 1E, F**). In addition, both Her-PEG-dMNPs and Mn-Her-PEG-dMNPs showed high biocompatibility and low cytotoxicity in HepG2 and HeLa cells after being cultured for 24 h (compared with the non-toxic control). The results showed that the cell survival rate was above 95% when the concentration of both probes reached 2 mg/ml (**Figures 2A, B**).

Radiosynthesis and Quality Control of the Nanoprobe

The radiochemical yield of ^{64}Cu-Her-PEG-dMNPs was 85.1%–89.6% and the specific activity was 4.77–8.25 Bq/mg, while ^{124}I-Her-PEG-dMNPs had a radiochemical yield of 90%–93% and a specific activity of 7.98–13.30 MBq/mg. The final purity of both radiotracers was >98% (**Figure 2C** and **Figure S2**). The *in-vitro* stability of ^{64}Cu/^{124}I-Her-PEG-dMNPs in PBS and 5% HSA was evaluated by



SCHEME 1 | Synthesis and characterization of $^{124}I/^{64}Cu$-Mn-Her-PEG-MNPs. ^{124}I, ^{64}Cu, and Mn^{2+}were used to construct a tumor special nanoprobe for MRI/PET imaging of the gastric cancer PDX model. All the numerical data are presented as mean ± SD. *P < 0.05; **P < 0.01; ***P < 0.001 by two-way ANOVA. The symbols **, *** indicates statistically significant, very statistically significant, and the expanded form of ns indicates no statistically significant.

radio-TLC. The PBS (pH = 7.4) and 5% HSA had more than 89% radioactive purity at 96 h (**Figure 2D** and **Figure S2**).

In-Vitro Cell Experiments

The cell uptake of ^{124}I-Her-PEG-dMNPs in HER2-expressed human ovarian carcinoma SKOV-3 cell showed a higher uptake than that of ^{64}Cu-Her-PEG-dMNPs in cells at 10 min, 30 min, and 1 h incubation time periods (**Figure 2E** and **Figure S2**). For SKOV-3 cells, the uptake of ^{124}I-Her-PEG-dMNPs reached 0.65% ± 0.31%, 1.56% ± 0.81%, and 3.61% ± 0.74% of the collected count to the total added count (% AD/10^5 cells) at 30, 60, and 120 min, while the uptake of ^{64}Cu-Her-PEG-dMNPs was 1.55% ± 1.29%, 2.43% ± 0.51%, and 4.03% ± 0.55% AD/10^5 cells, respectively. After blocking with the precursor, the uptake of ^{124}I-Her-PEG-dMNPs at 60 and 120 min was significantly reduced (1.56 ± 0.81 *vs.* 0.72 ± 0.06 at 1 h, $P = 0.008$; 3.61 ± 0.74 *vs.* 1.24 ± 0.43 for 2 h, $P = 0.002$) (**Figure 2E**). The uptake of ^{64}Cu-Her-PEG-dMNPs was significantly reduced at 1 and 2 h (2.43 ± 0.51 *vs.* 1.65 ± 0.41 for 1 h, $P = 0.002$; 4.03 ± 0.55 *vs.* 1.00 ± 0.12 for 2 h, $P = 0.001$) (**Figure S2**). These results suggest that $^{124}I/^{64}$Cu-Her-PEG-dMNPs have specific targeting ability against HER2-expressed tumor cells.

Pharmacokinetics of $^{124}I/^{64}$Cu-Her-PEG-dMNPs in KM Mice

The profile of radioactivity in the blood after tail injection of $^{124}I/^{64}$Cu-Her-PEG-dMNPs in KM healthy mice was delineated using GraphPad Prism software (**Figure 2F** and **Figure S2**). Circulating pharmacokinetics of $^{124}I/^{64}$Cu-Her-PEG-dMNPs were evaluated using a two-compartment model. The data represent mean values, representing the percentage (ID%/g) of normal (n = 5) KM mice per gram of blood injection dose. The model parameters were obtained by fitting the relationship between the percentage of injected dose per gram of blood (ID%/g) and post-injection time into the following equation: Ct (%ID/g) = SpanFast * exp(αt) + SpanSlow * exp(βt), where SpanFast and SpanSlow are the y-intercepts of the model, and α and β are the fast and slow rate constants, expressed as the reciprocal of the time axis. In this model, the equations of $^{124}I/^{64}$Cu-Her-PEG-dMNPs in KM mice were as follows: Ct (%ID/g) = 4.507 + 406.566 * exp(−17.81t) + 319.96 * exp (−0.384t); Ct (%ID/g) = 2.588 + 207.73 * exp(−0.466t) + 52.48 * exp (−0.053t). The half-life of $^{124}I/^{64}$Cu-Her-PEG-dMNP distribution stage was 1.485 and 0.03892 h, respectively, while the elimination stage was 12.87 and 1.803 h, respectively.

The Construction of the PDX Model and Immunohistochemistry Staining

All animal experiments were performed in accordance with the guidelines of the Animal Care and Use Committee of Peking University (Ethics Approval License: 2015KT08). HER2-positive patient-derived xenograft (PDX) models were built for multimodality imaging. When the tumor volume reached 0.5–1.0 cm in diameter, mice were randomly assigned. The represented immunohistochemistry results showed that the PDX tumor tissues exhibited high HER2 expression (**Figure 4C**).

Micro-PET/CT Imaging Studies in the PDX Nude Mice

We first used micro-PET/CT imaging to observe the accumulation of two radioactive tracers in HER2 highly expressing gastric cancer PDX model tumor site with $^{124}I/^{64}$Cu-Her-PEG-dMNPs. In recent years,

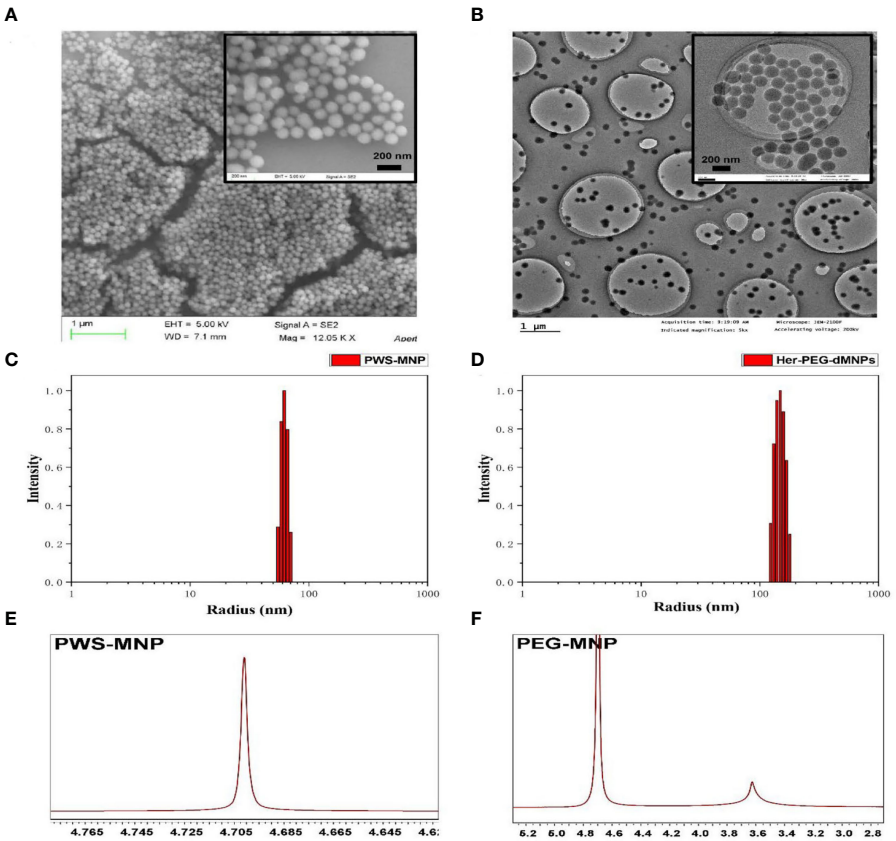


FIGURE 1 | Characterization of the nanoprobe. **(A)** SEM of dMNPs with a scale bar = 1 μm, and Her-PEG-dMNPs (reduction) with a scale bar = 200 nm. **(B)** TEM of dMNPs and Her-PEG-dMNPs (reduction) with scale bars = 1 μm and 200 nm. Hydrodynamic size distribution graphs of PWS-dMNP **(C)** and Her-PEG-dMNPs **(D)**. **(E)** ^{1}H-NMR spectra of PWS-MNP and **(F)** PEG-dMNPs in D_2O.

studies have shown that nanoprobes will be rapidly removed after *in-vivo* injection by human phagocytes, and their targeting function is also affected by the inevitable formation of protein crown. *In-situ* tumor injection, to some extent, can avoid such effects. The images of delayed acquisition are shown in **Figure 3A** and **Figure S3**. Both ^{124}I and ^{64}Cu showed excellent retention effects in the PDX model (n = 3) by *in-situ* injection compared with the control of the non-targeted molecular probe. ROI-based micro-PET/CT quantification analysis of those probes is shown in **Figures 3B, C**. The maximum standardized uptake values (SUVmaxs) of the liver were 0.24 ± 0.03, 0.42 ± 0.01, 0.38 ± 0.06, and 0.11 ± 0.01 at 2, 24, 48, and 72 h of 124I-Her-PEG-dMNPs, compared with the SUVmaxs of 0.87 ± 0.17, 1.47 ± 0.05, 1.72 ± 0.15, and 1.05 ± 0.05 at 2, 24, 48, and 72 h of ^{124}I-PEGdMNPs (**Figure 3D**). The SUVmaxs of tumor sites were 3.01 ± 0.59, 1.63 ± 0.07, 1.19 ± 0.06, and 1.04 ± 0.06 of 124I-Her-PEG-dMNPs, compared with the SUVs of 2.86 ± 0.52, 0.90 ± 0.04, 0.45 ± 0.16, and 0.07 ± 0.01 of ^{124}I-PEG-dMNPs (**Figure 3E**). The retention effect on the tumor was significantly different, indicating that the probe had good specificity for HER2-positive PDX tumor tissues. These results showed that the majority of ^{124}I/^{64}Cu-Her-PEG-dMNPs remained in the tumor at 2 h after injection, and a little part was distributed to the liver over time. The probe uptake was low in the liver and gradually decreased after 48 h, which was significantly different from the control group (^{124}I-PEG-dMNPs).

PET/MR Imaging and Analysis

Manganese (Mn) as a natural cell component has good biosafety and can be used in cellular and mitochondrial functions to avoid long-term side effects (33). Mn^{2+} saturation can increase the T1-weighted relaxation rate and reduce the dosage of the contrast agent (34). (^{124}I, Mn)-Her-PEG-dMNPs can be used for long-term multimodal imaging monitoring over the course of treatment. Therefore, we attempted to use PET/MRI dual enhancement imaging to monitor the high tumor uptake and retention of (^{124}I, Mn)-Her-PEG-dMNPs. The gastric cancer PDX tumor-bearing mice were used for imaging monitoring (n = 3). There were significant PET/MRI signals in the tumor site at 2 h after injection of the nanoprobe, and the signals were not concentrated in the liver or other phagocytic organs at 2, 24, and 72 h (**Figure 4A**). The T1-weighted MR signal of the tumor site slightly declined after 72 h, but the PET signal also showed strong retention at the tumor site.

Statistical analysis of the SUVmax and MR signals in the tumor site showed that the tumor/non-tumor (T/NT) values of

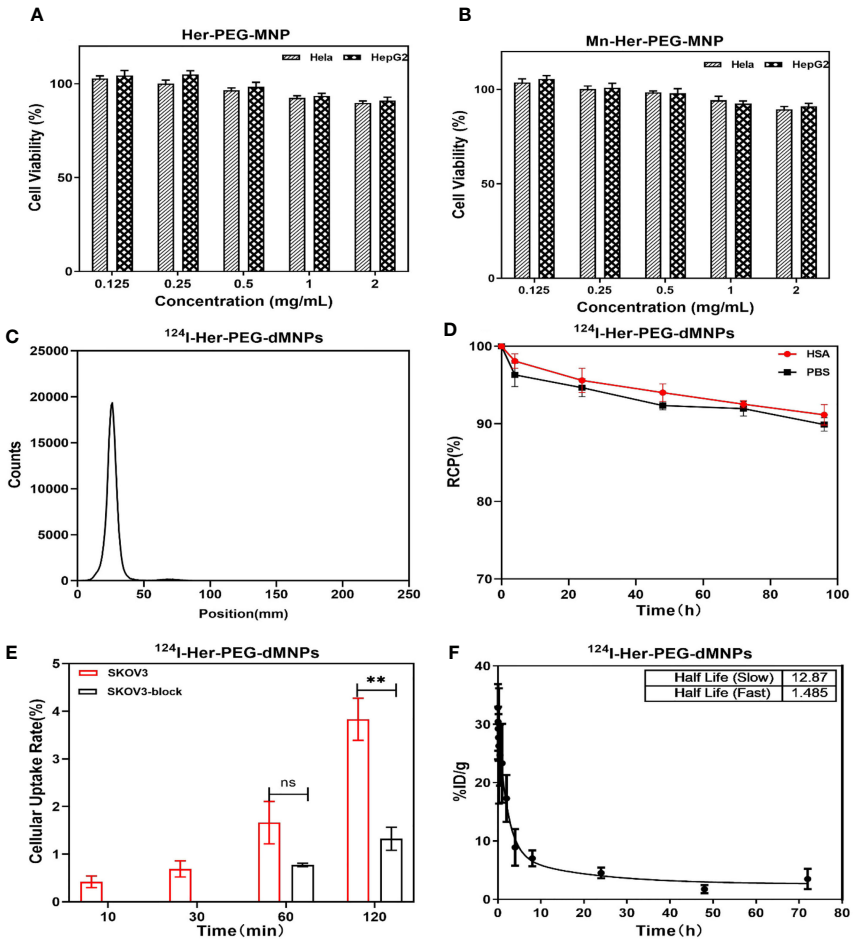


FIGURE 2 | MTT assay using HeLa and HepG2 cells with **(A)** Her-PEG-dMNP and **(B)** Mn-Her-PEG-dMNP concentrations of 0.125, 0.25, 0.5, 1, and 2 mg/ml after incubation at 37°C. **(C)** Radio-TLC chromatogram of ^{124}I-Her-PEG-dMNPs. **(D)** The *in-vitro* stability of ^{124}I-labeled Her-PEG-dMNPs in PBS (pH = 7.4) and 5% HSA. **(E)** *In-vitro* cell uptake of ^{124}I-Her-PEG-dMNPs in SKOV-3 cells. **(F)** Blood radioactivity profile of ^{124}I-Her-PEG-dMNPs. Data were expressed as mean ± SEM. All the numerical data are presented as mean ± SD. *P < 0.05; **P < 0.01; ***P < 0.001 by two-way ANOVA. The symbols ** indicates statistically significant.

the two signals decreased slowly (**Figure 4B**). However, the SUVmax signal and the MR signal in the liver were almost not significantly increased, which was basically consistent with the PET/MR images, indicating the strong retention effect of probes in the tumor site. This study truly realized the practical value of the theranostic functions of (^{124}I, Mn)-Her-PEG-dMNPs.

DISCUSSION

With a half-life of 12.7 h, ^{64}Cu is considered suitable for PET and tumor radionuclide therapy and can be used to radiologically label various peptides and other biologically related small molecules, monoclonal antibodies, proteins, and nanoparticles (35). With a longer half-life of 100.8 h, ^{124}I is suitable for *in-vivo* monitoring of drugs with long circulation (36). ^{64}Cu can label Her-PEG-dMNPs efficiently by adsorption alone. The classic iodine-labeled method with NBS was used as an oxidant to oxidize the active dihydroxyindole/indolequinone ring of Her-PEG-dMNPs for electrophilic substitution reaction labeling of ^{124}I. In the field of animal models, the PDX model is a new generation of tumor model, which was established by inoculating tumor tissue of surgical resection patients into immunodeficient mice. This model retains the histological and genetic characteristics of primary tumors and maintains the heterogeneity of the tumors of patients. The pharmacodynamic results are highly correlated with clinical practice. In this study, we synthesized and characterized a novel ^{124}I/^{64}Cu-Mn-Her-PEG-dMNP nanoprobe and further applied it to the micro-PET/CT and PET/MR imaging in HER2 highly expressing gastric cancer PDX model.

Trastuzumab has been successfully applied in the treatment of HER2 highly expressing gastric cancer, but its potential cardiotoxicity and drug resistance limit its clinical use (10). We have tentatively designed a novel nanoprobe to be coupled with trastuzumab to solve these problems. This nanoprobe fused active targeting and passive targeting, which can greatly improve the targeting efficiency and retention time in the tumor site and reduce the biological half-life of monoclonal

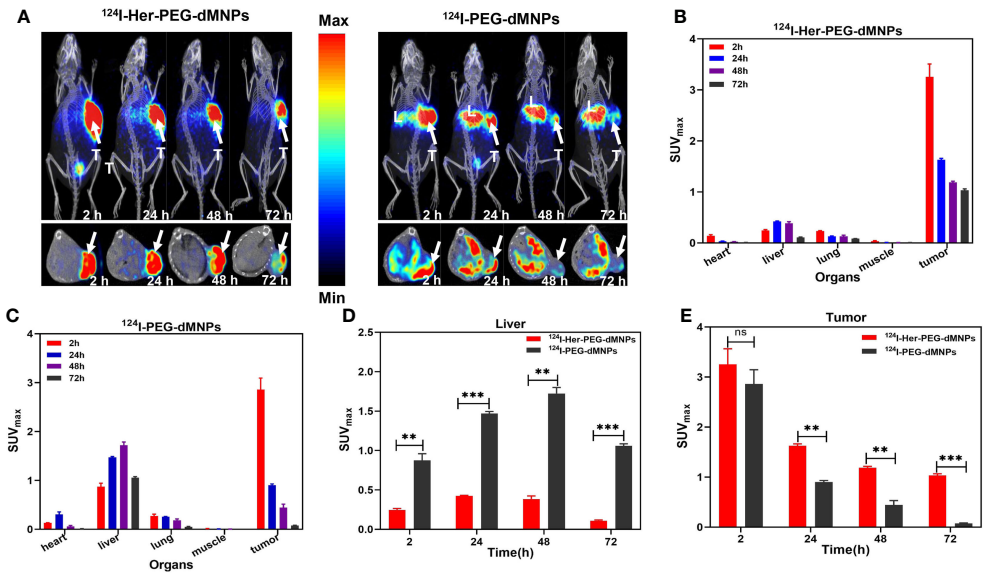


FIGURE 3 | **(A)** Micro-PET/CT imaging of ^{124}I-Her-PEG-dMNPs and ^{124}I-PEG-dMNPs in the gastric cancer PDX model at 2, 24, 48, and 72 h post-injection (the white arrows depicted the tumors; L, liver; T, tumor). **(B, C)** The SUVmax of the organs in the PDX model based on semiquantification of ROIs. The SUVmax of the liver **(D)** and tumor **(E)** based on micro-PET/CT imaging of ROI. The data were expressed as SUVmax ± SEM. All the numerical data are presented as mean ± SD. *P < 0.05; **P < 0.01; ***P < 0.001 by two-way ANOVA. The symbols **, *** indicates statistically significant, very statistically significant.

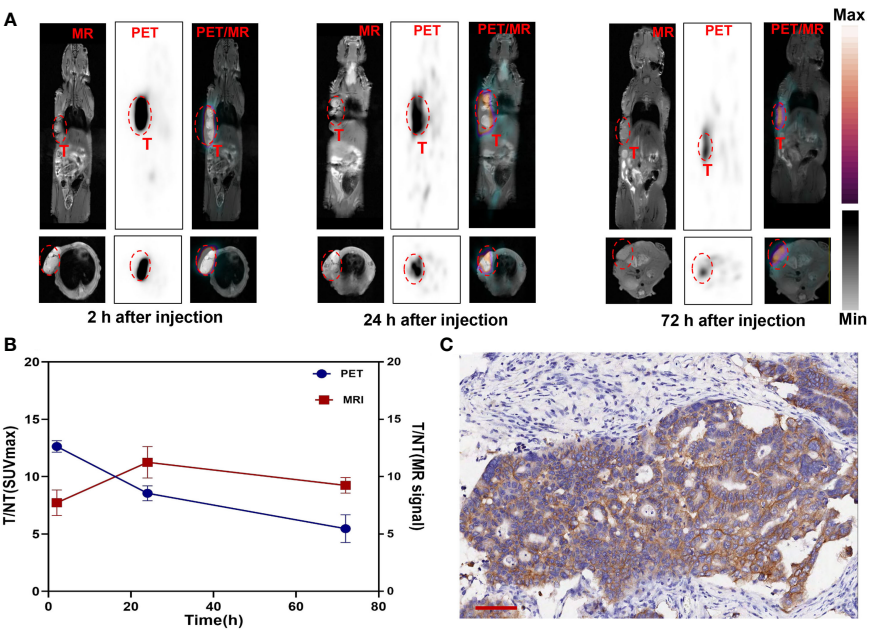


FIGURE 4 | **(A)** T1-weighted MR axial images of gastric cancer PDX model at 2, 24, and 72 h post-injection Mn-Her-PEG-dMNPs (tumor site enveloped by a red dotted line). **(B)** The radioactivity distribution of PET/MRI assessed by the T/NT ratios of SUVmax and MR signal changed over time at the tumor site as determined using the ROI. **(C)** Immunohistochemical staining for gastric cancer PDX model tumor tissues. Scale bar = 100 μm.

antibodies. In addition, the melanin nanoparticles used in this study have photothermal treatment functions, and they can also label therapeutic nuclides to construct integrated theranostics probes to reduce drug resistance during long-term treatment.

The dMNPs were synthesized by the oxidation and self-polymerization of dopamine in a simple method. We selected radionuclide ^{124}I and ^{64}Cu to construct a micro-PET/CT imaging nanoprobe for matching the pharmacokinetic distribution of

Her-PEG-dMNPs *in vivo*. Melanin has a strong chelating property for a variety of metal ions, including Cu^{2+}, Fe^{3+}, and Mn^{2+}. The melanin nanoparticles can perform metal adsorption and nuclide labeling without the introduction of chelating groups, which greatly simplifies the preparation process and reduces the possible probe heterogeneity. Mn^{2+} is a trace element in the human body and can be fully metabolized in living organisms (33). In contrast, although Gd^{3+} has a higher T1-weighted MR relaxation rate than Mn^{2+}, its potential brain toxicity remains an important issue (37).

In this study, the labeling and purification process of ^{124}I-Her-PEG-dMNPs took only 5 min with 98% radiochemical yield, suggesting a more convenient labeling synthesis with higher yield. However, the labeling and purification process of ^{64}Cu-Her-PEG-dMNPs took 1 h, and the average labeling rate was lower than ^{124}I-Her-PEG-dMNPs. After 96 h of incubation, the average radiochemical yield reached 89% of ^{124}I-Her-PEG-dMNPs and ^{64}Cu-Her-PEG-dMNPs, which proved that both radioactive probes have good stability *in vitro*. In the cell uptake experiment, the uptake of ^{124}I/^{64}Cu-Her-PEG-dMNPs in SKOV-3 cells, which have high HER2 expression, could be blocked by excessive Her-PEG-dMNPs. While both probes showed preliminary specificity for HER2 in SKOV-3 cells, ^{124}I-Her-PEG-dMNPs showed superior contrast due to better matching of full-length antibody pharmacokinetics, which has better application prospects. The micro-PET/CT and PET/MRI of different nanoprobes indicate that the ^{124}I-Her-PEG-dMNP probe has high specificity for the PDX model and also conforms to the EPR effect of nanoparticles (38). In MRI imaging, the magnetic resonance signal of the tumor site was significantly enhanced with the passage of time, which proved that Mn-Her-PEG-dMNPs had a good contrast effect. At present, the clinical application of MRI-targeted contrast agents have been rarely reported, which is related to the poor sensitivity of traditional contrast agents, and only excessive contrast agent injection can obtain the enhancement effect of image visibility. The new multimodal imaging probe constructed in this study was based on the high drug loading rate of nanoparticles, which significantly improved the load of the MRI contrast agent, and was expected to provide guidance for further development of magnetic resonance-targeted contrast agents.

CONCLUSIONS

We developed novel nanoprobes targeting a HER2 expression gastric cancer PDX model with micro-PET/CT and PET/MRI imaging to improve the retention time of monoclonal antibodies in tumor sites, reduce cardiac toxicity, and monitor the therapeutic effect in real time through multimodal imaging. Then, the safety of the nanoprobes was verified by cytotoxicity characterization, and its specificity and metabolism were evaluated by *in-vitro* cell uptake and pharmacokinetics. These results support the possibility of using the nanoprobe for the diagnosis of theranostics in gastric cancer.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The animal study was reviewed and approved by Peking University Cancer Hospital Animal Care and Use Committee.

AUTHOR CONTRIBUTIONS

ZY and HZ designed the study. XY, XG, and FW analyze data. LW and LX completed the experiment and wrote the paper together. All authors discuss the results and critique of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported, in part, by the National Natural Science Foundation of China (Nos.81960538, 81671733 and 81871386), the Science and Technology Foundation of Guizhou (No. [2019] 1021 and ZK[2021]471), Beijing Municipal Administration of Hospitals-Yangfan Project (ZYLX201816), Beijing Millions of Talent Projects A level funding (2019A38), Science Foundation of Peking University Cancer Hospital (2021-17), and open project funded by Key laboratory of Carcinogenesis and Translational Research) Ministry of Education/Beijing (2019 Open Project- 06).

ACKNOWLEDGMENTS

We gratefully appreciate all of the chemists, nurses, and technicians from the Department of Nuclear Medicine, Peking University Cancer Hospital, for their contributions to tracer administration and micro-PET/CT imaging.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.778728/full#supplementary-material

REFERENCES

1. Gschwind A, Fischer OM, Ullrich A. The Discovery of Receptor Tyrosine Kinases: Targets for Cancer Therapy. *Nat Rev Cancer* (2004) 4(5):361–70. doi: 10.1038/nrc1360
2. Akiyama T, Sudo C, Ogawara H, Toyoshima K, Yamamoto T. The Product of the Human c-erbB-2 Gene: A 185-Kilodalton Glycoprotein With Tyrosine Kinase Activity. *Sci* (1986) 232(4758):1644–6. doi: 10.1126/science.3012781
3. Yan M, Schwaederle M, Arguello D, Millis SZ, Gatalica Z, Kurzrock R. HER2 Expression Status in Diverse Cancers: Review of Results From 37,992 Patients. *Cancer Metastasis Rev* (2015) 34(1):157–64. doi: 10.1007/s10555-015-9552-6
4. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human Breast Cancer: Correlation of Relapse and Survival With Amplification of the

HER-2/Neu Oncogene. *Sci* (1987) 235(4785):177–82. doi: 10.1126/science.3798106
5. Scholl S, Beuzeboc P, Pouillart P. Targeting HER2 in Other Tumor Types. *Ann Oncol* (2001) 12(Suppl 1):S81–7. doi: 10.1093/annonc/12
6. Gebhart G, Flamen P, De Vries EG, Jhaveri K, Wimana Z. Imaging Diagnostic and Therapeutic Targets: Human Epidermal Growth Factor Receptor 2. *J Nucl Med* (2016) 57(Suppl 1):81S–8S. doi: 10.2967/jnumed.115.157941
7. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
8. Patel TH, Cecchini M. Targeted Therapies in Advanced Gastric Cancer. *Curr Treat Options Oncol* (2020) 21(9):70. doi: 10.1007/s11864-020-00774-4
9. Meric-Bernstam F, Johnson AM, Dumbrava EEI, Raghav K, Balaji K, Bhatt M, et al. Advances in HER2-Targeted Therapy: Novel Agents and Opportunities Beyond Breast and Gastric Cancer. *Clin Cancer Res* (2019) 25(7):2033–41. doi: 10.1158/1078-0432
10. Boku N. HER2-Positive Gastric Cancer. *Gastric Cancer* (2014) 17(1):1–12. doi: 10.1007/s10120-013-0252-z
11. Xavier C, Blykers A, Vaneycken I, D'Huyvetter M, Heemskerk J, Lahoutte T, et al. (18)F-Nanobody for PET Imaging of HER2 Overexpressing Tumors. *Nucl Med Biol* (2016) 43(4):247–52. doi: 10.1016/j.nucmedbio.2016.01.002
12. Xavier C, Vaneycken I, D'huyvetter M, Heemskerk J, Keyaerts M, Vincke C, et al. Synthesis, Preclinical Validation, Dosimetry, and Toxicity of 68Ga-NOTA-Anti-HER2 Nanobodies for iPET Imaging of HER2 Receptor Expression in Cancer. *J Nucl Med* (2013) 54(5):776–84. doi: 10.2967/jnumed.112.111021
13. Dijkers EC, Oude Munnink TH, Kosterink JG, Brouwers AH, Jager PL, de Jong JR, et al. Biodistribution of 89Zr-Trastuzumab and PET Imaging of HER2-Positive Lesions in Patients With Metastatic Breast Cancer. *Clin Pharmacol Ther* (2010) 87(5):586–92. doi: 10.1038/clpt.2010.12
14. Xu Y, Wang L, Pan D, Yan J, Wang X, Yang R, et al. Synthesis of a Novel 89Zr-Labeled HER2 Affibody and Its Application Study in Tumor PET Imaging. *EJNMMI Res* (2020) 10(1):58. doi: 10.1186/s13550-020-00649-7
15. Lam K, Chan C, Reilly RM. Development and Preclinical Studies of 64Cu-NOTA-Pertuzumab F(ab')2 for Imaging Changes in Tumor HER2 Expression Associated With Response to Trastuzumab by PET/CT. *MAbs* (2017) 9 (1):154–64. doi: 10.1080/19420862.2016.1255389
16. Qi S, Hoppmann S, Xu Y, Cheng Z. PET Imaging of HER2-Positive Tumors With Cu-64-Labeled Affibody Molecules. *Mol Imaging Biol* (2019) 1(5):907–16. doi: 10.1007/s11307-018-01310-5
17. Guo X, Zhu H, Zhou N, Chen Z, Liu T, Liu F, et al. Noninvasive Detection of HER2 Expression in Gastric Cancer by 64Cu-NOTA-Trastuzumab in PDX Mouse Model and in Patients. *Mol Pharm* (2018) 15(11):5174–82. doi: 10.1021/acs.molpharmaceut
18. Ulaner GA, Lyashchenko SK, Riedl C, Ruan S, Zanzonico PB, Lake D, et al. First-In-Human Human Epidermal Growth Factor Receptor 2-Targeted Imaging Using 89zr-Pertuzumab PET/CT: Dosimetry and Clinical Application in Patients With Breast Cancer. J. *Nucl Med* (2018) 59(6):900–6. doi: 10.2967/jnumed.117
19. Lee I, Lim I, Byun BH, Kim BI, Choi CW, Woo SK, et al. A Preliminary Clinical Trial to Evaluate 64Cu-NOTA-Trastuzumab as a Positron Emission Tomography Imaging Agent in Patients With Breast Cancer. *EJNMMI Res* (2021) 11(1):8. doi: 10.1186/s13550-021-00746-1
20. Jerusalem G, Lancellotti P, Kim SB. HER2+ Breast Cancer Treatment and Cardiotoxicity: Monitoring and Management. *Breast Cancer Res Treat* (2019) 177(2):237–50. doi: 10.1007/s10549-019-05303-y
21. Yamaguchi H, Tsuchimochi M, Hayama K, Kawase T, Tsubokawa N. Dual-Labeled Near-Infrared/(99m)Tc Imaging Probes Using PAMAM-Coated Silica Nanoparticles for the Imaging of HER2-Expressing Cancer Cells. *Int J Mol Sci* (2016) 17(7):1086. doi: 10.3390/ijms17071086
22. Gerson JN, Skariah S, Denlinger CS, Astsaturov I. Perspectives of HER2-Targeting in Gastric and Esophageal Cancer. *Expert Opin Investig Drugs* (2017) 26(5):531–40. doi: 10.1080/13543784
23. Farahavar G, Abolmaali SS, Gholijani N, Nejatollahi F. Antibody-Guided Nanomedicines as Novel Breakthrough Therapeutic, Diagnostic and Theranostic Tools. *Biomater Sci* (2019) 7(10):4000–16. doi: 10.1039/c9bm00931k
24. Demir E. A Review on Nanotoxicity and Nanogenotoxicity of Different Shapes of *Nanomaterials*. *J Appl Toxicol* (2021) 41(1):118–47. doi: 10.1002/jat.4061
25. Nagy A, Robbins NL. The Hurdles of Nanotoxicity in Transplant Nanomedicine. *Nanomed (Lond)* (2019) 14(20):2749–62. doi: 10.2217/nnm-2019-0192
26. Khalid K, Tan X, Mohd Zaid HF, Tao Y, Lye Chew C, Chu DT, et al. Advanced in Developmental Organic and Inorganic Nanomaterial: A Review. *Bioengineered* (2020) 11(1):328–55. doi: 10.1080/21655979.2020
27. Caldas M, Santos AC, Veiga F, Rebelo R, Reis RL, Correlo VM. Melanin Nanoparticles as a Promising Tool for Biomedical Applications - A Review. *Acta Biomater* (2020) 105:26–43. doi: 10.1016/j.actbio
28. Sun Y, Hong S, Ma X, Cheng K, Wang J, Zhang Z, et al. Recyclable Cu(i)/melanin Dots for Cycloaddition, Bioconjugation and Cell Labelling. *Chem Sci* (2016) 7(9):5888–92. doi: 10.1039/c6sc01536k
29. Simon JD. Spectroscopic and Dynamic Studies of the Epidermal Chromophores Trans-Urocanic Acid and Eumelanin. *Acc Chem Res* (2000) 33(5):307–13. doi: 10.1021/ar970250t
30. Liu Y, Ai K, Liu J, Deng M, He Y, Lu L. Dopamine-Melanin Colloidal Nanospheres: An Efficient Near-Infrared Photothermal Therapeutic Agent for *In Vivo* Cancer Therapy. *Adv Mater* (2013) 25(9):1353–9. doi: 10.1002/adma.201204683
31. Xu C, Shi S, Feng L, Chen F, Graves SA, Ehlerding EB, et al. Long Circulating Reduced Graphene Oxide-Iron Oxide Nanoparticles for Efficient Tumor Targeting and Multimodality Imaging. *Nanoscale* (2016) 8(25):12683–92. doi: 10.1039/c5nr09193d
32. Xie Q, Zhu H, Wang F, Meng X, Ren Q, Xia C, et al. Establishing Reliable Cu-64 Production Process: From Target Plating to Molecular Specific Tumor Micro-PET Imaging. *Molecules* (2017) 22(4):641. doi: 10.3390/molecules22040641
33. Xu W, Sun J, Li L, Peng X, Zhang R, Wang B. Melanin-Manganese Nanoparticles With Ultrahigh Efficient Clearance *In Vivo* for Tumor-Targeting T1 Magnetic Resonance Imaging Contrast Agent. *Biomater Sci* (2017) 6(1):207–15. doi: 10.1039/c7bm00635g
34. Xia L, Guo X, Liu T, Xu X, Jiang J, Wang F, et al. Multimodality Imaging of Naturally Active Melanin Nanoparticles Targeting Somatostatin Receptor Subtype 2 in Human Small-Cell Lung Cancer. *Nanoscale* (2019) 11 (30):14400–9. doi: 10.1039/c9nr04371c
35. Zhou H, Zhang Q, Cheng Y, Xiang L, Shen G, Wu X, et al. 64Cu-Labeled Melanin Nanoparticles for PET/CT and Radionuclide Therapy of Tumor. *Nanomed* (2020) 29:102248. doi: 10.1016/j.nano
36. Lubberink M, Herzog H. Quantitative Imaging of 124I and 86Y With PET. *Eur J Nucl Med Mol Imaging* (2011) 38 Suppl 1(Suppl 1):S10–8. doi: 10.1007/s00259-011-1768-2
37. Kanda T, Fukusato T, Matsuda M, Toyoda K, Oba H, Kotoku J, et al. Gadolinium-Based Contrast Agent Accumulates in the Brain Even in Subjects Without Severe Renal Dysfunction: Evaluation of Autopsy Brain Specimens With Inductively Coupled Plasma Mass Spectroscopy. *Radiol* (2015) 276(1):228–32. doi: 10.1148/radiol.2015142690
38. Fan Q, Cheng K, Hu X, Ma X, Zhang R, Yang M, et al. Transferring Biomarker Into Molecular Probe: Melanin Nanoparticle as a Naturally Active Platform for Multimodality Imaging. *J Am Chem Soc* (2014) 136(43):15185–94. doi: 10.1021/ja505412p

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Wen, Xia, Guo, Huang, Wang, Yang, Yang and Zhu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 19 November 2021
doi: 10.3389/fonc.2021.774117

Radiomics Analysis of Contrast-Enhanced CT for the Preoperative Prediction of Microvascular Invasion in Mass-Forming Intrahepatic Cholangiocarcinoma

Fei Xiang[1†], *Shumei Wei*[2†], *Xingyu Liu*[1], *Xiaoyuan Liang*[1], *Lili Yang*[3] *and Sheng Yan*[1*]

[1] Department of Hepatobiliary and Pancreatic Surgery, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China, [2] Department of Pathology, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China, [3] Department of Radiology, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

Background: Microvascular invasion (MVI) has been shown to be closely associated with postoperative recurrence and metastasis in patients with intrahepatic cholangiocarcinoma (ICC). We aimed to develop a radiomics prediction model based on contrast-enhanced CT (CECT) to distinguish MVI in patients with mass-forming ICC.

Methods: 157 patients were included and randomly divided into training (n=110) and test (n=47) datasets. Radiomic signatures were built based on the recursive feature elimination support vector machine (Rfe-SVM) algorithm. Significant clinical-radiologic factors were screened, and a clinical model was built by multivariate logistic regression. A nomogram was developed by integrating radiomics signature and the significant clinical risk factors.

Results: The portal phase image radiomics signature with 6 features was constructed and provided an area under the receiver operating characteristic curve (AUC) of 0.804 in the training and 0.769 in the test datasets. Three significant predictors, including satellite nodules (odds ratio [OR]=13.73), arterial hypo-enhancement (OR=4.31), and tumor contour (OR=4.99), were identified by multivariate analysis. The clinical model using these predictors exhibited an AUC of 0.822 in the training and 0.756 in the test datasets. The nomogram combining significant clinical factors and radiomics signature achieved satisfactory prediction efficacy, showing an AUC of 0.886 in the training and 0.80 in the test datasets.

Conclusions: Both CECT radiomics analysis and radiologic factors have the potential for MVI prediction in mass-forming ICC patients. The nomogram can further improve the prediction efficacy.

Keywords: intrahepatic cholangiocarcinoma, microvascular invasion, radiomics, computed tomography, nomogram

OPEN ACCESS

Edited by:
Guang Yang,
Imperial College London,
United Kingdom

Reviewed by:
Jue Jiang,
Memorial Sloan Ketter Cancer Center,
United States
Roberto Cannella,
University of Palermo, Italy
Changqiang Wu,
North Sichuan Medical College, China

***Correspondence:**
Sheng Yan
shengyan@zju.edu.cn

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *11 September 2021*
Accepted: *02 November 2021*
Published: *19 November 2021*

Citation:
Xiang F, Wei S, Liu X, Liang X, Yang L and Yan S (2021) Radiomics Analysis of Contrast-Enhanced CT for the Preoperative Prediction of Microvascular Invasion in Mass-Forming Intrahepatic Cholangiocarcinoma. Front. Oncol. 11:774117. doi: 10.3389/fonc.2021.774117

INTRODUCTION

Intrahepatic cholangiocarcinoma (ICC) arises from the epithelial cells of the intrahepatic bile ducts and is the second leading primary liver malignancy (1). ICC can be classified into intraductal, periductal infiltrative, and mass-forming types with macroscopic growth classification. The mass-forming ICC is the most predominant type, accounting for 80% to 90% of all ICC cases (2, 3). According to statistics, the incidence and mortality of ICC are continuously increasing worldwide over the years (4, 5). Surgery is the mainstay of therapy for localized, resectable ICC; Nevertheless, the prognosis remains unsatisfactory, with 5-year overall survival ranging from 15% to 23% (6, 7). High rates of recurrence and metastasis following resection are the leading causes of poor prognosis. In fact, recurrence is estimated to occur in 45% to 70% of patients with ICC (7–10).

Studies have revealed that microvascular invasion (MVI) is a considerable poor-prognostic factor in ICC. MVI has been shown to be closely associated with postoperative recurrence and metastasis (11–13). Tsukamoto et al. (14) reported that the absence of MVI and lymph node metastasis were the only two independent factors for recurrence-free survival over 5 years after liver resection for ICC. Hu et al. (15) showed that ICC patients with MVI exhibited enhanced aggressive behavior with a higher incidence of adjacent tissue infiltration, organ invasion, and satellite lesions; In addition, patients with MVI had a significantly worse disease-free survival (DFS) than patients without MVI. Moreover, Ercolani et al. (16) demonstrated that patients without MVI significantly experienced favorable median overall survival (OS) time than those patients with MVI in all types of cholangiocarcinoma. In contrast to macrovascular invasion, which can be evaluated by radiologic images, MVI can only be detected by postoperative pathological examination, limiting its value in the clinical setting. Recently, some studies have been performed to identify preoperative predictive markers for ICC patients with MVI. Laboratory parameters, including routine blood tests, liver function, and cancer biomarkers such as ALT, AFP, CA-199, have been screened and defined as predictive indicators (17, 18). However, despite the relative ease of obtaining these data, some parameters are controversial and have to be systematically evaluated in clinical practice. Radiological characteristics, such as tumor morphology, arterial phase enhancement pattern, tumor diameter, and apparent diffusion coefficient (ADC) values, were associated with MVI in ICC patients (18, 19). However, imaging features were assessed subjectively and may lead to interobserver variability.

Radiomics, an omics-based approach allowing for the extraction of quantitative features from raw medical images, has been used to perform objective and quantitative analysis of tumor heterogeneity and cancer phenotype (20). Radiomics has been widely utilized in predicting MVI for patients with hepatocellular carcinoma (HCC), and its feasibility and potential benefits have been proved (21–23). However, few studies have evaluated the prediction value of radiomics for ICC. Zhou et al. (24) exacted features from MRI images and a fusion radiomics signature comprising seven features was established for MVI prediction in ICC patients with an area under the receiver operating characteristic curve (AUC) of 0.85. Contrast-enhanced CT (CECT) is the most common imaging modality for diagnosis and assessment of ICC; however, there are currently no studies that evaluated the radiomics analysis of CECT for MVI prediction in patients with ICC.

Therefore, we aimed to verify whether radiomics analysis based on CECT could be useful to predict MVI in mass-forming ICC. Additionally, clinical-radiologic predictors were also evaluated and compared with radiomics analysis.

MATERIALS AND METHODS

Patient Characteristics

182 patients with pathologic diagnosis of ICC after hepatectomy were retrospectively identified in our hospital from March 2013 to May 2021. Inclusion criteria were as follows: (1) Malignancies were classified as mass-forming ICC; (2) CECT scans performed within two weeks before surgery; (3) MVI status was described on pathology reports. Exclusion criteria were: (1) Pathology-confirmed malignancies were mixed, periductal infiltrative or intraductal growing types of ICC; (2) prior intervention or partial hepatectomy; (3) Lack of contrast-enhanced CT scans or insufficient image quality; (4) Grossly tumor thrombus in the portal vein or bile duct tumor thrombosis. The detailed selection process is described in **Figure 1**. Of the 182 screened patients,157 patients were finally enrolled and randomly assigned to a training dataset (n=110) and a test dataset (n=47), with a split ratio of 7:3. The Ethics committee of our hospital approved the present retrospective study.

Clinical Characteristics and Radiological Evaluation

Demographic covariates were collected and compared between the MVI group and non-MVI group both in training and test datasets, including age, sex, hepatitis B immunology, platelet count (PLT), serum albumin (ALB), serum direct bilirubin (DB), serum total bilirubin (TB), γ-glutamyl transpeptidase (GGT), alkaline phosphatase (ALP), serum alanine aminotransferase (ALT), aspartate aminotransferase (AST), international normalized ratio (INR), prothrombin time (PT), carcinoembryonic antigen (CEA), Carbohydrate antigen199 (CA-199). Radiological features, including the number of segments involved, satellite nodules, lymph node status, intrahepatic duct dilatation, tumor contour, arterial rim enhancement, arterial hypo-enhancement, intratumor vascularity, hepatic capsular retraction, were blindly evaluated and recorded by two readers. The detail of radiological evaluation and our protocol for CECT scan acquisition is described in the **Supplementary Presentation**.

Histology

All resected specimens were examined and cross-checked by at least two senior pathologists. Seven tissues were harvested and examined for MVI diagnosis from the resected specimen, including the central and four sides of tumor tissues and two adjacent non-carcinoma tissues. MVI was defined as the invasion

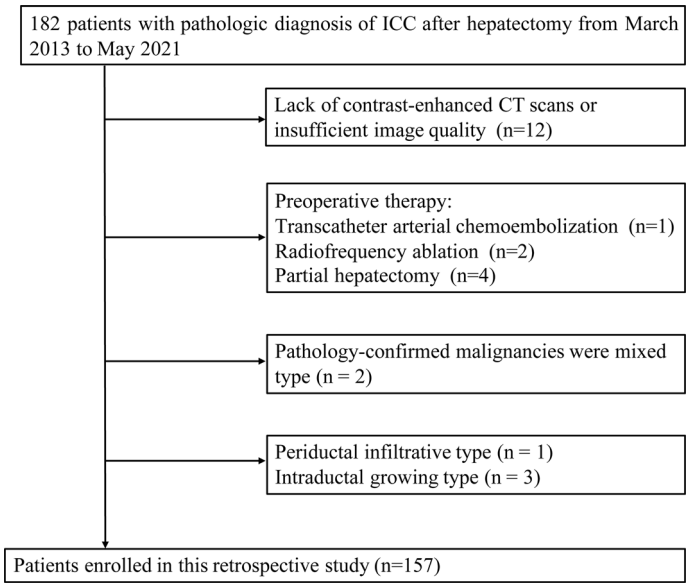


FIGURE 1 | Flow chart of patients recruitment in this study.

of tumor emboli into a vascular space that only can be detected on microscopy (25).

Radiomic Feature Extraction

The radiomics workflow is depicted in **Figure 2**. Region of interests (ROIs) of the whole tumor was contoured on arterial and portal venous phase of CT images using ITK-SNAP software (The specific ROIs segmentation was shown in the **Supplementary Presentation**). The Pyradiomics toolkit was used to extract features from each three-dimensional ROI (26). For each phase, 1130 radiomics features were extracted, including 18 first-order features, 14 shape features, 75 textural features, 279 Laplacian of Gaussian features (sigma=3.0,4.0,5.0), and 744 wavelet features. In total, 2260 features derived from arterial and portal phases were obtained for each patient. The parameter setting for radiomics features extraction and the

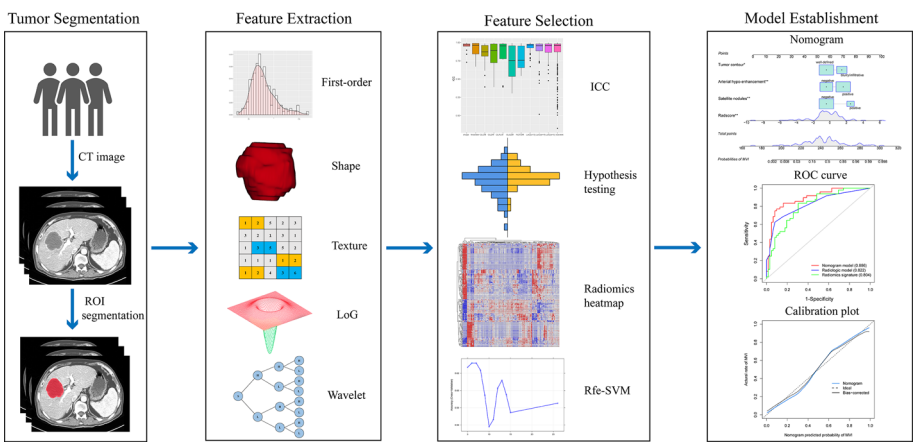


FIGURE 2 | Workflow for the radiomics process. After CT images were acquired, segmentation of the tumor was performed. The extracted radiomics features include first-order, shape, Laplacian of Gaussian, texture, and wavelet features. A four-step approach was performed for feature dimension reduction. Intra- and inter-class correlation coefficient (ICC) was used to evaluate the reproductivity of features. Values lower than 0.8 were eliminated. Student's t-test or Mann–Whitney-U test was performed to find the differential radiomics features. A heatmap shows the Pearson correlation coefficients matrix among radiomics features. Rfe-SVM was applied to develop the radiomics signature. A nomogram was constructed by integrating independent radiological predictors and portal phase image radiomics signature. The nomogram and radiologic model's discriminative ability were compared with the ROC curve analysis and quantified by the AUC. The calibration curve demonstrated good agreement between the actual and nomogram predicted probabilities.

detailed radiomics features is depicted in the **Supplementary Presentation** and **Supplementary Table**.

Radiomics Signature Construction

Before feature dimension reduction, values of radiomics features were transformed to a normal distribution by the z-score method. Then, feature dimension reduction and selection were performed in a four-step approach. Firstly, the intra- and inter-class correlation coefficient was used to evaluate the repeatability of radiomics features between the intra- and inter-observer. Features with intraclass and interclass correlation coefficients lower than 0.80 were eliminated. The specific process was described in the **Supplementary Presentation**. Secondly, Mann–Whitney-U test or Student's t-test was applied to find the differential radiomics features with effects of $p < 0.05$ were considered statistically significant. Thirdly, Pearson correlation coefficient > 0.75 was used as a cutoff to eliminate the highly correlated features. Lastly, a recursive feature elimination support vector machine (Rfe-SVM) method was performed to construct the radiomics signature. The formula of the radiomics signature was displayed in the **Supplementary Presentation**.

According to the process set up above, three radiomics signatures were developed based on features exacted from the arterial phase, portal phase, and pooling features of the two phases. Receiver operating characteristic (ROC) analysis was performed to assess the three signatures' predictive capability.

Development of the Clinical Model and the Nomogram

For constructing the clinical model, univariate analysis was applied to identify independent risk factors (including clinical characteristics and CECT radiological features) between the MVI and the non-MVI groups. The multivariate logistic model was built with significant factors ($p<0.05$) from the univariate analysis as inputs. Odds ratios (OR) and their 95% confidence intervals (CI) were used to express the estimate relative risk. The clinical-radiomics nomogram was built using the optimal radiomics signature and independent clinical-radiologic risk factors in the clinical model. Variance inflation factor (VIF) coefficients were examined to check the collinearity. Comparison of the ROC curves among nomogram model, radiomics signature, and clinical model was performed by using DeLong's test. A calibration curve was plotted to evaluate the calibration of the nomogram.

Statistical Analysis

The Chi-square test or Fisher exact test was used for categorical variables comparison. For continuous variables, Student's t-test or Mann-Whitney U test was used. We considered $p < 0.05$ (two-tailed) as statistically significant. Statistical analysis was conducted with SPSS (v. 26.0), R software (v. 3.6.1), and Python (v. 3.9.2).

RESULTS

Patient Characteristics

The baseline clinical-radiological characteristics are demonstrated in **Table 1**. The incidence of MVI between the training and test datasets shows no statistical difference (48/110 *vs*. 20/47, p=0.900). Patients with MVI were associated with larger tumor size, more liver segments invasion, and blurry or infiltrative tumor contour, which was confirmed in the training and test datasets. Significant differences were found in terms of age, PLT, satellite nodules, lymph node metastasis and arterial hypo-enhancement in the training dataset but not validated in the test dataset. Other baseline variables did not differ between the training and test datasets.

Radiomics Signature Construction

A total of 1130 radiomics features were identified from each imaging phase. For features identified from the arterial phase image, 319 features were excluded with intraclass and interclass correlation coefficients lower than 0.8 (**Figure S1**). Then, 811 features were subjected to statistical hypothesis testing with Student's t-test or Mann–Whitney-U test; 276 features were found significantly different between the MVI positive and non-MVI groups. Pearson correlation analysis found 254 features were highly correlated (correlation coefficient > 0.75) and were eliminated. Only 22 features were kept and subjected to Ref-SVM. Finally, 5 features were selected, and the arterial phase image radiomics signature was built.

The portal phase image radiomics signature with 6 features and fusion radiomics signature with 12 features were built with a similar process. The specific flow and the selected features of three radiomic signatures were shown in the **Figure S2** and **Table 2**. The arterial phase image radiomics signature, portal phase image radiomics signature, and fusion radiomics signature showed good discriminative abilities for MVI prediction, with AUCs of 0.776 (95% CI 0.688–0.863), 0.804 (95% CI 0.723–0.885), and 0.779 (95% CI 0.692–0.865) in the training dataset, and AUCs of 0.726 (95% CI 0.581–0.871), 0.769 (95% CI 0.630–0.908), and 0.763 (95% CI 0.627–0.898), in the test dataset (**Figure 3**). The portal phase image radiomics signature achieved slightly better predictive performance than the other two radiomics signatures, but no statistical differences were found.

Compared to non-MVI group, the MVI positive group had a significantly higher portal phase image Rad-score [median (IQR)] in the training dataset [0.549 (-0.302∼1.244) *vs* −0.916 (-1.829∼-0.050), $P < 0.001$]. Consistent results were obtained in the test dataset [1.151 (-0.075∼1.855) *vs* −0.505 (-1.032∼0.511), P=0.003] (**Figure S3**).

Development of the Radiologic Model and Nomogram

Among all baseline variables, three significant predictors including satellite nodules [OR=13.73 (3.14–59.93), P<0.001], arterial hypo-enhancement [OR=4.31 (1.55–11.94), P=0.005], and tumor contour [OR=4.99 (1.76–14.18), P=0.003] were identified in the training dataset by univariate analysis and further confirmed in multivariate analysis (**Table 3**). The radiologic model developed with the three independent risk factors exhibited AUC of 0.822 (95% CI: 0.741–0.903) and 0.756 (95% CI: 0.616–0.895) in the training and test datasets.

TABLE 1 | Baseline characteristics of patients.

Characteristics	Training dataset			Test dataset		
	MVI-positive (n=48)	MVI-negative (n=62)	p value	MVI-positive (n=20)	MVI-negative (n=27)	p value
Age, years, mean ± SD	61.2 ± 8.8	65.6 ± 8.5	0.008	56.0 ± 12.6	57.5 ± 8.8	0.616
Sex			0.413			0.706
male	27 (56.3)	30 (48.4)		10 (50)	15 (55.6)	
female	21(43.7)	32 (41.6)		10 (50)	12(44.4)	
HBV infection			0.682			0.209
Present	10 (20.8)	11 (17.7)		6 (30.0)	4 (14.8)	
absent	38 (79.2)	51(82.3)		14 (70.0)	23 (85.2)	
PLT, 10^9/L, mean ± SD	202.3 ± 56.6	173.5 ± 70.2	0.023	243.8 ± 116	203.1 ± 68.8	0.173
Alb, g/L, median (IQR)	40.7 (37.1-43.2)	40.0 (36.6-42.7)	0.341	40.4 (35.6-43.6)	42.1 (37.8-44.0)	0.890
TBIL, μmol/L, median (IQR)	13.9 (10.4-23.3)	13.8 (10.6-18.4)	0.109	11.7 (8.7-17.8)	11.1 (9.2-13.2)	0.309
DBIL, μmol/L, median (IQR)	3.0 (2.2-4.9)	2.9 (2.3-4.0)	0.091	2.4 (2.0-5.8)	2.3 (2-2.9)	0.258
ALT, U/L, median (IQR)	27.5 (19.3-48)	23 (15.8-31.3)	0.159	27 (16.5-52.5)	21 (14-36)	0.477
AST, U/L, median (IQR)	31.5 (21.3-48.8)	28 (22-35.6)	0.223	27 (22.3-47.3)	27 (20-34)	0.406
ALP, U/L, median (IQR)	118.5 (90-225.5)	95.5 (76.8-140.8)	0.202	141.5 (106.3-315.0)	134 (111-145)	0.152
GGT, U/L, median (IQR)	81.5 (42.3-179)	46 (31.8-76.3)	0.099	71.5 (55.8-200.8)	83 (44-208)	0.801
PT, mean ± SD	13.0 ± 1.0	13.3 ± 1.0	0.119	13.0 ± 1.5	13.1 ± 1.1	0.959
INR, mean ± SD	1.0 ± 0.10	1.0 ± 0.09	0.603	1.0 ± 0.3	1.0 ± 0.1	0.481
CEA> 5 ug/L	15 (31.3)	17 (27.4)	0.661	9 (45.0)	7 (25.9)	0.172
CA-199>37 ug/L	32 (66.7)	32 (51.6)	0.112	17 (85.0)	12 (44.5)	0.005
Tumor size	5.9 ± 2.6	4.4 ± 2.1	0.002	6.4 ± 2.1	4.6 ± 2.0	0.004
Liver cirrhosis	10 (20.8)	9 (14.5)	0.385	2 (10.0)	5 (18.5)	0.417
No. of segments involved			0.005			0.009
Single	26 (54.2)	50 (80.6)		9 (45.0)	22 (81.5)	
Two or more	22 (45.8)	12 (19.4)		11 (55.0)	5 (18.5)	
Satellite nodules	20 (41.7)	3 (4.8)	<0.001	7 (35.0)	5 (18.5)	0.200
lymph node metastasis	27 (56.3)	8 (12.9)	<0.001	13 (65.0)	11 (40.7)	0.100
Intrahepatic duct dilatation	14 (29.2)	16 (25.8)	0.695	8 (40.0)	7 (25.9)	0.306
Tumor contour			<0.001			0.002
Well-defined	21 (43.8)	52 (83.9)		5 (25.0)	19 (70.4)	
Blurry/infiltrative	27 (56.2)	10 (16.1)		15 (75.0)	8 (29.6)	
Arterial rim- enhancement	13 (27.1)	27 (43.5)	0.075	6 (30.0)	9 (33.3)	0.808
Arterial hypo-enhancement	34 (70.8)	28 (45.2)	0.007	14 (70.0)	17 (63.0)	0.615
Intratumor vascularity	27 (56.3)	34 (54.8)	0.883	5 (25.0)	5 (18.5)	0.591
Hepatic capsular retraction	11 (22.9)	14 (22.6)	0.967	2 (10.0)	6 (22.2)	0.270

MVI, microvascular invasion; HBV, hepatitis B virus; PLT, platelets; Alb, albumin; TBIL, total bilirubin; DBIL, direct bilirubin; ALT, alanine aminotransferase; AST, aspartate transaminase; ALP, alkaline phosphatase; GGT, γ-glutamyl transpeptidase; PT, prothrombin time; INR, international normalized ratio; CEA, carcinoembryonic antigen; CA-199, cancer antigen 19-9; SD, standard deviation; IQR, interquartile range.

We generated the nomogram using the independent predictors in the radiologic model and portal phase image radiomics signature with logistic regression (**Figure 4**). The VIFs for satellite nodules, arterial hypo-enhancement, tumor contour, and portal phase image radiomics signature were less than 10 (satellite nodules: 1.11; arterial hypo-enhancement: 1.14; tumor contour: 1.02; radiomics signature: 1.03), suggesting no collinearity between these variables. The nomogram demonstrated satisfactory prediction efficacy, with an AUC of 0.886 (95% CI: 0.823–0.949) and 0.80 (95% CI: 0.675–0.925) in the training and test datasets. The specific performances of nomogram are shown in **Table 4**. In the training dataset, the nomogram achieved higher AUC than the radiologic model (P =0.011) and portal phase image signature (P = 0.019) (**Figure 5A**). However, there were no statistical differences in the test dataset (nomogram vs radiologic model, portal phase image radiomics signature; P = 0.322, P = 0.642, respectively) (**Figure 5B**).

The calibration plots (**Figures 5C, D**) were consistent between the nomogram prediction and the actual observed probability. Decision curve analysis (**Figure S4**) revealed that the nomogram achieved highest net benefit compared with the radiologic model and radiomics signature. The Hosmer-Lemeshow test showed no significant both in the training (p=0.206) and test datasets (p=0.529), indicating the nomogram was well fitted. A representative case showing the discriminative ability of the nomogram is depicted in **Figure S5**.

DISCUSSION

With increasing recognition of MVI and its prognostic value after hepatectomy, preoperative MVI prediction has become a research hotspot in recent years. In fact, MVI status has been considered as an important event for preoperative surgical decision-making in China. Some studies recommend that anatomical liver resection or wide resection margin is a priority for HCC patients with MVI (27–30). In ICC patients, many studies suggested that a wide surgical margin may lead to favorable survival benefits (7, 31–33). Hu et al. (32) reported that

TABLE 2 | The list of selected features in three radiomics signatures.

Signature	Features selected	Feature name
Arterial phase image signature	5	log-sigma-5-0-mm-3D_firstorder_Variance
		wavelet-LHL_glcm_InverseVariance
		wavelet-LHL_gldm_DependenceVariance
		wavelet-HLH_glcm_Correlation
		wavelet-HHL_glszm_SizeZoneNonUniformity
Portal phase image signature	6	original_firstorder_Skewness
		wavelet-LLH_glcm_Correlation
		wavelet-HLL_glcm_InverseVariance
		wavelet-HHL_glszm_SizeZoneNonUniformity
		wavelet-LLL_glcm_Imc1
		wavelet-LLL_ngtdm_Strength
Fusion radiomics signature	12	PP_original_firstorder_Skewness
		PP_log-sigma-3-0-mm-3D_glszm_GrayLevelVariance
		PP_wavelet-LLH_glcm_Correlation
		PP_wavelet-HLL_glszm_LargeAreaLowGrayLevelEmphasis
		PP_wavelet-HLH_glcm_InverseVariance
		PP_wavelet-HHL_glszm_SizeZoneNonUniformity
		PP_wavelet-LLL_ngtdm_Strength
		AP_log-sigma-4-0-mm-3D_glszm_LargeAreaHighGrayLevelEmphasis
		AP_log-sigma-5-0-mm-3D_firstorder_Kurtosis
		AP_wavelet-LHL_glrlm_GrayLevelNonUniformityNormalized
		AP_wavelet-LLL_firstorder_Mean
		AP_wavelet-LLL_glcm_Contrast

PP, portal phase; AP, arterial phase.

very early recurrence (defined as recurrence within 6 months) after ICC hepatectomy mostly occurred in the surgical margin site. We think that MVI plays an important role in early recurrence after ICC hepatectomy, because MVI generally occurs on the tumor edges. Furthermore, Shao et al. (34) demonstrated that the distance of MVI from the tumor was associated with survival and recurrence in ICC patients, and it seemed that there was an incremental worsening DFS and OS as distance increased.

In this study, the radiomics approach and conventional clinical-radiologic method were used and compared to predict the presence of MVI in ICC patients. The two achieved comparable performance in the prediction of MVI. Furthermore, the nomogram combined radiomics signature and significant radiological factors achieved satisfactory discriminative ability, shown by an AUC of 0.886 in the training dataset and 0.80 in the test dataset. The radiomics features selected in this study were mainly wavelet transformation features consistent with other studies for MVI prediction (21, 22, 24). Among the six features in portal phase image signature, skewness measures the asymmetry of the gray level values distribution. Three GLCM features reflect the linear relation of gray level values or the local image homogeneity. GLSZM size zone nonuniformity reflects the image variability of size zone volumes. NGTDM strength is a measure of the primitives in an image. The wavelet transformation of raw

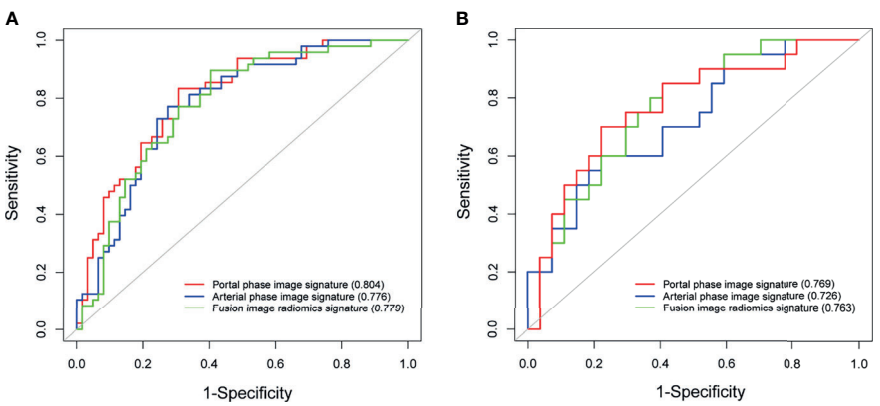


FIGURE 3 | Predictive performance of radiomic signatures for microvascular invasion. ROC curves of radiomic signatures in the training dataset **(A)**. ROC curves of radiomic signatures in the test dataset **(B)**.

TABLE 3 | Univariate and multivariate analyses of risk factors for MVI.

Variables	Univariate analysis			Multivariate analysis		
	OR	95%CI	P-value	OR	95% CI	P-value
Age (≥50 vs <50)	1.067	0.096-11.837	0.958			
Sex (male vs female)	0.571	0.118-2.762	0.486			
HBV infection	6.761	0.841-54.374	0.072			
Alb (>40 vs ≤ 40)	1.712	0.387-7.572	0.479			
TBIL (>50 vs ≤ 50)	2.087	0.329-13.235	0.435			
DBIL (>6.8 vs ≤ 6.8)	5.270	0.251-110.61	0.285			
ALT (>50 vs ≤ 50)	13.122	0.658-261.57	0.092			
AST (>40 vs ≤ 40)	0.219	0.017-2.756	0.240			
ALP (>125 vs ≤ 125)	1.619	0.174-15.069	0.672			
GGT (>50 vs ≤ 50)	1.214	0.206-7.149	0.830			
PT (>13 vs ≤ 13)	0.902	0.077-10.608	0.195			
INR (per 0.1 increase) (>1.0 vs ≤ 1.0)	0.208	0.012-3.619	0.281			
PLT	1.006	0.995-1.017	0.256			
CEA (>5 vs ≤ 5)	0.779	0.129-4.688	0.785			
CA-199 (>37 vs ≤ 37)	1.469	0.371-5.820	0.584			
Tumor size	0.990	0.953-1.029	0.617			
Cirrhosis	0.967	0.086-10.835	0.978			
No. of segments involved (single vs two/more)	2.244	0.508-9.519	0.287			
Satellite nodules	33.154	2.689-408.79	0.006	13.726	3.144-59.93	<0.001
lymph node metastasis	4.386	0.906-21.247	0.066			
Intrahepatic duct dilatation	0.252	0.034-1.882	0.179			
Tumor contour (well-defined vs blurry/infiltrative)	7.535	1.267-42.660	0.026	4.992	1.757-14.18	0.003
Arterial rim- enhancement	2.135	0.315-14.443	0.437			
Arterial hypo-enhancement	20.298	2.365-174.22	0.006	4.308	1.554-11.94	0.005
Intratumor vascularity	0.814	0.213-3.120	0.764			
Hepatic capsular retraction	1.400	0.256-7.662	0.698			

MVI, microvascular invasion; HBV, hepatitis B virus; Alb, albumin; TBIL, total bilirubin; DBIL, direct bilirubin; ALT, alanine aminotransferase; AST, aspartate transaminase; ALP, alkaline phosphatase; OR, odds ratios; GGT, γ-glutamyl transpeptidase; CI, confidence intervals; PT, prothrombin time; INR, international normalized ratio; PLT, platelets; CEA, carcinoembryonic antigen; CA-199, cancer antigen 19-9.

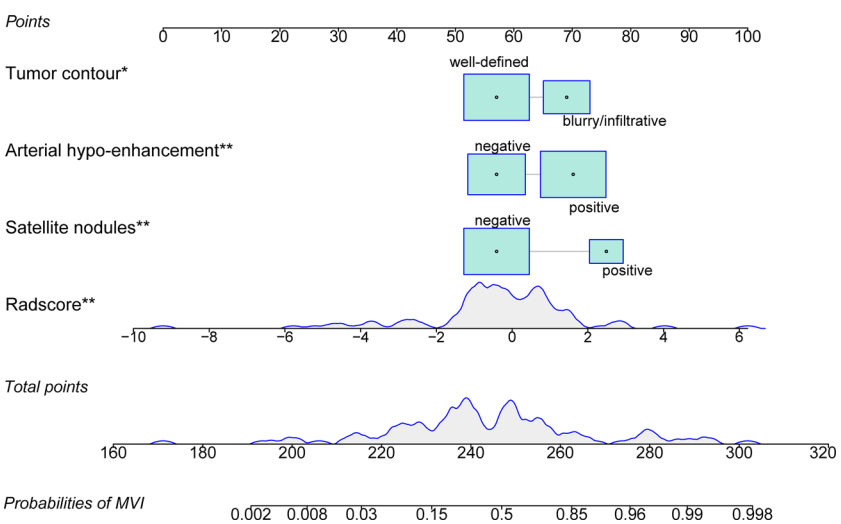


FIGURE 4 | The radiomics nomogram was developed by incorporating the portal phase image radiomics signature, satellite nodules, arterial hypo-enhancement, and tumor contour. *p < 0.05; **p < 0.01.

images can further reflect more accurate spatial changes across multiple dimensions (26). Features included in this study varied from Zhou's result (24), which may be attributed to the different image modalities analyzed.

From the comparison results of baseline characteristics, we found that the number of liver segments invaded, tumor size, tumor contour, age, PLT, satellite nodules, lymph node metastasis, arterial hypo-enhancement showed a significant

TABLE 4 | Performance of nomogram for MVI prediction.

Group	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)	AUC (95%CI)	Cut-off
Training dataset	77.1 (37/48)	90.3 (56/62)	84.5 (93/110)	86.0 (37/43)	83.5 (56/67)	0.886 (0.823–0.949)	>0.157
Test dataset	77.8 (21/27)	75.0 (15/20)	76.6 (36/47)	71.4 (15/21)	80.8 (21/26)	0.800 (0.675–0.925)	>0.157

PPV, positive predictive value; NPV, negative predictive value; AUC, area under the curve.

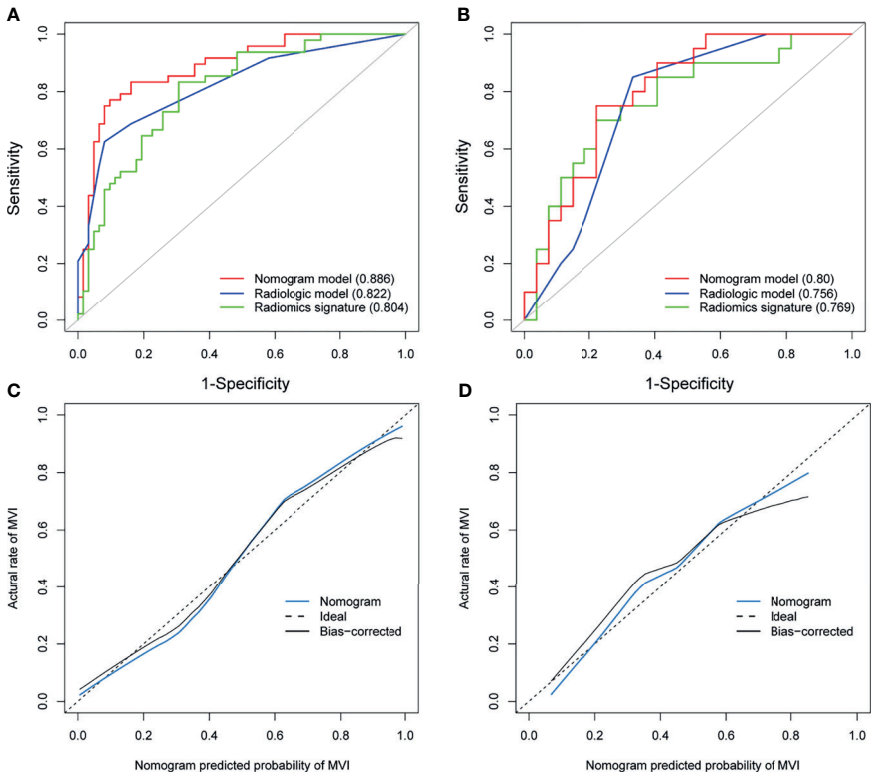


FIGURE 5 | Assessing the discriminative performance of the nomogram and comparison with other predictive models. The nomogram showed a significantly higher discriminative power than the radiomics signature and the radiologic model for the prediction of microvascular invasion in the training dataset **(A)**, but did not differ in the test dataset **(B)**. The calibration plots demonstrate that the nomogram-predicted probabilities were consistent with actual MVI incidence in the training **(C)** and test **(D)** datasets.

difference between the MVI group and non-MVI group in the training dataset. However, only tumor contour, satellite nodules, and arterial hypo-enhancement were selected as significant risk factors by multiple regression. Tumor morphology is a key feature for MVI prediction both in HCC and ICC. It seems that MVI-negative patients are inclined to have a well-defined, spherical, oval tumor contour, whereas those with MVI tend to show irregular, blurry or infiltrative contours (18, 19, 35, 36). Our results are in agreement with these reports. In terms of satellite nodules, micro-metastases invade into the portal venous system and subsequently spread to the tumor-bearing portal territories and eventually develop into microsatellite nodules (37, 38). Such a mechanism has long been accepted as the main cause of intrahepatic metastasis and postoperative recurrence (39, 40). Satellite nodules are common in ICC patients, and the incidence is reported as high as 30% (41–43). The enhancement pattern of the mass-forming ICC in the hepatic arterial phase of CECT could be classified into three types: the hypo-enhancement, the rim-enhancement, and the hyper-enhancement (44). Several studies suggest that the different enhancement patterns of mass-forming ICC show varied prognosis, and the hypo-enhancement is often associated with worse survival (44–46). Yugawa et al. (47) reported that among the three enhancement patterns, the lowest tumor microvessel density was found in the hypo-enhancement ICC and was often accompanied by larger tumor size, more frequent microvascular invasion, and a higher rate of intrahepatic and lymph node metastasis.

Our study has some limitations. Due to the highly aggressive behavior of ICC, only a low number of patients admitted to our hospital have the chance for surgical resection, thus leading to limited sample size in our study. Secondly, retrospective nature of the study may introduce inevitable selection bias. Thirdly, this

is a single-center study with internal validation was performed and in the absence of external validation. Therefore, further refinements with prospective multicenter studies are needed to check out our results.

CONCLUSION

The findings in this study verified that both radiomics analysis and radiologic factors have the potential for MVI prediction in mass-forming ICC patients. The advantage of radiomics is that it can detect microscopic structures and quantitatively measure the microscopic changes in tissue caused by disease. The radiologic method was convenient and fast and also demonstrated good diagnostic efficacy. The combined nomogram, which integrated radiologic factors and radiomics signature, further improved the predictive performance for MVI diagnosis.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Ethics committee of the Second Affiliated Hospital of Zhejiang University School of Medicine. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

FX and SY were responsible for the conception of the work. SW reviewed pathological pictures. FX, XLia and XLiu obtained the data. FX and LY segmented the images. FX and XLia analyzed the data. FX wrote the manuscript. SY critically revised the manuscript. All authors are accountable for the contents of this work. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from the National Natural Science Foundation of China (No.81572975) and Key research and development project of science and technology department of Zhejiang (No.2015C03053).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.774117/full#supplementary-material

REFERENCES

1. Bridgewater J, Galle PR, Khan SA, Llovet JM, Park JW, Patel T, et al. Guidelines for the Diagnosis and Management of Intrahepatic Cholangiocarcinoma. *J Hepatol* (2014) 60(6):1268–89. doi: 10.1016/j.jhep.2014.01.021
2. Chung T, Rhee H, Nahm JH, Jeon Y, Yoo JE, Kim YJ, et al. Clinicopathological Characteristics of Intrahepatic Cholangiocarcinoma According to Gross Morphologic Type: Cholangiolocellular Differentiation Traits and Inflammation- and Proliferation-Phenotypes. *HPB (Oxford)* (2020) 22(6):864–73. doi: 10.1016/j.hpb.2019.10.009
3. Meng ZW, Pan W, Hong HJ, Chen JZ, Chen YL. Macroscopic Types of Intrahepatic Cholangiocarcinoma and the Eighth Edition of AJCC/UICC TNM Staging System. *Oncotarget* (2017) 8(60):101165–74. doi: 10.18632/oncotarget.20932
4. Beal EW, Tumin D, Moris D, Zhang XF, Chakedis J, Dilhoff M, et al. Cohort Contributions to Trends in the Incidence and Mortality of Intrahepatic Cholangiocarcinoma. *Hepatobiliary Surg Nutr* (2018) 7(4):270–6. doi: 10.21037/hbsn.2018.03.16
5. Khan SA, Genus T, Morement H, Murphy A, Rous B, Tataru D. Global Trends in Mortality From Intrahepatic and Extrahepatic Cholangiocarcinoma. *J Hepatol* (2019) 71(6):1261–2. doi: 10.1016/j.jhep.2019.07.024
6. Qu WF, Zhou PY, Liu WR, Tian MX, Jin L, Jiang XF, et al. Age-Adjusted Charlson Comorbidity Index Predicts Survival in Intrahepatic Cholangiocarcinoma Patients After Curative Resection. *Ann Transl Med* (2020) 8(7):487. doi: 10.21037/atm.2020.03.23
7. Spolverato G, Yakoob MY, Kim Y, Alexandrescu S, Marques HP, Lamelas J, et al. The Impact of Surgical Margin Status on Long-Term Outcome After Resection for Intrahepatic Cholangiocarcinoma. *Ann Surg Oncol* (2015) 22(12):4020–8. doi: 10.1245/s10434-015-4472-9
8. Nickkholgh A, Ghamarnejad O, Khajeh E, Tinoush P, Bruckner T, Kulu Y, et al. Outcome After Liver Resection for Primary and Recurrent Intrahepatic Cholangiocarcinoma. *BJS Open* (2019) 3(6):793–801. doi: 10.1002/bjs5.50217
9. Langella S, Russolillo N, Ossola P, Luzzi AP, Casella M, Lo Tesoriere R, et al. Recurrence After Curative Resection for Intrahepatic Cholangiocarcinoma: How to Predict the Chance of Repeat Hepatectomy? *J Clin Med* (2021) 10(13):2820. doi: 10.3390/jcm10132820
10. Spolverato G, Kim Y, Alexandrescu S, Marques HP, Lamelas J, Aldrighetti L, et al. Management and Outcomes of Patients With Recurrent Intrahepatic Cholangiocarcinoma Following Previous Curative-Intent Surgical Resection. *Ann Surg Oncol* (2016) 23(1):235–43. doi: 10.1245/s10434-015-4642-9
11. Tsilimigras DI, Sahara K, Wu L, Moris D, Bagante F, Guglielmi A, et al. Very Early Recurrence After Liver Resection for Intrahepatic Cholangiocarcinoma: Considering Alternative Treatment Approaches. *JAMA Surg* (2020) 155(9):823–31. doi: 10.1001/jamasurg.2020.1973
12. Ali SM, Clark CJ, Mounajjed T, Wu TT, Harmsen WS, Reid-Lombardo KM, et al. Model to Predict Survival After Surgical Resection of Intrahepatic Cholangiocarcinoma: The Mayo Clinic Experience. *HPB (Oxford)* (2015) 17(3):244–50. doi: 10.1111/hpb.12333
13. Wang T, Zhang J, Wang W, Yang X, Kong J, Shen S, et al. Development and Validation of Nomograms for Predicting Cancer-Specific Survival in Elderly Patients With Intrahepatic Cholangiocarcinoma After Liver Resection: A Competing Risk Analysis. *Cancer Manag Res* (2020) 12:11015–29. doi: 10.2147/cmar.S272797
14. Tsukamoto M, Yamashita YI, Imai K, Umezaki N, Yamao T, Okabe H, et al. Predictors of Cure of Intrahepatic Cholangiocarcinoma After Hepatic Resection. *Anticancer Res* (2017) 37(12):6971–5. doi: 10.21873/anticanres.12164
15. Hu LS, Weiss M, Popescu I, Marques HP, Aldrighetti L, Maithel SK, et al. Impact of Microvascular Invasion on Clinical Outcomes After Curative-Intent Resection for Intrahepatic Cholangiocarcinoma. *J Surg Oncol* (2019) 119(1):21–9. doi: 10.1002/jso.25305
16. Ercolani G, Dazzi A, Giovinazzo F, Ruzzenente A, Bassi C, Guglielmi A, et al. Intrahepatic, Peri-Hilar and Distal Cholangiocarcinoma: Three Different Locations of the Same Tumor or Three Different Tumors? *Eur J Surg Oncol* (2015) 41(9):1162–9. doi: 10.1016/j.ejso.2015.05.013

17. Tang Z, Liu WR, Zhou PY, Ding ZB, Jiang XF, Wang H, et al. Prognostic Value and Predication Model of Microvascular Invasion in Patients With Intrahepatic Cholangiocarcinoma. *J Cancer* (2019) 10(22):5575–84. doi: 10.7150/jca.32199
18. Ma X, Liu L, Fang J, Rao S, Lv L, Zeng M, et al. MRI Features Predict Microvascular Invasion in Intrahepatic Cholangiocarcinoma. *Cancer Imaging* (2020) 20(1):40. doi: 10.1186/s40644-020-00318-x
19. Zhou Y, Wang X, Xu C, Zhou G, Liu X, Gao S, et al. Mass-Forming Intrahepatic Cholangiocarcinoma: Can Diffusion-Weighted Imaging Predict Microvascular Invasion? *J Magn Reson Imaging* (2019) 50(1):315–24. doi: 10.1002/jmri.26566
20. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169
21. Xu X, Zhang HL, Liu QP, Sun SW, Zhang J, Zhu FP, et al. Radiomic Analysis of Contrast-Enhanced CT Predicts Microvascular Invasion and Outcome in Hepatocellular Carcinoma. *J Hepatol* (2019) 70(6):1133–44. doi: 10.1016/j.jhep.2019.02.023
22. Zhang X, Ruan S, Xiao W, Shao J, Tian W, Liu W, et al. Contrast-Enhanced CT Radiomics for Preoperative Evaluation of Microvascular Invasion in Hepatocellular Carcinoma: A Two-Center Study. *Clin Transl Med* (2020) 10(2):e111. doi: 10.1002/ctm2.111
23. Yang L, Gu D, Wei J, Yang C, Rao S, Wang W, et al. A Radiomics Nomogram for Preoperative Prediction of Microvascular Invasion in Hepatocellular Carcinoma. *Liver Cancer* (2019) 8(5):373–86. doi: 10.1159/000494099
24. Zhou Y, Zhou G, Zhang J, Xu C, Wang X, Xu P. Radiomics Signature on Dynamic Contrast-Enhanced MR Images: A Potential Imaging Biomarker for Prediction of Microvascular Invasion in Mass-Forming Intrahepatic Cholangiocarcinoma. *Eur Radiol* (2021) 31(9):6846–55. doi: 10.1007/s00330-021-07793-1
25. Cong WM, Bu H, Chen J, Dong H, Zhu YY, Feng LH, et al. Practice Guidelines for the Pathological Diagnosis of Primary Liver Cancer: 2015 Update. *World J Gastroenterol* (2016) 22(42):9279–87. doi: 10.3748/wjg.v22.i42.9279
26. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* (2017) 77(21):e104–7. doi: 10.1158/0008-5472.Can-17-0339
27. Shimada S, Kamiyama T, Yokoo H, Orimo T, Wakayama K, Einama T, et al. Clinicopathological Characteristics of Hepatocellular Carcinoma With Microscopic Portal Venous Invasion and the Role of Anatomical Liver Resection in These Cases. *World J Surg* (2017) 41(8):2087–94. doi: 10.1007/s00268-017-3964-0
28. Zhong XP, Zhang YF, Mei J, Li SH, Kan A, Lu LH, et al. Anatomical Versus Non-Anatomical Resection for Hepatocellular Carcinoma With Microscope Vascular Invasion: A Propensity Score Matching Analysis. *J Cancer* (2019) 10(17):3950–7. doi: 10.7150/jca.32592
29. Hirokawa F, Hayashi M, Miyamoto Y, Asakuma M, Shimizu T, Komeda K, et al. Outcomes and Predictors of Microvascular Invasion of Solitary Hepatocellular Carcinoma. *Hepatol Res* (2014) 44(8):846–53. doi: 10.1111/hepr.12196
30. Yang P, Si A, Yang J, Cheng Z, Wang K, Li J, et al. A Wide-Margin Liver Resection Improves Long-Term Outcomes for Patients With HBV-Related Hepatocellular Carcinoma With Microvascular Invasion. *Surgery* (2019) 165(4):721–30. doi: 10.1016/j.surg.2018.09.016
31. Watanabe Y, Matsuyama Y, Izumi N, Kubo S, Kokudo N, Sakamoto M, et al. Effect of Surgical Margin Width After R0 Resection for Intrahepatic Cholangiocarcinoma: A Nationwide Survey of the Liver Cancer Study Group of Japan. *Surgery* (2020) 167(5):793–802. doi: 10.1016/j.surg.2019.12.009
32. Hu LS, Zhang XF, Weiss M, Popescu I, Marques HP, Aldrighetti L, et al. Recurrence Patterns and Timing Courses Following Curative-Intent Resection for Intrahepatic Cholangiocarcinoma. *Ann Surg Oncol* (2019) 26(8):2549–57. doi: 10.1245/s10434-019-07353-4
33. Liu H, Lin L, Lin Z, Chen Y, Huang Q, Ding L, et al. Impact of Surgical Margin Width on Long-Term Outcomes for Intrahepatic Cholangiocarcinoma: A Multicenter Study. *BMC Cancer* (2021) 21(1):840. doi: 10.1186/s12885-021-08560-7
34. Shao C, Chen J, Chen J, Shi J, Huang L, Qiu Y. Histological Classification of Microvascular Invasion to Predict Prognosis in Intrahepatic Cholangiocarcinoma. *Int J Clin Exp Pathol* (2017) 10(7):7674–81.
35. Renzulli M, Brocchi S, Cucchetti A, Mazzotti F, Mosconi C, Sportoletti C, et al. Can Current Preoperative Imaging Be Used to Detect Microvascular Invasion of Hepatocellular Carcinoma? *Radiology* (2016) 279(2):432–42. doi: 10.1148/radiol.2015150998
36. Meng XP, Wang YC, Zhou JY, Yu Q, Lu CQ, Xia C, et al. Comparison of MRI and CT for the Prediction of Microvascular Invasion in Solitary Hepatocellular Carcinoma Based on a Non-Radiomics and Radiomics Method: Which Imaging Modality Is Better? *J Magn Reson Imaging* (2021) 54(2):526–36. doi: 10.1002/jmri.27575
37. Makuuchi M, Hasegawa H, Yamazaki S. Ultrasonically Guided Subsegmentectomy. *Surg Gynecol Obstet* (1985) 161(4):346–50.
38. Shindoh J, Makuuchi M, Matsuyama Y, Mise Y, Arita J, Sakamoto Y, et al. Complete Removal of the Tumor-Bearing Portal Territory Decreases Local Tumor Recurrence and Improves Disease-Specific Survival of Patients With Hepatocellular Carcinoma. *J Hepatol* (2016) 64(3):594–600. doi: 10.1016/j.jhep.2015.10.015
39. Zhou XP, Quan ZW, Cong WM, Yang N, Zhang HB, Zhang SH, et al. Micrometastasis in Surrounding Liver and the Minimal Length of Resection Margin of Primary Liver Cancer. *World J Gastroenterol* (2007) 13(33):4498–503. doi: 10.3748/wjg.v13.i33.4498
40. Shindoh J, Hasegawa K, Inoue Y, Ishizawa T, Nagata R, Aoki T, et al. Risk Factors of Post-Operative Recurrence and Adequate Surgical Approach to Improve Long-Term Outcomes of Hepatocellular Carcinoma. *HPB (Oxford)* (2013) 15(1):31–9. doi: 10.1111/j.1477-2574.2012.00552.x
41. Schiffman SC, Nowacki MR, Spencer L, McMasters KM, Scoggins CR, Martin RC. Molecular Factors Associated With Recurrence and Survival Following Hepatectomy in Patients With Intrahepatic Cholangiocarcinoma: A Guide to Adjuvant Clinical Trials. *J Surg Oncol* (2014) 109(2):98–103. doi: 10.1002/jso.23459
42. Tarchi P, Tabrizian P, Prigoff J, Schwartz M. Outcomes of Resection for Solitary ≤5 Cm Intrahepatic Cholangiocarcinoma. *Surgery* (2018) 163(4):698–702. doi: 10.1016/j.surg.2017.09.058
43. Zhao YJ, Chen WX, Wu DS, Zhang WY, Zheng LR. Differentiation of Mass-Forming Intrahepatic Cholangiocarcinoma From Poorly Differentiated Hepatocellular Carcinoma: Based on the Multivariate Analysis of Contrast-Enhanced Computed Tomography Findings. *Abdom Radiol (NY)* (2016) 41(5):978–89. doi: 10.1007/s00261-015-0629-z
44. Fujita N, Asayama Y, Nishie A, Ishigami K, Ushijima Y, Takayama Y, et al. Mass-Forming Intrahepatic Cholangiocarcinoma: Enhancement Patterns in the Arterial Phase of Dynamic Hepatic CT - Correlation With Clinicopathological Findings. *Eur Radiol* (2017) 27(2):498–506. doi: 10.1007/s00330-016-4386-3
45. Teraoku H, Morine Y, Uyama N, Ikemoto T, Iwahashi S, Saito Y, et al. Role of Central Hypo-Enhancement in the Hepatic Arterial Phase of Dynamic Computed Tomography in Patients With Mass-Forming Intrahepatic Cholangiocarcinoma. *World J Surg* (2020) 44(7):2350–8. doi: 10.1007/s00268-020-05456-8
46. Kim SA, Lee JM, Lee KB, Kim SH, Yoon SH, Han JK, et al. Intrahepatic Mass-Forming Cholangiocarcinomas: Enhancement Patterns at Multiphasic CT, With Special Emphasis on Arterial Enhancement Pattern–Correlation With Clinicopathologic Findings. *Radiology* (2011) 260(1):148–57. doi: 10.1148/radiol.11101777
47. Yugawa K, Itoh S, Yoshizumi T, Iseda N, Tomiyama T, Toshima T, et al. Prognostic Impact of Tumor Microvessels in Intrahepatic Cholangiocarcinoma: Association With Tumor-Infiltrating Lymphocytes. *Mod Pathol* (2021) 34(4):798–807. doi: 10.1038/s41379-020-00702-9

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Xiang, Wei, Liu, Liang, Yang and Yan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 02 December 2021
doi: 10.3389/fonc.2021.781798

Automatic Sequence-Based Network for Lung Diseases Detection in Chest CT

Jinkui Hao[1,2†], Jianyang Xie[1†], Ri Liu[3†], Huaying Hao[1], Yuhui Ma[1,2], Kun Yan[3], Ruirui Liu[4], Yalin Zheng[5], Jianjun Zheng[4], Jiang Liu[1,6], Jingfeng Zhang[3] and Yitian Zhao[1,7,8*]*

[1] Cixi Institute of Biomedical Engineering, Ningbo Institute of Material Technology and Engineering, Chinese Academy of Sciences, Ningbo, China, [2] School of Optical Technology, University of Chinese Academy of Sciences, Beijing, China, [3] Hwa Mei Hospital, University of Chinese Academy of Sciences, Ningbo, China, [4] School of Medicine, Ningbo University, Ningbo, China, [5] Department of Eye and Vision Science, University of Liverpool, Liverpool, United Kingdom, [6] Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China, [7] Zhejiang International Scientific and Technological Cooperative Base of Biomedical Materials and Technology, Ningbo Institute of Material Technology and Engineering, Chinese Academy of Sciences, Ningbo, China, [8] Zhejiang Engineering Research Center for Biomedical Materials, Ningbo Institute of Material Technology and Engineering, Chinese Academy of Sciences, Ningbo, China

OPEN ACCESS

Edited by:
Guang Yang,
Imperial College London,
United Kingdom

Reviewed by:
Anand Nayyar,
Duy Tan University, Vietnam
Seyedali Mirjalili,
Torrens University Australia, Australia
Tao Zhou,
Nanjing University of Science and Technology, China
Zhili Chen,
Shenyang Jianzhu University, China

***Correspondence:**
Yitian Zhao
yitian.zhao@nimte.ac.cn
Jingfeng Zhang
jingfeng.zhang@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 23 September 2021
Accepted: 01 November 2021
Published: 02 December 2021

Citation:
Hao J, Xie J, Liu R, Hao H, Ma Y, Yan K, Liu R, Zheng Y, Zheng J, Liu J, Zhang J and Zhao Y (2021) Automatic Sequence-Based Network for Lung Diseases Detection in Chest CT. Front. Oncol. 11:781798. doi: 10.3389/fonc.2021.781798

Objective: To develop an accurate and rapid computed tomography (CT)-based interpretable AI system for the diagnosis of lung diseases.

Background: Most existing AI systems only focus on viral pneumonia (e.g., COVID-19), specifically, ignoring other similar lung diseases: e.g., bacterial pneumonia (BP), which should also be detected during CT screening. In this paper, we propose a unified sequence-based pneumonia classification network, called SLP-Net, which utilizes consecutiveness information for the differential diagnosis of viral pneumonia (VP), BP, and normal control cases from chest CT volumes.

Methods: Considering consecutive images of a CT volume as a time sequence input, compared with previous 2D slice-based or 3D volume-based methods, our SLP-Net can effectively use the spatial information and does not need a large amount of training data to avoid overfitting. Specifically, sequential convolutional neural networks (CNNs) with multi-scale receptive fields are first utilized to extract a set of higher-level representations, which are then fed into a convolutional long short-term memory (ConvLSTM) module to construct axial dimensional feature maps. A novel adaptive-weighted cross-entropy loss (ACE) is introduced to optimize the output of the SLP-Net with a view to ensuring that as many valid features from the previous images as possible are encoded into the later CT image. In addition, we employ sequence attention maps for auxiliary classification to enhance the confidence level of the results and produce a case-level prediction.

Results: For evaluation, we constructed a dataset of 258 chest CT volumes with 153 VP, 42 BP, and 63 normal control cases, for a total of 43,421 slices. We implemented a comprehensive comparison between our SLP-Net and several state-of-the-art methods across the dataset. Our proposed method obtained significant performance without a

large amount of data, outperformed other slice-based and volume-based approaches. The superior evaluation performance achieved in the classification experiments demonstrated the ability of our model in the differential diagnosis of VP, BP and normal cases.

Keywords: deep learning, CT, CNN, ConvLSTM, lung diseases

1 INTRODUCTION

COVID-19, the latest in viral pneumonia diseases, is an acute respiratory syndrome that has spread rapidly around the world since the end of 2019, having a devastating effect on the health and well-being of the global population (1, 2). To diagnose viral pneumonia (limited to COVID-19 in our work), reverse transcription-polymerase chain reaction (RT-PCR) has widely been accepted as the gold standard. However, shortages of equipment and strict requirements for testing environments limit the rapid and accurate screening of suspected subjects. Furthermore, RT-PCR testing is also reported to suffer from a high false-negative rate (3), with a low sensitivity of only 71%. In clinical practice, radiological imaging techniques, e.g., X-rays and computed tomography (CT), have also been demonstrated to be effective in diagnosis, and also follow-up assessment and evaluation of disease evolution (4, 5). CT is the most widely used imaging technique, due to its high resolution and three-dimensional (3D) view, and its relatively high detection sensitivity of around 98% (6). For example, the study (5) found that the dynamic lesion process of viral pneumonia (from ground-glass opacity in the early stage to pulmonary consolidation in the late stage) can be observed in CT scans, and its CT manifestations have been emphasized.

Bacterial and viral pathogens are the two leading causes of pneumonia, but require very different forms of management (7). Bacterial pneumonia requires urgent referral for immediate antibiotic treatment, while viral pneumonia is treated with supportive care. Therefore, accurate classification of different types of pneumonia is imperative for timely diagnosis and treatment. However, the imaging features of viral and bacterial infections are not often compared, and the only imaging feature that was significantly different between the viral and bacterial lung infection was the frequency of diffuse airspace disease (8). In the case of a typical viral pneumonia, in clinical practice, it is difficult to accurately differentiate viral pneumonia from bacterial pneumonia. See **Figure 1** as an example. In clinical practice, especially in primary medical institutions, the consistency of imaging diagnosis of pneumonia pathogens is poor (9–11). Moreover, it is time consuming for radiologists to read CT volumes that contain hundreds of 2D slices. As such, it is of great practical significance to quickly and accurately identify pathogens to guide individualized anti-infectious treatment and minimize and delay the occurrence of drug resistance.

As an emerging technology in medical image analysis, artificial intelligence (AI) has been widely employed for lesion segmentation, and for clinical assessment and diagnosis of lung-related diseases *via* radiological imaging (12, 13). Recently, many novel AI techniques for viral pneumonia have been presented (1). For instance, Ouyang et al. (14) proposed a dual-sampling attention network for the differential diagnosis of COVID-19 from Community Acquired Pneumonia (CAP) (14, 15), and Fan et al. (16) introduced an automatic COVID-19 lung infection lesion segmentation method using a deep network. These works

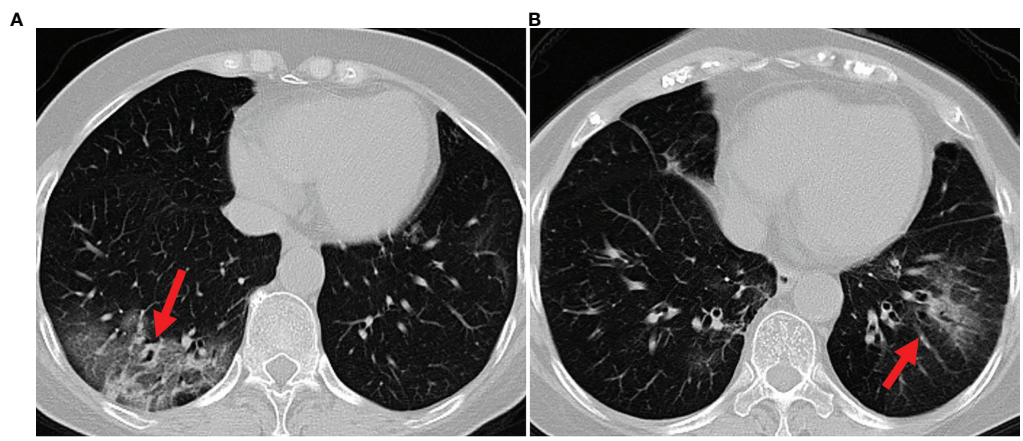


FIGURE 1 | Example axial CT slices of viral pneumonia **(A)**, bacterial pneumonia **(B)**. Accurate classification of different types of pneumonia is imperative for timely diagnosis and treatment. However, viral pneumonia and bacterial pneumonia display similar appearances in a CT image, which makes it difficult to accurately differentiate a patient with viral pneumonia from a case of bacterial pneumonia.

are useful for detecting and controlling of the spread of COVID-19. However, there are very few studies on differentiating COVID-19 from other etiological pneumonias, despite success in using deep learning (DL) approaches to discriminate bacterial and viral pneumonias in pediatric chest radiographs (17, 18).[1]

Further, most existing works make use of 2D CT slices, and the lack of continuity information makes it impossible to capture the true spatial distribution of the lesion in the lungs. To this end, some recent studies have attempted to use entire 3D volumes to train a 3D classification or segmentation model directly (14, 19) , achieving a slightly better performance than the approaches based on 2D slices. However, these 3D volume based approaches greatly increase the computational load, and require much more powerful and expensive hardware configurations. Additionally, 3D volumes may contain large portions of redundant information, which leads to great difficulty in accurately identifying small lesions. The imaging appearance of viral and bacterial lung infection has considerable similarity, and that, in any individual case, the viral pneumonia cannot reliably be distinguished from bacterial infections (8, 20). For example, viral pneumonia and bacterial pneumonia have some image features in common, such as ground glass opacities and interstitial changes in the peripheral zone of lungs, and accompanied by partial consolidation. The only imaging feature that was significantly different between the viral and bacterial lung infection was the frequency of diffuse airspace disease (8). Precise characterization of the spatial morphology of the infected regional lesions is essential to distinguish the two infection types by CT imaging.

In this paper, we treat the spatially continuous 2D CT slices as a time sequence and proposed a unified sequence-based pneumonia classification network (SLP-Net) for differentiating viral pneumonia (VP) from bacterial pneumonia (BP) and normal control cases. Our network comprises a CNN encoder and the ConvLSTM module, and sequence attention maps are used for auxiliary classification. As stated above, the precise characterization of the lesion is the key to distinguish the different pneumonia types. The combination of these components ensures the model pay more attention to spatial morphology of the lesion during the decision making. Specifically, the encoder with multi-scale receptive fields is first used to extract local representations of the sequence. Then we apply the ConvLSTM to acquire spatial information of these sequence features, modeling the distribution of the lesion. To optimize the SLP-Net, we introduce a novel adaptive-weighted cross-entropy (ACE) loss, with a view to ensuring that as many valid features from the previous images as possible are encoded into the subsequent CT image. Given the fact that the final diagnosis conclusion needs to be made for each patient, case-based prediction rather than a slice- or sequence-based prediction is more valuable. To obtain case-based prediction, in addition to the classification result of the sequence, we also use sequence attention maps to aid the case-level classification, aiming to enhance the confidence of the results. We collect a dataset of 258 chest CT volumes (153 VP, 42 BP, and 63 normal control cases). The experimental results show that the proposed SLP-Net achieves an accurate classification performance of viral pneumonia, bacterial pneumonia, and normal control, which could benefit the large-scale screening and control of viral pneumonia, and also enable efficient treatment for different types of pneumonia.

We organize the remainder of this paper as follows. In Section 2, the existing methods of AI-enpowered viral pneumonia analysis are briefly reviewed. In Section 3 we give detailed descriptions of collected datasets. Section 4 introduces the proposed SLP-Net. In Section 5, we present the experimental results and discuss the effectiveness, robustness, and efficiency of the SLP-Net. Section 6 concludes the paper and indicates directions for future work.

2 RELATED WORK

AI-based medical image analysis plays an essential role in the global fight against COVID-19, and a considerable number of approaches have been proposed in the past five months. This body of work on COVID-19 has focused primarily on two problems: lesion segmentation (16, 21, 22), and automated screening (23–31). For example (16), recently introduced a parallel partial decoder to aggregate high-level features, using an implicit reverse attention and explicit edge-attention to model boundaries and enhance representations so as to identify infected regions from 2D chest CT slices. To alleviate the shortage of labeled data, a semi-supervised segmentation framework based on a randomly selected propagation strategy was applied by (21). They proposed a relational approach, in which a non-local neural network module was introduced to efficiently learn both visual and geometric relationships among all convolutional features.

However, automated viral pneumonia (e.g., COVID-19) screening has attracted even more attention. For instance (32), introduced a COVID-19 detection method with multi-task DL approaches, using an inception residual recurrent convolutional neural network (CNN) with transfer learning. Their detection model achieved 84.67% accuracy from X-ray images (33). proposed a deep features fusion and ranking technique to detect COVID-19 in its early phase. They employed a pre-trained CNN structure to obtain a set of features, which were subsequently fused and evaluated with a support vector machine (SVM) classifier. In the classification task of COVID-19 and no COVID-19, their proposed method obtained 98.27% accuracy on their own dataset (34). applied a modified residual network, called DeepPneumonia, based on ResNet50 for slice-level classification, and could discriminate the COVID-19 patients from the bacteria pneumonia patients with an AUC of 0.95 (23). built multiple deep convolutional neural models for classifying chest X-ray images into normal and COVID-19 cases, which obtained 96.1% accuracy (35). proposed a unified latent representation to explore multiple features describing CT images from different views, a method that can completely encode information from different features aspects and is endowed with a promising class structure for separability. Performance in diagnosis for COVID-19 and community-

[1] https://github.com/HzFu/COVID19.

acquired pneumonia (CAP) is 95.5% in terms of accuarcy. An infection size-aware random forest method was introduced by (15) for the differentiation of COVID-19 from CAP, in which patients were automatically categorized into groups with different extensions of infected lesion sizes, followed by generation of random forests with each group for classification. The method achieved an accuracy of 89.4% in discriminating COVID-19 from CAP.

However, all of the above mentioned works are based on 2D images, the spatial correlation between consecutive CT scans is neglected by most slice-based methods, despite this being essential for the screening of lung diseases. A variety of volume-based methods have been proposed in an attempt to address this deficiency (19). proposed an attention-based deep 3D multi-instance learning method to screen COVID-19 from 3D chest CT sacns, using a weakly supervised learning framework that incorporates an attention mechanism into deep multi-instance learning, achieving an accuracy of 97.9% (14). proposed a 3D CNN to diagnose COVID-19 from CAP, in which a novel online attention module is combined with a dual-sampling strategy. The online attention module focuses on the infected regions when making diagnostic decisions. The dual-sampling strategy mitigates the imbalanced distribution in the sizes of infected regions between COVID-19 and CAP. Their method was evaluated in private dataset and achieved an accuracy of 87.5%. These 3D volume based approaches greatly increase the computational load, and require much more powerful and expensive hardware configurations. Additionally, 3D volumes may contain large portions of redundant information, which leads to great difficulty in accurately identifying small lesions. Overall, 2D slice based methods cannot take advantage of spatial continuity information and 3D volume based methods require much more expensive hardware configurations. In order to take advantage of the complementary information of 2D slices and 3D volumes, we treat the spatially continuous 2D CT slices as a time sequence, and divide the volume into multiple different temporal sequences of consecutive slices as the input.

3 MATERIALS AND METHODS

3.1 Materials

A total of 258 subjects were enrolled into this study, with 258 CT volumes, corresponding to 43,421 slices. Of the 258 subjects, 42 patients were confirmed positive for BP by clinical diagnosis (age: 59.5 ± 27.2; male/female: 36/6), 153 patients were positive for VP, confirmed by RT-PCR (age: 52.3 ± 12.7; male/female: 68/85), and 63 were control subjects (age: 35.8 ± 11.7; male/female: 33/30). The CT volumes of normal and VP patients were captured between January 29, 2020 and February 18, 2020, and the BP data was collected between January 2, 2019 and February 19, 2020. There is no statistically significant difference between the ages of the VP and BP subjects (P = >0.05), but both groups are significantly older than patients in the normal group (P < 0.001). CT examinations of all the enrolled patients were performed on a ScintCare CT16 (Minfound Inc, China) with standard chest imaging protocols. All the patients underwent CT scans during the end-inspiration without the administration of contrast material. Related parameters for chest CT scanning were listed as follows: field of view (FOV), 360 mm; tube voltage, 120 kV; tube current, 240 mA; helical mode; slice thickness, 5 mm; pitch, 1.5; collimation 16 × 1.2 mm; gantry rotation speed, 0.5 s/r; matrix, 512 × 512; software version, syngo CT 2014A; mediastinal window: window width of 350 HU, with a window level of 40 HU; and lung window: window width of 1,300 HU, with a window level of −500 HU.

CT volumes were retrospectively collected according to the history of laboratory investigations (e.g., sputum culture and reverse transcription-polymerase chain reaction), which we can generate the case-level labels. Meanwhile, professional radiologists (from the Hwa Mei Hospital, University of Chinese Academy of Sciences, Ningbo, China.) picked out the slice containing the infected region in each volume for the subsequent automatic generation of sequence labels: each volume was be divided into overlapping sequences containing n slices, with k overlapping slices between two sequences. If a sequence from VP volume contains the infected slice(s), the label for that sequence is 1; if it is from BP volume and contains infected slice(s), the label is 2; the label from normal volume is 0. It is worth noting that k and n are both hyperparameters, and in *Sensitivities to Hyperparameters* we discuss how to choose the values of these, as well as their impact on the classification results.

The sequences generated from the volume of VP and BP were not used for training if they did not contain any slices with lesion regions. If all normal slices from patients are used for training instead of excluding them, it will increase the proportion of normal control samples in the training set.The data imbalance may cause the model to over-fit the normal samples and tend to predict the samples with the lesion as normal. To avoid this, we exclude the normal slice from patients. For training and evaluation of the proposed method, as shown in **Table 1**, we split 258 volumes into 168 volumes (95 VP, 30 BP, and 43 normal controls) for training and 90 (58 VP, 12 BP, and 20 normal controls) volumes for testing.

3.2 Proposed Method

The architecture of our SLP-Net is shown in **Figure 2**, which consists of two main components: sequence CNNs, and ConvLSTM. The sequence CNNs with multi-scale receptive fields are employed to extract more discriminative high-level features from the CT sequence, while the ConvLSTM captures the axial dimensional dynamics of features. In addition, a sequence attention map is utilized as an auxiliary means to integrate the output of the network and obtain prediction results with a higher level of confidence. Finally, an adaptive-weighted cross-entropy (ACE) loss is used to optimize the whole model.

3.2.1 Sequence CNN With Multi-Scale Receptive Fields

A typical CNN model consists of a stack of convolution layers, interleaved with non-linear downsampling operations (e.g., max pooling) and point-wise nonlinearities (e.g., ReLU). The residual

TABLE 1 | Characteristics of training and testing CT dataset for identifying viral pneumonia (VP) from bacterial pneumonia (BP) and normal controls.

Cohort	VP		BP		Normal controls		Total	
	Volumes	Slices	Volumes	Slices	Volumes	Slices	Volumes	Slices
Training set	95	15,931	30	5,068	43	7,310	168	28,309
Testing set	58	7,832	12	3,107	20	4,173	90	15,112
Total	153	23,763	42	8,175	63	11,483	258	43,421

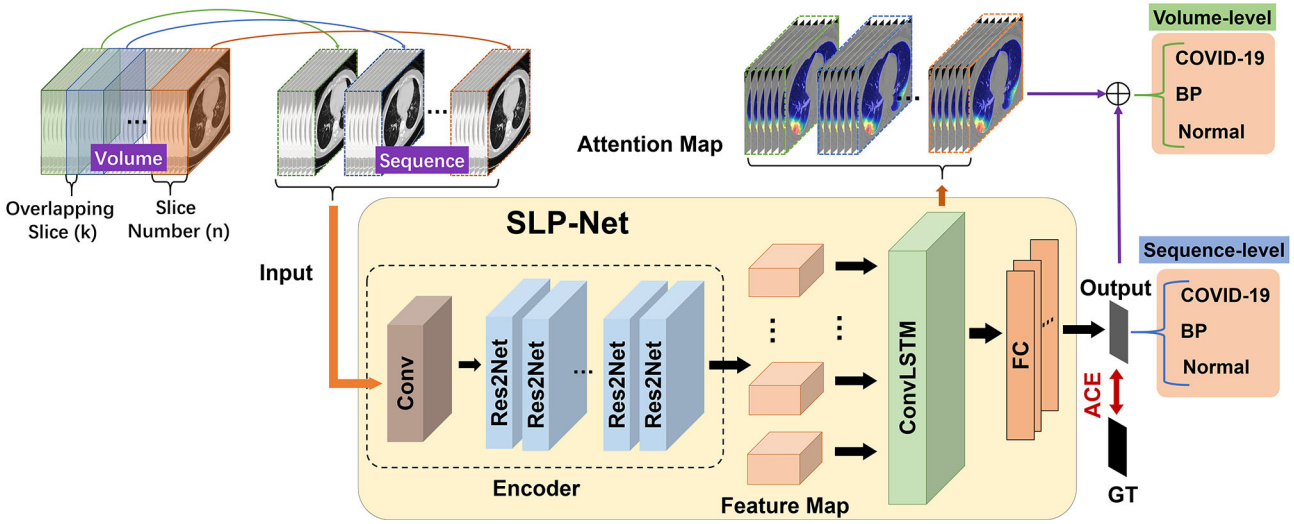


FIGURE 2 | Flowchart of our whole system for differentiating between viral pneumonia and bacterial pneumonia in a chest CT volume. Each volume is divided into overlapping sequences containing *n* slices during the training phase, such that the overlapping slices between two sequences are *k*. When predicting each volume during the testing phase, in addition to using the model to obtain the classification results of the sequence, we also introduce sequence attention maps for auxiliary classification to enhance the confidence level of the results. GT, ground truth; ACE, adaptive-weighted cross-entropy loss; FC, fully connected layer.

shortcut used in ResNet can reduce the over-fitting of the model, so that the depth of the network can be greater and achieve better performance. Taking into consideration the problems of overfitting and parameter cost, we employed ResNet (36) as our encoder backbone. The first four feature-extracting blocks are retained, without the average-pooling layer and the fully-connected layers.

Since VP and BP reveal similar appearances in CT images, we aim to obtain more discriminative features by employing multi-scale information, in order to distinguish them more accurately. Unlike most existing methods (37–39) that improve multi-scale ability by utilizing features with different resolutions, we apply a recently proposed multi-scale receptive fields technique (40) to enhance representation ability at a more granular level. Specifically, we apply a modified bottleneck with multi-scale ability, the Res2Net module, to replace a group of 3 × 3 filters used in the original bottleneck block of ResNet. As shown in **Figure 3**, after the 1 × 1 convolution, feature maps are split into *s* feature map subsets, denoted by x_i. Then, apart from x_1, each x_i goes through a corresponding 3 × 3 convolutional operator, denoted by $K_i(\cdot)$, where y_i is the output of $K_i(\cdot)$. The output of $K_{i-1}(\cdot)$is added to x_i, then sent to the next group of filters $K_i(\cdot)$. Thus, y_i can be defined as follows:

$$y_i = \begin{cases} x_i & i = 1; \\ K_i(x_i) & i = 2; \\ K_i(x_i + y_{i-1}) & 2 < i \le s. \end{cases} \quad (1)$$

In order to better integrate the information from different scales, all outputs y_i, where $i \in \{1,2, \ldots, s\}$, are concatenated and passed through a 1×1 convolution. Such splitting and concatenation strategies can force the convolution to process features more efficiently. Note that each 3×3 convolution operation $K_i(\cdot)$receives information from all the feature splits $\{x_j, j \le i\}$. Each time x_j performs a 3×3 convolution, the size of the receptive field will increase. Due to the combinatorial effect, the output of the Res2Net block contains different combinations of receptive field sizes/scales.

3.2.2 ConvLSTM With ACE Loss

Although the conventional fully-connected LSTM (FC-LSTM) can handle sequences of any length and capture long-term dependencies (41), it contains too much redundancy for spatial data, which is a critical problem for image sequences. Inspired by video object detection (42), we apply ConvLSTM (43) to process the feature sequences from the encoder.

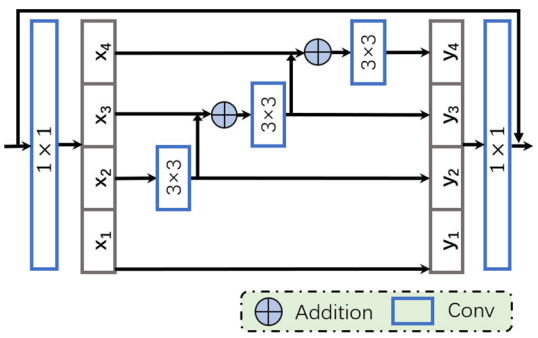


FIGURE 3 | A Res2Net module is utilized to extract more discriminative features.

As the convolutional counterpart of the FC-LSTM, the ConvLSTM introduces the convolution operation into the input-to-state and state-to-state transitions. The ConvLSTM can model axial dimensional dependencies while preserving spatial information. As with the FC-LSTM, the ConvLSTM unit (see **Figure 4**) includes an input gate i_t, a memory cell C_t, a forget gate f_t and an output gate o_t. The memory cell C_t, acting as an accumulator of the state information, is accessed, updated and cleared through self-parameterized controlling gates: i_t, o_t, and f_t. If the input gate is switched on, the new data is accumulated into the memory cell once an input arrives. Similarly, the past cell status C_{t-1} will be forgotten if the forget gate f_t is activated. The output gate o_t further controls whether the latest memory cell's value C_t will be transmitted to the final state H_t. With the above definitions, the ConvLSTM can be formulated as follows:

$$
\begin{aligned}
i_t &= \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci}\, C_{t-1} + b_i),\\
f_t &= \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf}\, C_{t-1} + b_f),\\
C_t &= f_t\, C_{t-1} + i_t\, \tanh\left(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c\right),\\
o_t &= \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co}\, C_t + b_o),\\
H_t &= o_t\, \tanh\left(C_t\right),
\end{aligned}
\tag{2}
$$

where '*' denotes the convolution operator, '∘' denotes the Hadamard product, and σ is the sigmoid activation function. X_t and $\mathcal{H}_t$ are the input and output of the ConvLSTM at time step t (t indicates the tth frame in a CT image sequence, and slices will be referred to as frames in the sequel.), and i_t, f_t, and o_t indicate the input, forget and output gates, respectively. b_i, b_f, and b_o are the bias of the input gate, forget gate, and output gate. A memory cell C_t stores the historical information. All the gates i, f, o, memory cell C, hidden state $\mathcal{H}$ and the learnable weights W are 3D tensors. Input sequences ? are fed into a ConvLSTM block, which captures the long and short-term memory of sequences and contains both axial dimensional information, for use in implicitly learning axial dimensional dynamics and efficiently fusing axial dimensional features.

We define L_t as the output of the ConvLSTM layer at time step t. The output of the ConvLSTM layer is fed to the fully-connected (FC) layers, which transform the features into a space that makes the output easier to classify. The outputs of the FC layers are defined as O_t at time step t. Ideally, the longer the image sequence, and the more classification information ConvLSTM processes, the higher the confidence of classification. From this perspective, it is sufficient to use the output of the final time step for classification without further processing. However, in practice, due to differences in the distribution of lesions on different slices, there may be some useful information that has not been accumulated in the memory

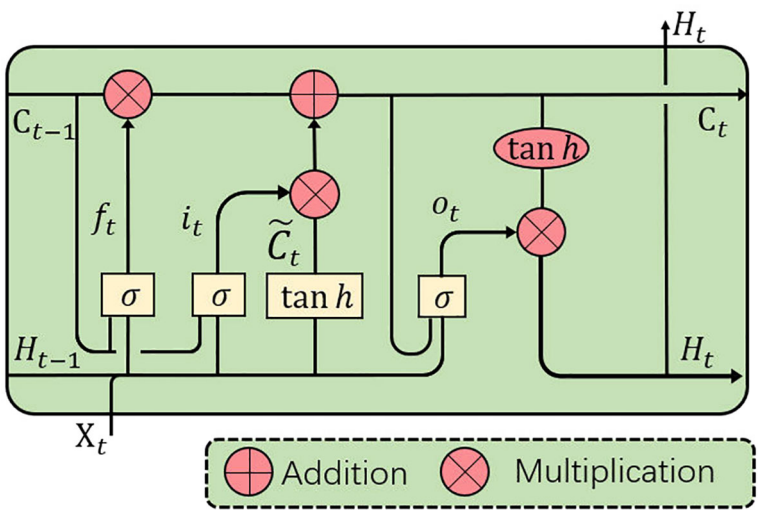


FIGURE 4 | ConvLSTM is utilized to implicitly learn axial dimensional dynamics and efficiently fuse axial dimensional features.

cell. In order to enchance the memory ablity of ConvLSTM for CT sequence at different slices and ensure that as much valid information from the previous slices as possible are encoded, we propose to use all the intermediate outputs of every time step as our feature for identification. A better ConvLSTM means that the longer the sequence it processes and the more comprehensive information it considers, the more confident it identifies the input. From this perspective, instead of minimizing the loss on the final time step, we define a new adaptive-weighted cross-entropy (ACE) loss to use all the intermediate outputs of every time step weighted by w_t:

$$\mathcal{L}_{ACE} = \frac{1}{n}\sum_{t=1}^{n}\sum_{p=1}^{P} - w_t[y_p \log (C_p(O_t))], \tag{3}$$

where C and p denote the classifier and classification label, respectively, and n denotes the number of images in a sequence. $C_p(O_t)$ indicates the classifier C, which correctly identifies the final output O at time step t, $y_p \in \{0, 1, 2\}$ are the label values; and P=3 denotes the total number of labels. Finally, w_t is the weight of each frame in a sequence. weight. We let each group of two weight items constitute the arithmetic sequence.

Since the importance of the information contained in different slices is different, it is not reasonable to use the equal weights. Due to the output of the final time step has taken into account all other previous slices , it contains the most information, and the further away from the last output, the less information it contains. Moreover, the number of slices in a sequence is a hyper-parameter, we adopt an adaptive weighting scheme. The output of the final time step should be assigned the maximum weight, and the farther away from the final time step, the smaller the weight. Specifically, we let each group of two weight items constitute the arithmetic sequence: the sum of which is 1. The first two items are taken as 0.01, namely, $w_1 = w_2 = 0.01$, and the subsequent weights can then be calculated according to the hyperparameter n and weight w_1. The ACE loss ensures that the features of the previous CT images in the sequence can be encoded into the later image.

3.2.3 Auxiliary Diagnosis With Attention Maps

Deciding which type (VP, BP, or normal) the entire volume belongs to based on the prediction results of the sequence is a critical step in auxiliary diagnosis. For higher confidence, in addition to using the model to obtain the classification result of the sequence, we also utilized the Grad-CAM (44) technology to generate attention maps of the sequence to assist the prediction. Grad-CAM is a method for producing visual interpretations for CNNs in the form of class-specific saliency maps. A saliency map, L_t^c, is produced for each image input based on the activation from k filters, A_{ij}^k, at the final convolutional layer. To make the method applicable to image sequences, the activations for all timesteps t in the sequence are considered (Eq. R1).

$$L_{ijt}^c = \sum_a w_{kt}^c A_{ijt}^k;\ w_{kt} = \frac{1}{Z}\sum_{ij}\frac{\partial F^c}{\partial A_{ijt}^k} \tag{R1}$$

where Z is a normalizing constant and F^c is the network output for the class c. i, j are pixel location of filter A^k. In the visualization examples shown in **Figure 5**, stronger class activation map (CAM) areas are indicated with lighter colors.

During the prediction phase, we can obtain the classification results of sequences belonging to a volume. In addition, we apply Grad-CAM to generate the response heat map of each sequence. A volume containing m sequences will be classified as viral pneumonia (VP) sample if it meets both of following criteria: (a) More sequences are classified as viral pneumonia (VP) then bacterial pneumonia (BP) in this volume; (b) There are two adjacent sequences with the category of viral pneumonia whose activation regions have an intersecting area of more than 50%. If the second criterion is not satisfied for VP, the bacterial type sequences are checked if there are two adjacent sequences with an intersecting area of more than 50% of the activation region, and if so the sample is classified as bacterial type, otherwise it is classified as normal. Notably, if there are the same number of VP and BP sequences, the one with the greater average sequence probability value is treated as the dominant category. The possibility output of the network dominates the classification on the volume level, and Attention Map is used as an auxiliary during the decision making.

3.3 Evaluation Metrics

We employ the commonly used metrics for multi-class classification to measure performance: e.g., weighted sensitivity (Sen, also known as recall), specificity (Spe), accuracy (Acc), and balanced accuracy (B-Acc, a.k.a. balanced classification rate). In order to reflect the tradeoff between sensitivity and specificity, and evaluate the quality of our classification results more reliably, a kappa analysis and F-measure ($F1$ score) are also provided following (45). These two measures are more robust than other percentage agreement measures, as they take into account the possibility of the agreement occurring by chance. The weighted sensitivity (Sen), specificity (Spe), accuracy (Acc), and balanced accuracy (B-Acc) are defined as:

$$Spe = \sum_{i=1}^{P} w_i \frac{TN_i}{TN_i + FP_i},\ Sen = \sum_{i=1}^{P} w_i \frac{TP_i}{TP_i + FN_i},$$

$$B - Acc = \frac{(Sen + Spe)}{2},\ Acc = \sum_{i=1}^{P} w_i \frac{TP_i + TN_i}{TP_i + FN_i + FP_i + TN_i},$$

where TP_i indicates the number of true positives, TN_i—the number of true negatives, FP_i—the number of false positives, and FN_i—the number of false negatives for the i–*th* classification label; and w_i represents the percentage of images whose ground truth labels are i. The kappa values and F-measure ($F1$ score, a.k.a. Dice score) are defined as follows:

$$p_o = \sum_{i=1}^{P} w_i \frac{TP_i}{n}, p_e = \sum_{i=1}^{P} \frac{a_i * b_i}{n * n},\ Pre = \sum_{i=1}^{P} w_i \frac{TP_i}{TP_i + FP_i},$$

$$Kappa = \frac{p_o - p_e}{1 - p_e},\ F1 = 2 \cdot \frac{Pre \cdot Sen}{Pre + Sen},$$

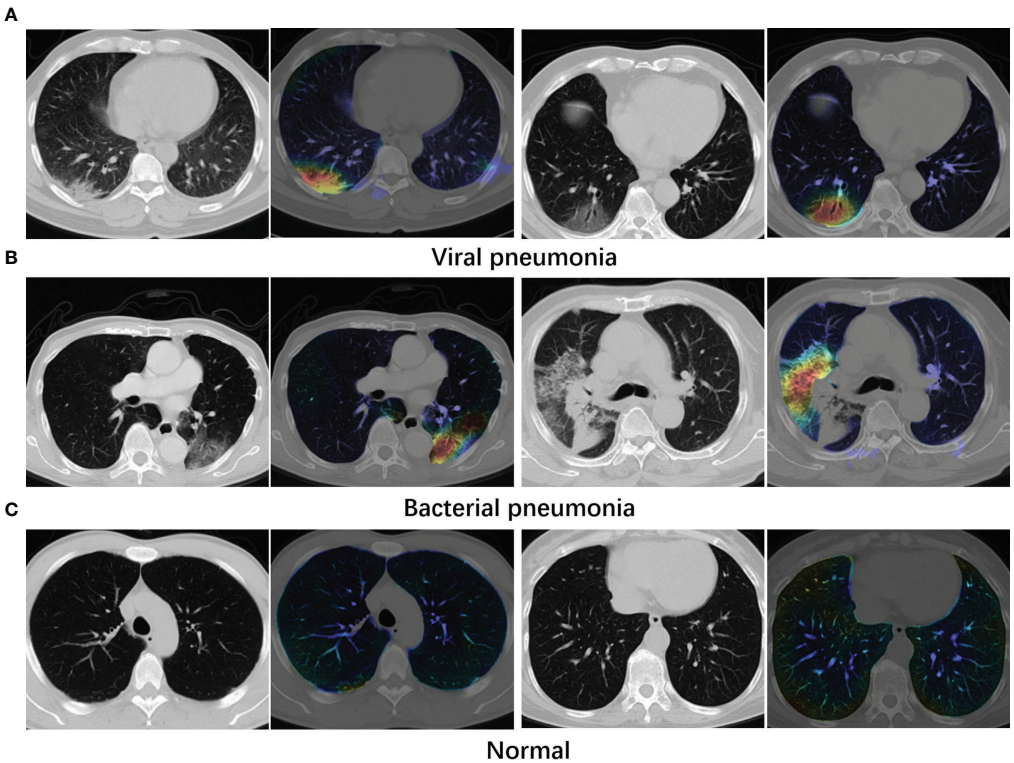


FIGURE 5 | Examples of attention maps obtained with Grad-CAM. **(A)** Viral pneumonia cases. **(B)** Bacterial pneumonia cases. **(C)** Normal cases. Lighter colors indicate the stronger response regions. From the maps, the infected regions receive greater attention.

where a_i denotes the true sample number of each class, b_i denotes the predicted sample number of each class, n denotes the total sample number, and P denotes the number of classes. Note that *kappa* values between 0.81 to 1.00 indicate almost perfect agreement, values between 0.61 and 0.80 exhibit substantial agreement, values of 0.41–0.60 exhibit moderate agreement and values less than 0.40 exhibit poor to fair agreement. The $F1$ score reaches its best value at 1 and worst at 0. We also present the ROC curves and the area under ROC curve (AUC) for VP against BP.

3.4 Implementation Details

The proposed method was implemented in the publicly available Pytorch library. The combination of CNN and ConvLSTM makes the model more complex. To accelerate convergence, we first trained a CNN classification network with a labeled 2D slice. After removing the FC layer, the encoder is used as the initialization parameter of SLP-Net. During the training phase of SLP-Net, CNN and ConvLSTM are jointly trained in an end-to-end manner using Adam optimizer. In practice, we found that CNN pre-training does speed up the convergence of the model, but has no effect on the final classification performance. The learning rate was gradually decreasing starting from 0.0001, and the momentum was set to 0.9. In addition, online data enhancement was employed to enlarge the training sequence data. The same data enhancement was used for all images in a sequence: we implemented data augmentation in a random way, including brightness, color, contrast, and sharpness transformation from 90 to 110%. We set a random seed from 1 to 4 for the enhancement.

4 RESULTS

4.1 Classification Performances

To compare the classification performance, we evaluated the detection ability of the model at both sequence and volume levels. All the existing pneumonia detection methods are accomplished using 2D CT slices or 3D volumes. To further verify whether the features containing both axial dimensional and spatial information captured by our model could benefit detection performance, we compared the proposed method to other classic classification models using 2D slices: AlexNet (46), VGG19 (47), InceptionV3 (48), ResNet34 (36), and Xception (49). Due to the lack of sufficient training data and the GPU memory constraint, we cannot apply 3D CNNs on complete CT volumes. In order to compare the proposed SLP-Net with other 3D deep learning architectures, we apply CT sequence data, which can be considered as 3D data, to train 3D models, including C3D (50), I3D (51), and S3D (52). We report the detection results for slice/sequence-level and case-level in **Table 2**. We applied a similar strategy to that in *Auxiliary*

TABLE 2 | Classification results for VP, BP and normal controls by different methods.

Method	Slice/Sequence-Level					Case-level				
	Kappa	F1	B-Acc	Sen	Spe	Kappa	F1	B-Acc	Sen	Spe
AlexNet	0.5207	0.5680	0.6872	0.6375	0.7370	0.6889	0.7381	0.8207	0.8274	0.8140
VGG19	0.6258	0.6574	0.7502	0.7152	0.7853	0.7709	0.8000	0.8571	0.8690	0.8601
ResNet34	0.6767	0.7100	0.7948	0.7783	0.8112	0.8489	0.8598	0.9045	0.9048	0.9043
InceptionV3	0.5177	0.6107	0.7156	0.6978	0.7333	0.7692	0.7976	0.8607	0.8631	0.8582
Xception	0.6802	0.6776	0.7688	0.7252	0.8124	0.8500	0.8631	0.9048	0.9107	0.9061
C3D	0.7382	0.7442	0.8232	0.8013	0.8450	0.8616	0.8729	0.9158	0.9183	0.9132
I3D	0.7437	0.7513	0.8403	0.8302	0.8132	0.8481	0.8952	0.9229	0.9167	0.9290
S3D	0.7525	0.7332	0.8106	0.7696	0.8517	0.8696	0.8796	0.9297	0.9133	0.9157
SLP-Net	**0.8280**	**0.8123**	**0.8665**	**0.8397**	**0.8934**	**0.9263**	**0.9291**	**0.9523**	**0.9524**	**0.9521**

The best performance of all the methods is highlighted in bold.

Diagnosis With Attention Maps section to determine the prediction result of a volume when using the 2D models. First, Grad-CAM was used to generate the activated maps of 2D slices, and binary activated maps can be obtained through thresholding. If five consecutive slices in a volume were predicted as indicative of viral pneumonia, and the intersection area of their activated area exceeds 50% of the union area, the volume was considered to be indicative of viral pneumonia.

As can be observed, the 3D networks achieve higher performances than the 2D networks, which confirms the importance of the combination of axial dimensional and spatial information for accurate detection results. At a slice/sequence level, our SLP-Net outperformed other methods in terms of *kappa*, *F*1 and *Sen* by a large margin, as well as achieving the best performance at a volume level. In addition, **Figure 6** shows the confusion matrices of VGG-19, Resnet34, Xception, C3D,

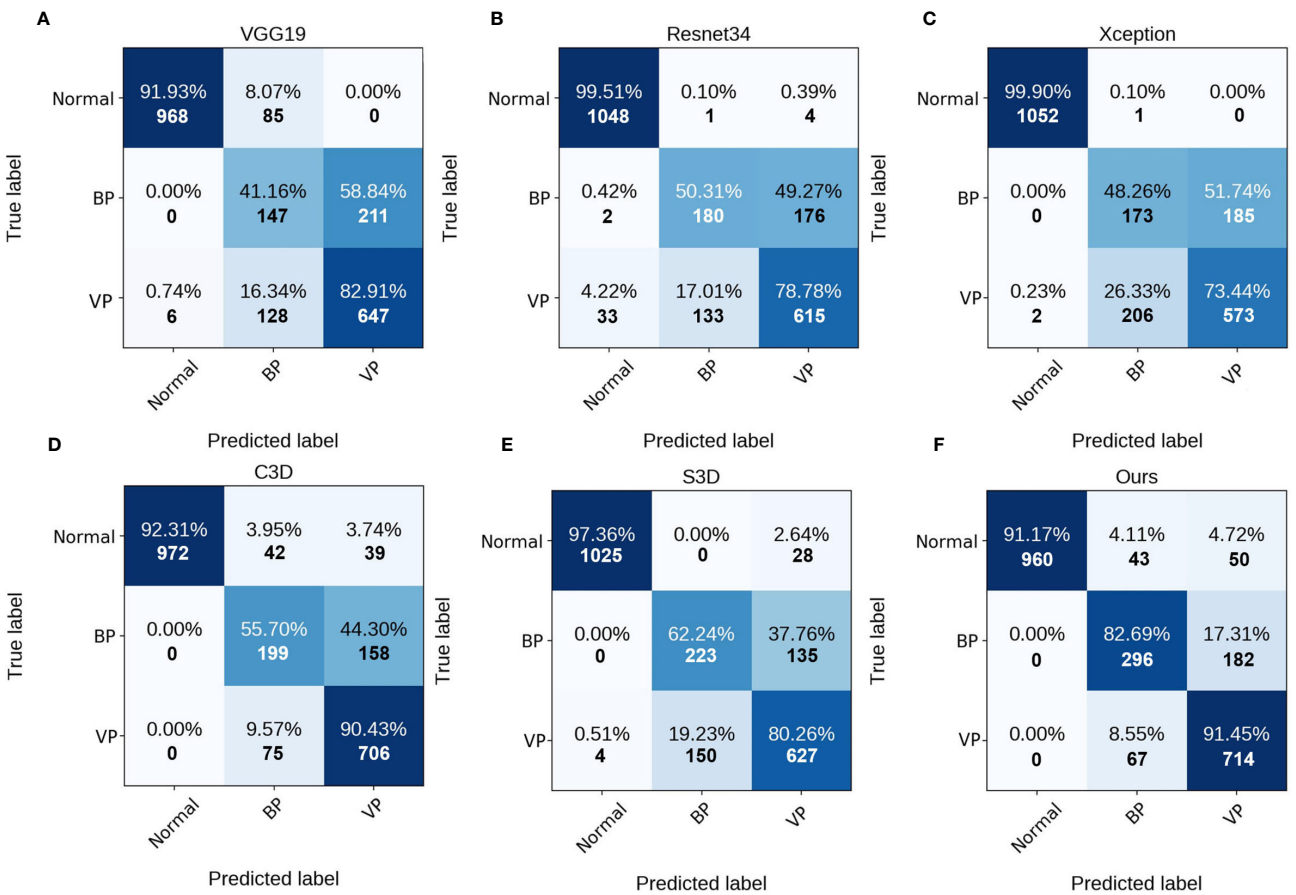


FIGURE 6 | Confusion matrices of the different methods at slice/sequence level. **(A–F)** are the results of VGG19, Resnet34, Xception, C3D, S3D and ours, respectively. The numbers in the confusion matrices denote the percentage (above) and number (below) of the predicted class.

S3D, and our method over the dataset. These results further indicate the superiority of the performance of our approach. As stated in the Introduction section, the difficulty of pneumonia diagnosis is the differentiation between BP and VP. Accordingly, we conducted experiments on the dataset contained VP and BP samples only. Results are shown in **Table 3** and **Figure 7**. We may observe that the proposed method again produces the best performance compared to the other methods. **Table 3** gathers all the performances of these models. The results show that the 3D-based method is generally better than the 2D-based method, mainly because the 3D input provides richer spatial information, which allows the model to learn and extract the subtle differences in the spatial distribution of different diseases, which is especially important for difficult samples with similar lesion appearance. In this regard, our proposed method not only utilizes the 3D information, but also explicitly focuses on the lesion area in the decision-making process through attention map, which makes the classification results more reliable. This idea can be applied to many different medical image-based classification tasks, since the similarity of lesion appearance is a problem in many scenarios.

Table 4 summarizes space and time cost of different methods. For fair comparison of inference time, we test all these models with PyTorch. Our SLP-Net had the best time efficiency and achieved smallest model size because it didn't use 3D convolution operations.

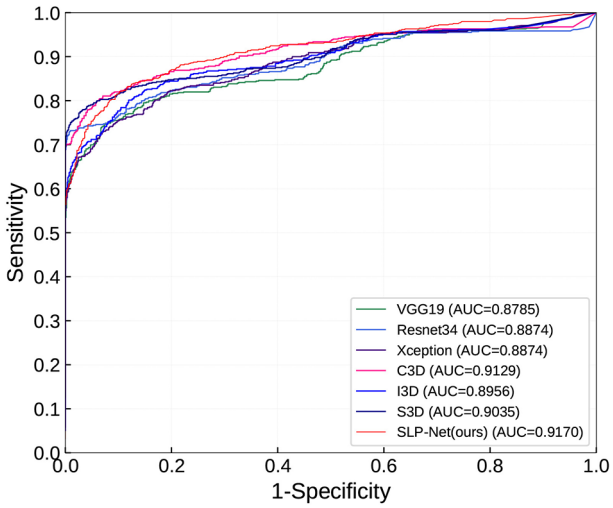


FIGURE 7 | ROC curves in classifying viral pneumonia (VP) and bacterial pneumonia (BP) of the compared models and the proposed method.

5 ANALYSIS AND DISCUSSION

5.1 Sensitivities to Hyperparameters

In **Table 5** we investigate the sequence settings, i.e., n and k, denote the number of slices per sequence and the number of overlapping slices between two sequences, respectively. By default we set $n = 10$ and $k = 5$. As can be observed, the performance was greatly affected by the value assigned to n. When n is set very small ($n = 5$), the *kappa* and $F1$ drop by the considerable margin of 3%, demonstrating that the more axial dimensional information ConvLSTM encodes, the more the model benefits. However, when n is greater than 15, the model performance will decline. This may be due to the fact that as the slice number increases, there will be less training data, resulting in the model not being fully trained. **Table 5** also shows that our result is impacted just marginally when k is within a scale of 5-7. The performance of the model mainly depends on the abundance of the information contained in the sequence, that is, the more information contained in the sequence, the better the classification performance. Compared to n = 10, the sequence provide less timing information when n = 5, so the classification performance will decrease. If n is too large (e.g., n = 20), the performance will decrease due to fewer training samples.

5.2 Ablation Study

Our SLP-Net employs three main components to form the classification framework: a sequence CNNs with multi-scale receptive fields, a ConvLSTM module, and a carefully designed ACE loss. In this subsection, we analyze and discuss the network under different scenarios to validate the performance of each key component of our model, and the results of different combinations of these modules are reported in **Figure 8**.

5.2.1 Effectiveness of ConvLSTM

To explore the contribution of the ConvLSTM, we use a ResNet50 pretrained on ImageNet as the backbone. As shown in **Figure 8**, a backbone + ConvLSTM + Res2Net method clearly

TABLE 3 | Comparison of different methods in classifying viral pneumonia (VP) and bacterial pneumonia (BP), at a slice/sequence level.

Method	Acc	Sen	Spe	AUC (*p*-value)
AlexNet	0.7327	0.8182	0.6650	0.8700 (p <0.001)
VGG19	0.7776	0.8197	0.7442	0.8785 (p <0.001)
ResNet34	0.7939	0.8305	0.7649	0.8874 (p <0.001)
InceptionV3	0.7429	0.8028	0.6955	0.8697 (p <0.001)
Xception	0.8170	0.7704	0.8438	0.8874 (p <0.001)
C3D	0.8218	**0.869**	0.7844	0.9129 (p <0.05)
I3D	0.8320	0.8274	0.8356	0.8956 (p <0.001)
S3D	0.8361	0.8413	0.8319	0.9035 (p <0.01)
SLP-Net	**0.8476**	0.8459	**0.8490**	**0.9170**

P-value is calculated by Delong's test.
The best performance of all the methods is highlighted in bold.

TABLE 4 | Model size and inference time of different methods.

	C3D	I3D	S3D	SLP-Net
Model size (MB)	39.2	48.7	42.3	34.4
Time (ms)	41.4	59.0	47.1	39.5

TABLE 5 | Effect of different settings of hyperparameter *n* and *k* on the results.

Method	kappa	F1	Acc	B-Acc	Sen	Spe
n = 5, k = 3	0.7376	0.7469	0.8546	0.8272	0.8096	0.8449
n = 10, k = 3	0.7652	0.7701	0.8684	0.8453	0.8307	0.8599
n = 10, k = 5	0.8241	0.8091	0.9012	0.8641	0.8471	0.8911
n = 10, k = 7	0.8280	0.8123	0.9034	0.8665	0.8397	0.8934
n = 15, k = 5	0.7453	0.7557	0.8577	0.8362	0.8231	0.8493
n = 20, k = 5	0.7329	0.7068	0.8577	0.7930	0.7450	0.8410

Here, n *and* k *denote the number of slices in the sequence and the number of overlapping slices between two sequences, respectively.*

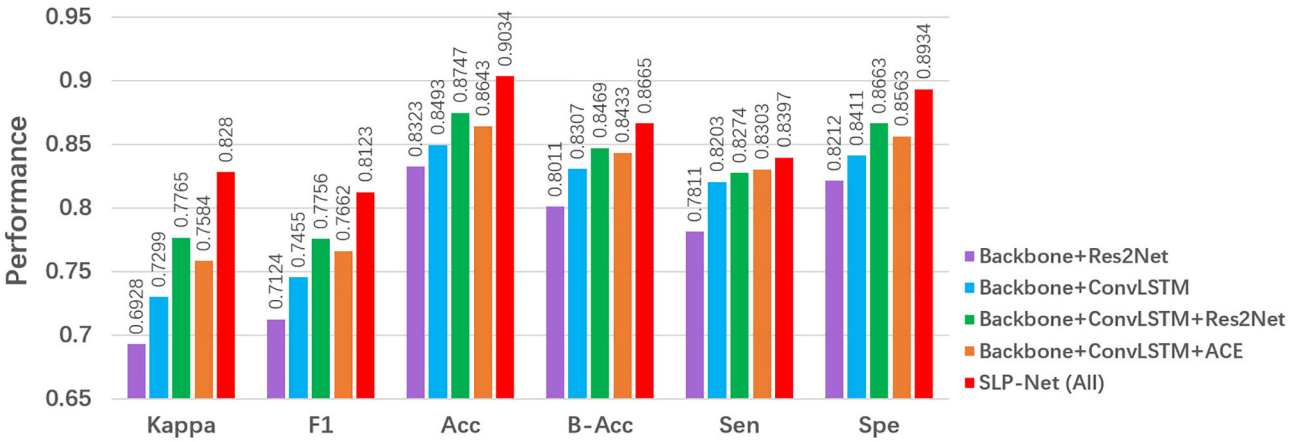


FIGURE 8 | Ablation studies of our SLP-Net.

TABLE 6 | Ablation study of Attention Map in classifying VP, BP, and Normal controls at the case-level.

Method	Kappa	F1	B-Acc	Sen	Spe
Without Attention Map	0.8731	0.9090	0.9167	0.9174	0.9160
With Attention Map	0.9263	0.9291	0.9523	0.9524	0.9521

outperformes the backbone + Res2Net, with improvement of about 6% in *F*1. This shows that the ConvLSTM is capable of extracting the axial dimensional and spatial information, thus memorizing the change in appearance that corresponds to axial dimensional information, and improving the performance in identifying and discriminating between VP and BP.

5.2.2 Effectiveness of the Res2Net module

We investigated the importance of the multi-scale sequence module, i.e., Res2Net. From **Figure 8**, we observe that a backbone + ConvLSTM + Res2Net model outperformed the backbone model in terms of major metrics, i.e., *kappa* and *F*1. This suggests that introducing the Res2Net module enables the encoder to capture more discriminative features to accurately differentiate VP from BP.

5.2.3 Effectiveness of ACE

Finally, we investigate the importance of the ACE loss. From the results in **Figure 8**, it may clearly be observed that the ACE Loss effectively improves the classification performance in our model. One possible reason is that, with the ACE loss, the ConvLSTM explores the axial dimensional dynamics of appearance features in CT sequences, and these features are further aggregated for classification purposes.

5.2.4 Effectiveness of Attention Maps

To investigate the contribution of the Attention Map, we added an additional experiment—case-level classification without Attention Map. Specifically, sequence-level classification results were first obtained using SLP-Net, and if there were VP or BP sequences in a volume, the type with more number is used as the category of the whole volume. If it does not contain VP and BP sequence, it is classified as normal. Notably, if there are the same number of VP and BP sequences, the one with the greater average sequence probability value is treated as the dominant category. **Table 6** shows the result, where SLP-Net with Attention Map as auxiliary achieves better performance than without Attention Map. This demonstrates that with the aid of attention map, the distribution of lesions can be considered simultaneously in the decision-making process, thus improving the performance of case-level classification.

6 DISCUSSION AND CONCLUSIONS

6.1 Limitations

Although our method achieves better results in the pneumonia classification task compared to other methods, this work still has some limitations. Firstly, we used the multiscale feature technique Res2Net in the feature extraction part, but did not further explore the hyperparameter settings in it, and although we believe that careful selection of hyperparameters may further improve the classification performance, no additional experiments were conducted in this work to compare the impact of different hyperparameters since this is not the focus of our work. Secondly, the model is not evaluated on an external dataset. To our knowledge, there are no publicly available 3D CT datasets for different types of pneumonia classification tasks, and it is difficult to collect compliant data from multiple centers due to various conditions. We intend to evaluated the performance of our model on external datasets in the future.

6.2 Conclusion

Hospitals are beginning to use CT imaging in the diagnosis of viral pneumonia, and it is vital to improve the sensitivity of diagnostic methods so as to reduce the incidence of false negatives. AI-empowered image acquisition workflows are effective, and may also aid in protecting clinicians from viral pneumonia (e.g., COVID-19) infection. Although several effective AI-based COVID-19 diagnosis or lesion segmentation methods have been introduced recently, automated differentiation of viral pneumonia from other types of pneumonia is still a challenging task. The motivation of this study was to employ AI techniques to alleviate the problem posed by the fact that even radiologists are hard pressed to distinguish VP from BP, as they share very similar presentations of infection lesion characteristics in CT images.

In this paper, we have proposed a novel viral pneumonia detection network, named SLP-Net. By contrast with previous 2D slice-based or 3D volume-based methods, we considered continuous CT images as time sequences. Our model first utilized the sequence CNNs with multi-scale receptive fields to extract a sequence of higher-level representations. The feature sequences were then fed into a ConvLSTM to capture axial dimensional features. Finally, in order to ensure that as many valid features from previous slice as possible are encoded into the later CT slices, a novel ACE loss was proposed to optimize the output of the SLP-Net. Furthermore, during the prediction phase, we used sequence attention maps for auxiliary classification to predict each volume, which can enhance the confidence level of the results. Overall, in order to accurately distinguish VP from BP and normal subjects, we used the sequence CNNs with multi-scale receptive fields to extract more differentiating features, and then applied a ConvLSTM to capture axial dimensional features of the CT sequence, thereby obtaining features containing both axial dimensional and spatial information. The superior evaluation performance achieved in the classification experiments demonstrate the ability of our model in the differential diagnosis of VP, BP and normal cases. Although we only evaluated our method on the CT dataset of pneumonia, it can be adapted to any other 3D medical image classification problems, such as lung cancer imaging analysis, and the identification of Alzheimer's disease. In future work we will further validate our models on even larger datasets, and seek its implementation in real clinical settings.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Hwa Mei Hospital, University of Chinese Academy of Sciences. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JH, JX, and RL were involved in data analysis and interpretation, and drafting and revising the manuscript. HH, YM, KY, RRL, YLZ, and JJZ were involved in data analysis and interpretation. JL, JFZ, and YTZ were involved in study conceptualization, supervision, revising the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported in part by the Zhejiang Provincial Natural Science Foundation of China (LZ19F010001, and LQ20F030002), in part by the Key Project of Ningbo Public Welfare Science and Technology (2021S107), and in part by the Youth Innovation Promotion Association CAS (2021298).

REFERENCES

1. Wang C, Horby PW, Hayden FG, Gao GF. A Novel Coronavirus Outbreak of Global Health Concern. *Lancet* (2020) 395:470–3. doi: 10.1016/S0140-6736(20)30185-9
2. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical Features of Patients Infected With 2019 Novel Coronavirus in Wuhan, China. *Lancet* (2020) 395:497–506. doi: 10.1016/S0140-6736(20)30183-5
3. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology* (2020) 296:E32–40. doi: 10.1148/radiol.2020200642
4. Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology* (2020) 296:E115–7. doi: 10.1148/radiol.2020200432
5. Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of Chest Ct for Covid-19: Comparison to Rt-Pcr. *Radiology* (2020) 296:E115–7. doi: 10.1148/radiol.2020200432
6. Pan F, Ye T, Sun P, Gui S, Liang B, Li L, et al. Time Course of Lung Changes at Chest CT During Recovery From Coronavirus Disease 2019 (COVID-19). *Radiology* (2020) 295:715–21. doi: 10.1148/radiol.2020200370
7. McLuckie A. *Respiratory Disease and its Management*. London, UK: Springer Science & Business Media (2009).
8. Miller WT, Mickus TJ, Barbosa JE, Mullin C, Van VM, Shiley K, et al. CT of Viral Lower Respiratory Tract Infections in Adults: Comparison Among Viral Organisms and Between Viral and Bacterial Infections. *Am J Roentgenology* (2011) 197:1088–95. doi: 10.2214/AJR.11.6501
9. Jae YP, Rosemary F, Richard S, Neil S, Nicholas J. *The Accuracy of Chest Ct in the Diagnosis of Covid-19: An Umbrella Review. Website* (2021). Available at: https://www.cebm.net/covid-19/the-accuracy-of-chest-ct-in-the-diagnosis-of-covid-19-an-umbrella-review/.
10. Nambu A, Ozawa K, Kobayashi N, Tago M. Imaging of Community-Acquired Pneumonia: Roles of Imaging Examinations, Imaging Diagnosis of Specific Pathogens and Discrimination From Noninfectious Diseases. *World J Radiol* (2014) 6:779. doi: 10.4329/wjr.v6.i10.779
11. Wootton D, Feldman C. The Diagnosis of Pneumonia Requires a Chest Radiograph (X-Ray)—Yes, No or Sometimes? *Pneumonia* (2014) 5:1–7. doi: 10.15172/pneu.2014.5/464
12. Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, et al. Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation, and Diagnosis for Covid-19. *IEEE Rev Biomed Eng* (2020) 14:4–15. doi: 10.1109/RBME.2020.2987975
13. Dong D, Tang Z, Wang S, Hui H, Gong L, Lu Y, et al. The Role of Imaging in the Detection and Management of Covid-19: A Review. *IEEE Rev Biomed Eng* (2020) 14:16–29. doi: 10.1109/RBME.2020.2990959
14. Ouyang X, Huo J, Xia L, Shan F, Liu J, Mo Z, et al. Dual-Sampling Attention Network for Diagnosis of Covid-19 From Community Acquired Pneumonia. *IEEE Trans Med Imaging* (2020) 39:2595–605. doi: 10.1109/TMI.2020.2995508
15. Shi F, Xia L, Shan F, Song B, Wu D, Wei Y, et al. Large-Scale Screening to Distinguish Between Covid-19 and Community-Acquired Pneumonia Using Infection Size-Aware Classification. *Phys Med Biol* (2021) 66:065031. doi: 10.1088/1361-6560/abe838
16. Fan D, Zhou T, Ji G, Zhou Y, Chen G, Fu H, et al. Inf-Net: Automatic Covid-19 Lung Infection Segmentation From Ct Images. *IEEE Trans Med Imaging* (2020) 39:2626–37. doi: 10.1109/TMI.2020.2996645
17. Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *cell* (2018) 172:1122–31. doi: 10.1016/j.cell.2018.02.010
18. Rajaraman S, Candemir S, Kim I, Thoma G, Antani S. Visualization and Interpretation of Convolutional Neural Network Predictions in Detecting Pneumonia in Pediatric Chest Radiographs. *Appl Sci* (2018) 8:1715. doi: 10.3390/app8101715
19. Han Z, Wei B, Hong Y, Li T, Cong J, Zhu X, et al. Accurate Screening of Covid-19 Using Attention-Based Deep 3d Multiple Instance Learning. *IEEE Trans Med Imaging* (2020) 39:2584–94. doi: 10.1109/TMI.2020.2996256
20. Kohno N, Ikezoe J, Johkoh T, Takeuchi N, Tomiyama N, Kido S, et al. Focal Organizing Pneumonia: CT Appearance. *Radiology* (1993) 189:119–23. doi: 10.1148/radiology.189.1.8372180
21. Xie W, Jacobs C, Charbonnier JP, Van Ginneken B. Relational Modeling for Robust and Efficient Pulmonary Lobe Segmentation in Ct Scans. *IEEE Trans Med Imaging* (2020) 39:2664–75. doi: 10.1109/TMI.2020.2995108
22. Wang G, Liu X, Li C, Xu Z, Ruan J, Zhu H, et al. A Noise-Robust Framework for Automatic Segmentation of Covid-19 Pneumonia Lesions From Ct Images. *IEEE Trans Med Imaging* (2020) 39:2653–63. doi: 10.1109/TMI.2020.3000314
23. Narin A, Kaya C, Pamuk Z. Automatic Detection of Coronavirus Disease (Covid-19) Using X-Ray Images and Deep Convolutional Neural Networks. *Pattern Anal Appl* (2021) 24:1207–20. doi: 10.1007/s10044-021-00984-y
24. Bai HX, Wang R, Xiong Z, Hsieh B, Chang K, Halsey K, et al. AI Augmentation of Radiologist Performance in Distinguishing COVID-19 From Pneumonia of Other Etiology on Chest CT. *Radiology* (2020) 296:E156–65. doi: 10.1148/radiol.2020201491
25. Barstugan M, Ozkaya U, Ozturk S. *Coronavirus (Covid-19) Classification Using Ct Images by Machine Learning Methods*. Konya, Turkey: arXiv preprint arXiv:2003.09424 (2020).
26. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, et al. Artificial Intelligence Distinguishes Covid-19 From Community Acquired Pneumonia on Chest Ct. *Radiology* (2020) 0:200905. doi: 10.1148/radiol.2020200905
27. Hwang EJ, Park S, Jin KN, Kim JI, Choi SY, Lee J, et al. Development and Validation of a Deep Learning-Based Automated Detection Algorithm for Major Thoracic Diseases on Chest Radiographs. *JAMA Network Open* (2019) 2:e191095–e191095. doi: 10.1001/jamanetworkopen.2019.1095
28. Wang X, Deng X, Fu Q, Zhou Q, Feng J, Ma H, et al. A Weakly-Supervised Framework for Covid-19 Classification and Lesion Localization From Chest Ct. *IEEE Trans Med Imaging* (2020) 39:2615–25. doi: 10.1109/TMI.2020.2995965
29. Roy S, Menapace W, Oei S, Luijten B, Fini E, Saltori C, et al. Deep Learning for Classification and Localization of Covid-19 Markers in Point-of-Care Lung Ultrasound. *IEEE Trans Med Imaging* (2020) 39:2676–87. doi: 10.1109/TMI.2020.2994459
30. Wang J, Bao Y, Wen Y, Lu H, Luo H, Xiang Y, et al. Prior-Attention Residual Learning for More Discriminative Covid-19 Screening in Ct Images. *IEEE Trans Med Imaging* (2020) 39:2572–83. doi: 10.1109/TMI.2020.2994908
31. Oh Y, Park S, Ye JC. Deep Learning Covid-19 Features on Cxr Using Limited Training Data Sets. *IEEE Trans Med Imaging* (2020) 39:2688–700. doi: 10.1109/TMI.2020.2993291
32. Alom MZ, Rahman M, Nasrin MS, Taha TM, Asari VK. *Covid_mtnet: Covid-19 Detection With Multi-Task Deep Learning Approaches*. Dayton, OH, USA: arXiv preprint arXiv:2004.03747 (2020).
33. Özkaya U, Öztürk Ş, Barstugan M. Coronavirus (Covid-19) Classification Using Deep Features Fusion and Ranking Technique. In: *Big Data Analytics and Artificial Intelligence Against COVID-19: Innovation Vision and Approach*. Konya, Turkey: Springer (2020). p. 281–95.
34. Song Y, Zheng S, Li L, Zhang X, Zhang X, Huang Z, et al. Deep Learning Enables Accurate Diagnosis of Novel Coronavirus (Covid-19) With Ct Images. *IEEE/ACM Trans Comput Biol Bioinf* (2021) 34:102–6. doi: 10.1109/TCBB.2021.3065361
35. Kang H, Xia L, Yan F, Wan Z, Shi F, Yuan H, et al. Diagnosis of Coronavirus Disease 2019 (Covid-19) With Structured Latent Multi-View Representation Learning. *IEEE Trans Med Imaging* (2020) 39:2606–14. doi: 10.1109/TMI.2020.2992546
36. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. Las Vegas, NV, USA: CVPR (2016). p. 770–8.
37. Chen CF, Fan Q, Mallinar N, Sercu T, Feris R. *Big-Little Net: An Efficient Multi-Scale Feature Representation for Visual and Speech Recognition*. New Orleans, USA: arXiv preprint arXiv:1807.03848 (2018).
38. Chen Y, Fan H, Xu B, Yan Z, Kalantidis Y, Rohrbach M, et al. Drop an Octave: Reducing Spatial Redundancy in Convolutional Neural Networks With Octave Convolution. *CVPR* (2019), 3435–44. doi: 10.1109/ICCV.2019.00353
39. Cheng B, Xiao R, Wang J, Huang T, Zhang L. *High Frequency Residual Learning for Multi-Scale Image Classification*. Cardiff, Wales: arXiv preprint arXiv:1905.02649 (2019).
40. Gao S, Cheng M, Zhao K, Zhang X, Yang M, Torr P. Res2net: A New Multi-Scale Backbone Architecture. *IEEE Trans Pattern Anal Mach Intell* (2021) 43:652–62. doi: 10.1109/TPAMI.2019.2938758
41. Hochreiter S, Schmidhuber J. Long Short-Term Memory. *Neural Comput* (1997) 9:1735–80. doi: 10.1162/neco.1997.9.8.1735

42. Song H, Wang W, Zhao S, Shen J, Lam KM. *Pyramid Dilated Deeper Convlstm for Video Salient Object Detection*. Munich, Germany: ECCV (2018) p. 715–31.
43. Shi X, Chen Z, Wang H, Dit-Yan, Yeung, Wai-kin W, et al. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. In: *International Conference on Neural Information Processing Systems*. Montreal, Canada: NIPS (2015) p. 802–C810.
44. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. *Grad-Cam: Visual Explanations From Deep Networks via Gradient-Based Localization*. Venice, Italy: ICCV (2017) p. 618–26.
45. Zhu J, Chen N, Xing EP. *Infinite Latent Svm for Classification and Multi-Task Learning*. Granada, Spain: NIPS (2011) p. 1620–8.
46. Krizhevsky A, Sutskever I, Hinton GE. *Imagenet Classification With Deep Convolutional Neural Networks*. Nevada, USA: NIPS (2012) p. 1097–105.
47. Simonyan K, Zisserman A. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. San Diego, California, USA: arXiv preprint arXiv:1409.1556 (2014).
48. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. *Rethinking the Inception Architecture for Computer Vision*. Las Vegas, NV, USA: CVPR (IEEE (2016) p. 2818–26.
49. Chollet F. *Xception: Deep Learning With Depthwise Separable Convolutions*. Honolulu, HI, USA: CVPR (2017) p. 1251–8.
50. Tran D, Bourdev L, Fergus R, Torresani L, Paluri M. *Learning Spatiotemporal Features With 3d Convolutional Networks*. Santiago, Chile: ICCV (2015) p. 4489–97.
51. Carreira J, Zisserman A. *Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset*. Honolulu, HI, USA: CVPR (2017) p. 6299–308.
52. Xie S, Sun C, Huang J, Tu Z, Murphy K. *Rethinking Spatiotemporal Feature Learning: Speed-Accuracy Trade-Offs in Video Classification*. Munich, Germany: ECCV (2018) p. 305–21.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Hao, Xie, Liu, Hao, Ma, Yan, Liu, Zheng, Zheng, Liu, Zhang and Zhao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 20 December 2021
doi: 10.3389/fonc.2021.792420

Molecular Imaging of Ulex Europaeus Agglutinin in Colorectal Cancer Using Confocal Laser Endomicroscopy (With Video)

Weijun Wang[1,2†], Shuxin Tian[1,2,3†], Xin Jiang[1], Suya Pang[1], Huiying Shi[1], Mengke Fan[1], Zeyu Wang[1], Weiwei Jiang[1], Weiqian Hu[1], Xueyan Xiao[1] and Rong Lin[1]*

[1] Department of Gastroenterology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [2] Department of Gastroenterology, National Health Commission (NHC) Key Laboratory of Prevention and Treatment of Central Asia High Incidence Diseases, the First Affiliated Hospital of Medical College, Shihezi University, Shihezi, China, [3] Department of Gastroenterology, the First Affiliated Hospital of Medical College, Shihezi University, Shihezi, China

OPEN ACCESS

Edited by:
Kyung Hyun Sung,
UCLA Health System, United States

Reviewed by:
Evgenii Belykh,
Newark, United States
Gaetano Luglio,
University of Naples Federico II, Italy

***Correspondence:**
Rong Lin
selinalin35@hotmail.com

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 10 October 2021
Accepted: 23 November 2021
Published: 20 December 2021

Citation:
Wang W, Tian S, Jiang X, Pang S, Shi H, Fan M, Wang Z, Jiang W, Hu W, Xiao X and Lin R (2021) Molecular Imaging of Ulex Europaeus Agglutinin in Colorectal Cancer Using Confocal Laser Endomicroscopy (With Video). Front. Oncol. 11:792420. doi: 10.3389/fonc.2021.792420

Background and Study Aims: Previous studies have identified that colorectal cancer has different fucosylation levels compared to the normal colon. Ulex europaeus agglutinin-I (UEA-I), which specifically combines with α1-2 fucose glycan, is usually used to detect fucosylation levels. Therefore, we used confocal laser endomicroscopy (CLE) to investigate fluorescently labeled UEA-Fluorescein isothiocyanate (FITC) for detecting colonic cancer.

Patients and Methods: We stained frozen mouse colon tissue sections of normal, adenoma, and adenocarcinoma species with UEA-FITC to detect fucosylation levels in different groups. White light endoscopy and endocytoscopy were first used to detect the lesions. The UEA-FITC was then stained in the mice and human colon tissues *in vitro*. The CLE was used to detect the UEA-FITC levels of the corresponding lesions, and videos were recorded for quantitation analysis. The diagnostic accuracy of UEA-FITC using CLE was evaluated in terms of sensitivity and specificity.

Results: The UEA expression level in colorectal cancer was lower than that in normal intestinal epithelium. The fluorescence intensity ratio of UEA-FITC in colorectal cancer was significantly lower than that in normal tissue detected by CLE in both mice and humans. The combination of UEA-FITC and CLE presented a good diagnostic accuracy with a sensitivity of 95.6% and a specificity of 97.7% for detecting colorectal cancer. The positive and negative predictive values were 91.6% and 95.6%, respectively. Overall, 95.6% of the sites were correctly classified by CLE.

Conclusions: We developed a new imaging strategy to improve the diagnostic efficacy of CLE by using UEA-FITC.

Keywords: colorectal cancer, UEA-FITC, CLE, diagnostic efficacy, colorectal cancer imaging target

INTRODUCTION

Combining cancer-specific molecular imaging technologies with optical imaging agents provides a novel technique for cancer detection. White light endoscopy (WLE) is the initial technique used to detect gastrointestinal tract tumors (1). Nevertheless, WLE has several disadvantages: it is insensitive to detecting multifocal and flat tumors, and is insufficient to judge tumor demarcation lines (2). These disadvantages may affect the detection and complete resection of tumors, which may influence the prognosis of cancer. Cancer-specific molecular imaging agents labeled by fluorescence may augment the distinction between tumors and adjacent normal or benign tissues. Consequently, development of molecular imaging agents that can provide enhanced tumor visualization is of great significance.

Colorectal cancer (CRC) is the third most commonly diagnosed malignancy and the fourth leading cause of cancer-related deaths worldwide (3). The global morbidity rate was high at 10.2%, and the mortality rate was high at 9.2% (4). The prognosis of colon cancer is closely related to its clinical stage. The 5-year survival rate for early colon cancer is 90%, and the 5-year survival rate for advanced colon cancer is 14% (5). Therefore, early diagnosis and treatment are essential to improve patient survival.

Colon tumors mostly develop from adenomatous polyps and look similar, but the treatment strategies are completely different (6, 7). Adenomatous polyps can be removed by cold snare polypectomy and endoscopic mucosal resection (EMR); colon tumors can be removed by endoscopic submucosal dissection (ESD), according to the invasion (6). In addition, when deep submucosal infiltration is suspected or confirmed, as the lymph node metastasis increases, the corresponding surgical operations must be performed, and adjuvant chemotherapy may be needed to eradicate the cancer (7). Clinically, appropriate treatment strategies should be selected according to the benign and malignant diseases of the colon.

As the main tool to detect, biopsy, and resect gastrointestinal tumors, WLE has several shortcomings, including difficulty in recognizing flat tumors and difficulty in estimating tumor boundaries (1). Over the past two decades, some novel endoscopic techniques have emerged as promising tools to improve cancer identification and guide endoscopic resection over standard WLE assessment (8–10). Narrow-band imaging (NBI), using a narrowed bandwidth filter (specific blue and green wavelengths), can accentuate structural mucosal patterns and mucosal/submucosal vessels (8). Endocytoscopy (EC) with ultra-high magnification enables *in vivo* observation of both structural and cellular atypia during routine endoscopic examination (9). Confocal laser endomicroscopy (CLE) is another new enhanced technology. CLE can magnify mucosa up to 1000 times at a subcellular level of resolution up to 250 μm below the mucosal surface using fluorescein as the imaging agent (10). Despite the enhanced images provided by these newly developed technologies, several limitations remain in their applications. For NBI, in some cases, it is not easy to distinguish inflammatory lesions from tumor lesions, thereby contributing to high false-positive rates (11). For EC, the procedure time is prolonged and the recognition theory system has yet to be perfected (12). For CLE, fluorescein stains equally all cell types (including normal cells and cancer cells), and a real-time image version to differentiate cancer may be a great challenge (10).

CLE using a fluorescent-labeled molecular probe has been reported in order to identify colonic dysplasia and cancer (13). Herein, we also attempted to develop a new imaging approach to increase the diagnostic efficacy of the available imaging skills by employing a new fluorescein imaging agent differentially expressed in normal and tumor lesions. Previous studies have identified that colonic cancer has different fucosylation levels than the normal colon (14). Fucosylation is a type of glycosylation modification that adds fucose (6-deoxy-L-galactose) to oligosaccharides/proteins (15). It has been reported that in the absence of fucosylation, dysplasia appeared and proceeded to cancer mainly influencing the right colon and cecum (14). Ulex europaeus agglutinin-I (UEA-I), which specifically binds to α1-2 fucose glycan, is specifically used to detect fucosylation levels (16).

Therefore, we investigated the expression level of UEA in normal intestinal epithelium and colon cancer, with the purpose of estimating fluorescently labeled UEA-fluorescein isothiocyanate (FITC) as an intestinal imaging agent for detecting colonic cancer with optical imaging.

MATERIALS AND METHODS

Study Design

This study aimed to evaluate UEA as an endoscopic molecular imaging marker for colon cancer. The differential expression of α (1,2) fucosylation in normal colorectal epithelium and colorectal tumor specimens was first demonstrated by UEA-FITC staining in a mouse azoxymethane/dextran sodium sulfate (AOM/DSS)-induced colitis-associated colorectal cancer model. Fluorescently labeled UEA was investigated as a molecular imaging marker to identify colorectal cancer in mouse and human colonic specimens. UEA-FITC was administered to the intestinal mucosa, and FITC was detected using CLE.

In mice, an AOM/DSS-induced colitis-associated colorectal cancer model was used, and the detailed methods are provided below. For human colon specimens, all imaged colon specimens were selected on the basis of the endoscopic diagnosis of adenoma or adenocarcinoma. First, WLE was used to detect the susceptive lesions. Then, endocytoscopy and CLE were used to observe the details. All imaged specimens were biopsied and assessed by a pathologist who was blinded to the endoscopic results. For UEA-FITC imaging with CLE, analysis of biopsied regions was reliant on both CLE judgment and histopathology. Histopathological diagnosis was considered the gold standard. This study was approved by the Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Animal Experiment

Juvenile male C57BL/6J mice (7 ± 1 weeks old, body weight 20 ± 1 g, Beijing HFK Bio-Technology Co., LTD, China) were housed in individually ventilated cages (four animals per cage) at the SPF facility of Huazhong University of Science and Technology under controlled environmental conditions (temperature 22 ± 2°C; relative humidity 60%–70%) with free access to standard laboratory chow and tap water, and maintained under a regular 12/12-h light/dark cycle. All animal care and experimental procedures were approved by the Animal Care Ministry of Health and were performed in accordance with national and EU guidelines for the handling and use of experimental animals. Animal studies were reported in compliance with the ARRIVE guidelines.

The AOM/DSS-induced colitis-associated colorectal cancer model was developed as previously described (17). Briefly, male mice were treated with AOM (10 mg/kg). Five days later, the mice were feed with 1%–2% DSS for 7 days, followed by regular water drinking for 14 days. This cycle was repeated for three times. During the whole trial, body weight, diarrhea, and macroscopic bleeding of the mice were examined. At the end of the animal experiment, after overnight fasting, mice were sacrificed under anesthesia, and colorectums were collected for WLE, CE, and CLE observation.

Endocytoscopy

All colonoscopic inspections were conducted by well-trained endoscopists (more than 1,500 colonoscopies) using an integrated-type EC (Olympus, Tokyo, Japan). This instrument has a single lens on its tip with a hand lever; the one-touch switch or the hand lever enables consecutively increasing magnification power from the conventional endoscopic image to ultra-magnification power, without changing the scope. This device has a 520-fold magnification power with a focus depth of 50 μm, which can image the gland duct lumens and the shape of nuclei in the epithelial superficial layer. The images were obtained using 0.05% crystal violet and 1% methylene blue. The intestinal mucosa of mice was examined *in vitro* because of its narrow lumen, and the human intestinal mucosa was observed *in vivo*. Uniformly sized fusiform nuclei and roundish lumens regularly represent normal mucosa. Slit-like smooth lumens and uniform fusiform or roundish nuclei indicate adenoma. Vague gland formation and agglomeration of distorted nuclei suggested adenocarcinoma (**Figures 2A** and **3A**).

UEA-FITC Imaging With CLE

After examination by WLE and EC, UEA-FITC staining was used for CLE. Colorectums were washed with saline; then, UEA-FITC (1/50 mg/ml) and IgG-FITC (1/50 mg/ml) were incubated for 2–3 min. Tissue-bound UEA-FITC was detected using a clinical CLE system. This probe-based system (pCLE) (Cellvizio Endomicroscopy System; Mauna Kea Technologies, Paris, France) is commercially available and consists of a flexible miniprobe that may be introduced through the working channel of a standard endoscope. This device was used with a 2.8-mm fiber optic probe. The probe provided a field of view of 240 μm and acquired videos. A typical CLE image of stained normal mucosa displays a hexagonal, honeycomb appearance of vessel architecture along with regular, round luminal openings, surrounded by a homogeneous layer of epithelial cells. Irregular epithelium and villiform- or tubular-shaped crypts all indicate adenoma. Adenocarcinoma causes disorders and destruction of the glands. Videos were taken, and data from entire video sequences (12 frames/s, 10- to 20-s videos) were used for quantitation analysis. For mice, 10 videos from normal mice (n = 10), 19 videos of 19 adenoma specimens from mice with adenoma (n = 5), and 20 videos of 20 tumor specimens from mice with tumors (n = 5) were analyzed. For humans, 10 videos from normal humans (n = 10), six videos of six adenoma specimens from patients with adenoma (n = 6), and four videos of four tumor specimens from patients with tumors (n = 4) were analyzed. To validate the quantitative analysis, we set to cancel the automatic control function of the laser power when using the CLE system. After examination by CLE, representative lesions were biopsied for histopathological correlation.

Statistical Analysis

Statistical analysis was performed using SPSS software. All data are presented as the mean (95% confidence interval). One-way analysis of variance tests were used to determine the significance of UEA expression levels in different groups. For the diagnostic accuracy of UEA-FITC using CLE, sensitivity was calculated as TP/(TP + FN) and specificity was calculated as TN/(TN + FP), where TP equaled the number of true-positive cases, FN equaled the number of false-negative cases, TN equaled the number of true-negative cases, and FP equaled the number of false-positive cases.

RESULTS

UEA-FITC Expression and Distribution

Immunofluorescence was conducted to detect UEA-FITC in frozen mouse tissue sections of normal intestinal epithelium, adenoma, and colorectal tumor; however, the distribution was different (**Figure 1A**). The single layer of normal intestinal epithelium cells showed widely distributed UEA expression on the mucosal epithelial layer, whereas UEA expression on the mucosal epithelial layer decreased in cancer tissue. Furthermore, the UEA expression level of colorectal adenoma was between them both (**Figure 1B**). This difference in distribution indicates that UEA may be a feasible agent for the endoscopic molecular imaging of colorectal cancer.

Endoscopic Molecular Imaging for Mice

An *ex vivo* endoscopic molecular imaging protocol was devised for mice (**Figure 2**). Fresh mouse colorectums were obtained from a mice AOM/DSS-induced colitis-associated colorectal cancer model (n = 10). For its narrow intestinal cavity, the colorectum was opened longitudinally for the sake of white light visual examination of the intestinal epithelium using a camera.

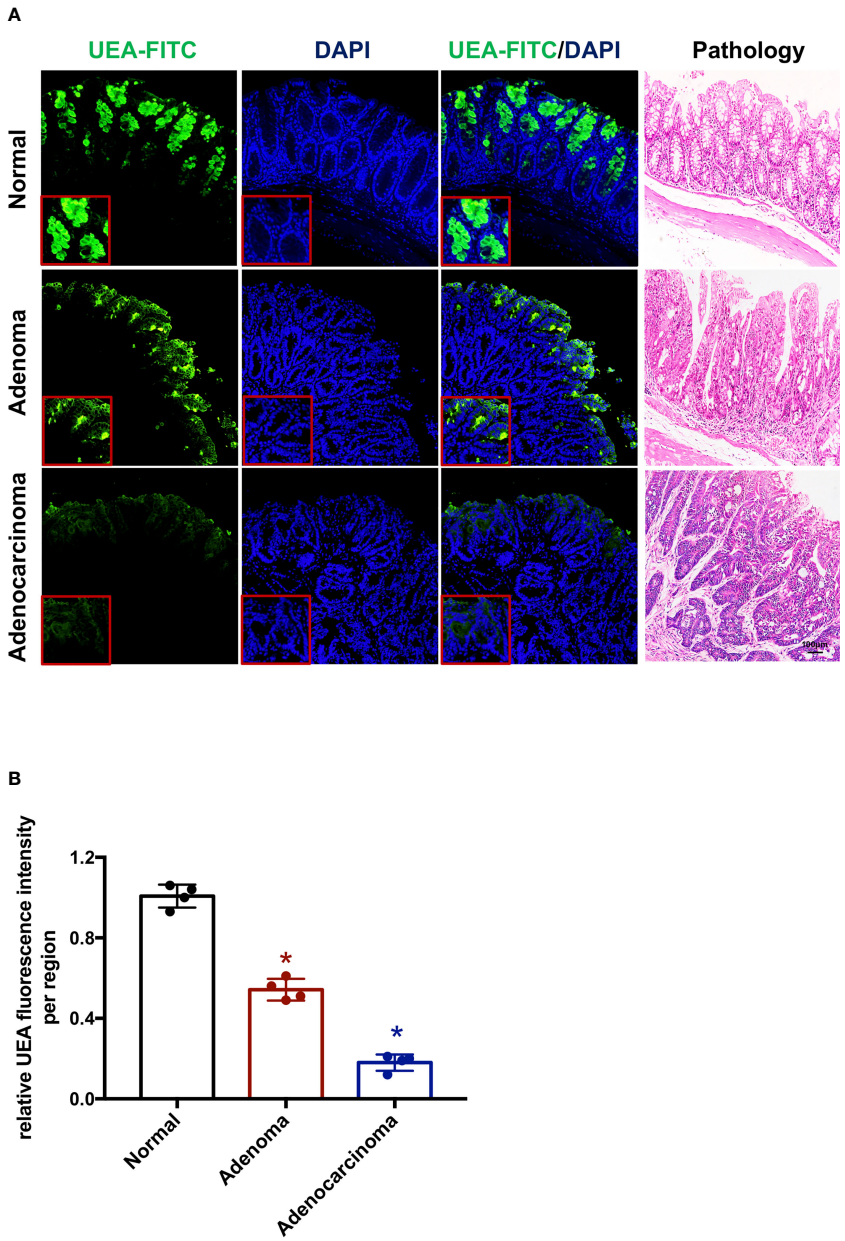


FIGURE 1 | UEA-FITC expression and distribution in colorectal cancer. **(A)** Immunofluorescence of colorectum tissue sections showing UEA distribution in normal intestinal epithelium, adenoma, and cancer. The corresponding pathologic pictures were also shown. Scale bars, 250 μm. **(B)** Cancer/adenoma-to-normal fluorescence signal ratio of each colorectum (n = 4) imaged with immunofluorescence. Error bars, SD. *$p < 0.05$. UEA, ulex europaeus agglutinin; FITC, fluorescein isothiocyanate; SD, standard deviation.

After spraying 0.05% crystal violet plus 1% methylene blue onto the surface of a polyp or adenocarcinoma together with its surrounding normal mucosa, EC was used to observe the surface of the mucosa. Typical distinct EC images of stained normal mucosa, adenoma, and adenocarcinoma are shown in **Figure 2A**. Subsequently, a UEA labeled with fluorescein isothiocyanate (FITC) was instilled after washing the mucosa with saline. IgG-FITC was also used as a negative control (**Supplementary Figure 1**). Colorectums incubated with UEA-FITC and IgG-FITC were imaged using CLE. Normal, adenoma, or adenocarcinoma regions in each colorectum were imaged and biopsied. A pathologist was blinded to the imaging results and analyzed the histologic results. All colorectums were confirmed to have colorectal cancer on the basis of the histologic results. Representative frames of normal intestinal, adenoma, and adenocarcinoma lesions in each colorectum with corresponding histopathology taken by confocal videos are also shown in **Figure 2A**.

To quantitatively analyze the UEA-FITC signal in confocal videos, the average fluorescence intensity of all frames in the video was calculated as the mean fluorescence intensity. Videos from a normal region, adenoma region, and cancer (endoscopically diagnosed) were applied to determine the fluorescence intensity ratio of UEA-FITC coupling to cancer, adenoma, and normal tissue (**Figure 2B**). The results showed that the fluorescence intensity ratio of UEA-FITC in cancer tissues was significantly lower than that in normal tissues ($p < 0.05$).

Endoscopic Molecular Imaging for Humans

For human colon specimens, all specimens were selected on the basis of endoscopic diagnosis of adenoma or adenocarcinoma ($n = 10$). The adenoma and adenocarcinoma lesions were removed endoscopically by EMR, ESD, or by surgery. The observation procedure for human colorectums was the same as that for mouse colorectums. The colorectums were first investigated by WLE and EC, and then by CLE. After incubation with UEA-

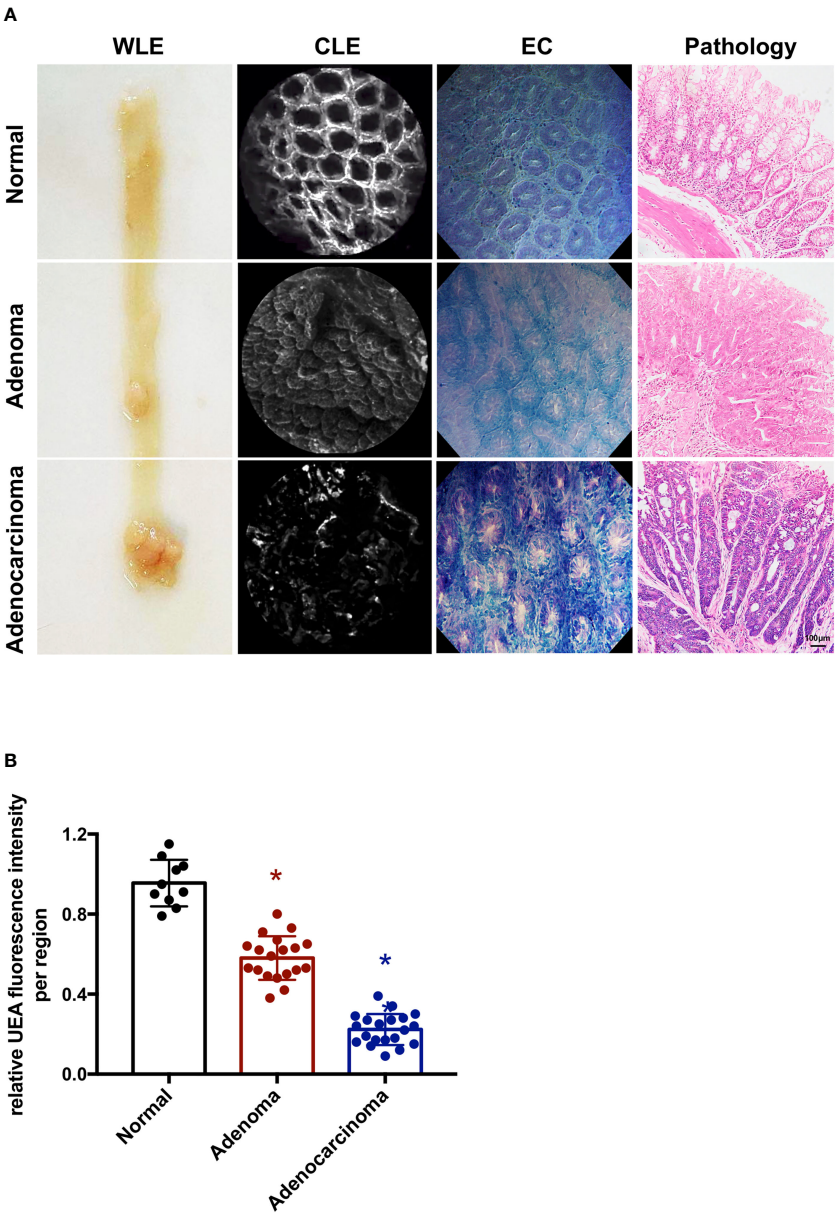


FIGURE 2 | Endoscopic molecular imaging of mouse colorectal cancer using UEA-FITC and CLE. **(A)** Representative frames of WLE, CLE, and EC acquired from normal, adenoma, and cancer lesions, and the corresponding pathologic pictures. Scale bars, 100 μm. **(B)** Cancer/adenoma-to-normal fluorescence signal ratio of each colorectum ($n = 4$) imaged with CLE. Error bars, SD. *$p < 0.05$. UEA, ulex europaeus agglutinin; FITC, fluorescein isothiocyanate; WLE, white light endoscopy; CLE, confocal laser endomicroscopy; EC, endocytoscopy; SD, standard deviation.

FITC, fiber optic probe was done. IgG-FITC was used for the negative control (**Supplementary Figure 1**). The lesions were imaged and evaluated by a pathologist who was blinded to the imaging results. All 10 colorectal tumors were pathologically confirmed to have adenomas or cancers by histology. Representative frames from confocal videos taken of adenoma or adenocarcinoma lesions and para-cancer normal tissues with corresponding histopathology are shown in **Figure 3A**. The average fluorescence intensity of all frames in a confocal video was determined so as to quantitatively analyze the UEA-FITC signal. For each colorectum, videos from a normal region, adenoma region, and cancer were obtained to calculate the fluorescence intensity ratio of UEA-FITC coupling to cancer, adenoma, and normal tissue (**Figure 3B**). The results showed that compared to normal tissue, the fluorescence intensity ratio of UEA-FITC in cancer was significantly lower ($p < 0.05$).

Diagnostic Accuracy of UEA-Targeted Imaging

Confocal images, including mouse and human specimens, were used prospectively to predict the histopathology. Overall, 69 different locations in 20 colorectums, including 20 locations

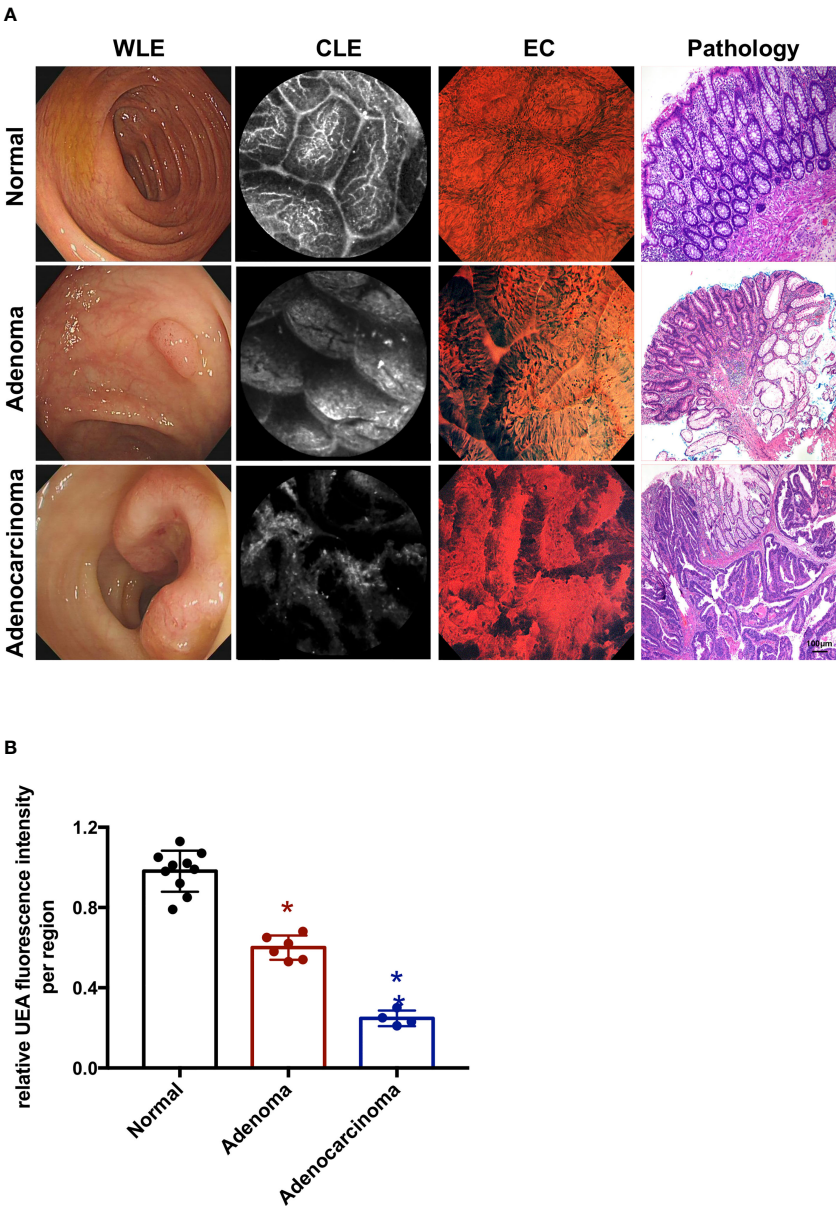


FIGURE 3 | Endoscopic molecular imaging of human colorectal cancer using UEA-FITC and CLE. **(A)** Representative frames of WLE, CLE, and EC acquired from normal, adenoma, and cancer lesions, and the corresponding pathologic pictures. Scale bars, 100 μm. **(B)** Cancer/adenoma-to-normal fluorescence signal ratio of each colorectum ($n = 4$) imaged with CLE. Error bars, SD. *$p < 0.05$. UEA, ulex europaeus agglutinin; FITC, fluorescein isothiocyanate; WLE, white light endoscopy; CLE, confocal laser endomicroscopy; EC, endocytoscopy; SD, standard deviation.

from normal mucosa, 25 locations from adenoma, and 24 locations from cancer, were analyzed using a blind method. CLE was applied to evaluate all sites before biopsy, and histological results were compared with CLE judgment. The results of CLE judgment and histology are displayed in **Table 1**. The sensitivity and specificity of CLE for malignant lesion detection were 95.6% and 97.7%, respectively. The positive and negative predictive values were 91.6% and 95.6%, respectively. The overall correction rate for all sites was 95.6% by the CLE.

DISCUSSION

The enhancement of endoscopic detection of colorectal cancer can increase the efficiency of tumor excision and decrease cancer recurrence and progression. Here, we show three main novel findings: (1) UEA is highly expressed in the normal intestinal epithelium but less in colorectal cancer, (2) the combination of UEA-FITC and CLE helped to identify patients with colorectal cancer, and (3) the real-time assessment by CLE conducted on UEA-FITC-targeted locations has a good diagnostic accuracy for colorectal cancer (sensitivity 96.4% and specificity 97.5%).

Previous studies have identified that colonic cancer has different fucosylation levels compared to normal colons (14). It has been reported that α1-3/4 fucosylation in the sera of colon cancer patients was higher than that in other groups (18). However, another study demonstrated that in the deficiency of fucosylation, dysplasia appeared and progressed to adenocarcinoma, mainly affecting the right colon and cecum (14). UEA-I, which combines with α1-2 fucose glycan, is usually used to detect α1-2 fucosylation levels. Our present study demonstrated decreased UEA levels in colonic cancer compared to normal colon tissue (**Figure 1**). This difference in distribution indicates that UEA may be a feasible agent for the endoscopic molecular imaging of colorectal cancer.

UEA-I is a lectin that specifically couples with α-linked fucose (16). Lectins are highly stereospecific proteins that can recognize numerous sugar structures and interact with glyco-conjugate complexes to form reversible bindings (19). Above all, they are abundantly discovered in plants and animals. Therefore, the UEA-I used here is a safe protein that is suitable for human use, although in our study, it was used *in vitro*. Furthermore, as the colorectum is a cavity organ, UEA-I can be administered locally without significant systemic absorption. Therefore, UEA-I is a viable and safe target for molecular imaging. The organ imaging protocol developed here facilitates the clinical use of UEA-targeted imaging for the detection and resection of colorectal cancer.

CLE combined with UEA-FITC can reveal differences in the morphology of the staining pattern. As shown in **Figures 2A** and **3A**, a typical CLE image of stained normal mucosa displays a hexagonal, honeycomb appearance of vessel architecture along with regular, round luminal openings, surrounded by a homogeneous layer of epithelial cells. Irregular epithelium and villiform- or tubular-shaped crypts all indicate adenoma. Adenocarcinoma causes disorders and destruction of the glands. These different morphologies could help to diagnose normal, adenoma, and adenocarcinoma lesions in the first step.

In addition, CLE combined with UEA-FITC was investigated for the digital nature of the confocal images, allowing for semi-quantitative analysis. *In vivo* endoscopic molecular imaging combining fluorescent agents and CLE has also been reported for esophageal lesions (20) and gastric dysplasia (21). In esophageal applications, the addition of a 3-biomarker panel (cyclin A, p53, and aneuploidy) led to a 50% reduction in the false-positive rate of CLE, obtaining a sensitivity of 89.2% and a specificity of 88.9% for any grade of dysplasia, respectively (20). In the gastric study, targeted H-ferritin was demonstrated in resected human early gastric cancers with good diagnostic efficacy (100% sensitivity and 90% specificity) using CLE (21). In previous studies, quantitation was determined using images obtained from video sequences (20, 21). In this study, data of normal and cancer regions collected from the whole confocal video sequences (12 frames/s, 10- to 20-s videos) were incorporated for analysis. We measured the UEA-FITC fluorescence intensity ratio of tumor tissue to normal tissue using this unbiased method (**Figures 2B** and **3B**).

Although CLE allowed for quantitation, the small field of view impeded efficient examination of the entire colon. WLE was first used to observe the entire colon, then EC was applied to detect suspicious lesions, and CLE was subsequently used. The combination of UEA-FITC and CLE has a good diagnostic accuracy for colorectal cancer (sensitivity 95.6% and specificity 97.7%), which is quite higher than the data reported on CLE alone (sensitivity 81%–91.4% and specificity 76%–85.7%) (22, 23). Estimating the benign and malignant lesions and determining the boundary of the malignant lesions are the key points for the diagnosis and treatment of colorectal cancer. The combination of UEA-FITC and CLE may help improve colorectal cancer detection and the thoroughness of tumor resection.

This study had several limitations. First, this study was limited to an *ex vivo* investigation, as required by the local ethics committee. However, despite the small number of human specimens, given the experience of our *ex vivo* human

TABLE 1 | Correlation between confocal imaging and histology.

Confocal diagnosis	Sites	Histology		
		Normal	Adenoma	Adenocarcinoma
Normal	20	20	0	0
Adenoma	25	2	22	1
Adenocarcinoma	24	0	2	22

investigation, we are confident that topical administration *in vivo* of UEA-FITC would highlight adenoma and adenocarcinoma lesions in patients as well. Secondly, the relatively high cost and long learning curve of the procedure make it difficult for clinical use. Nevertheless, with the development of the economy and technology, we believe that it will be feasible in the near future. Thirdly, the actual advantages with respect to histology as the gold standard made this new approach more reliable, but it still could provide very reliable information.

In summary, we provide new concepts for identifying and validating molecular imaging targets for colorectal cancer. UEA can act as a cancer-specific imaging agent to improve the detection and treatment of CRC. The present imaging approach utilizes the simplicity of accessing the colorectal lumen, the well-established intraintestinal route for administering drugs in the colorectum, and the availability of clinical level imaging tools involving EC and CLE. We demonstrate that this method is feasible and may be adaptable for further development and validation of imaging agents in other hollow organs. Together, this research provides evidence for UEA as a colorectal cancer imaging target with the potential for clinical translation.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology, Affiliated to the Huazhong University of Science and Technology. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements. The animal study was reviewed and approved by Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology, Affiliated to the Huazhong University of Science and Technology. Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

WW and ST: Study design, manuscript preparation, statistical analysis, and study coordinator. XJ and SP: Animal experiment and data collection. HS, MF, and WJ: Patient experiment and data collection. ZW and WH: Animal data acquisition, and data analysis and interpretation. XX: Revising the manuscript. RL: Study design, manuscript revision, study supervision, and foundation support. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 81900580, 81974068, and 81770539), the Natural Science Foundation of Hubei Province (No. 2017CFA061), the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2020-PT330-003), and the Open Research Fund of NHC Key Laboratory of Prevention and Treatment of Central Asia High Incidence Diseases.

ACKNOWLEDGMENTS

The authors appreciate the language correction from Editage.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.792420/full#supplementary-material

Supplementary Figure 1 | Endoscopic molecular imaging of colorectal cancer using IgG-FITC and CLE in mouse and human.

Supplementary Video 1 | A video of CLE combined with UEA to detect normal colon tissue in mice CLE, confocal laser endomicroscopy; UEA, Ulex europaeus agglutinin.

Supplementary Video 2 | A video of CLE combined with UEA to detect colon adenoma tissue in mice CLE, confocal laser endomicroscopy; UEA, Ulex europaeus agglutinin.

Supplementary Video 3 | A video of CLE combined with UEA to detect colorectal cancer in mice CLE, confocal laser endomicroscopy; UEA, Ulex europaeus agglutinin.

REFERENCES

1. Chiu PWY, Uedo N, Singh R, Gotoda T, Ng EKW, Yao K, et al. An Asian Consensus on Standards of Diagnostic Upper Endoscopy for Neoplasia. *Gut* (2019) 68:186–97. doi: 10.1136/gutjnl-2018-317111
2. van der Sommen F, Curvers WL, Nagengast WB. Novel Developments in Endoscopic Mucosal Imaging. *Gastroenterology* (2018) 154:1876–86. doi: 10.1053/j.gastro.2018.01.070
3. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Patterns and Trends in Colorectal Cancer Incidence and Mortality. *Gut* (2017) 66:683–91. doi: 10.1136/gutjnl-2015-310912
4. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492
5. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, et al. Colorectal Cancer Statistics, 2020. *CA Cancer J Clin* (2020) 70(3):145–64. doi: 10.3322/caac.21601
6. Ferlitsch M, Moss A, Hassan C, Bhandari P, Dumonceau JM, Paspatis G, et al. Colorectal Polypectomy and Endoscopic Mucosal Resection (EMR): European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline. *Endoscopy* (2017) 49:270–97. doi: 10.1055/s-0043-102569

7. Benson AB 3rd, Venook AP, Cederquist L, Chan E, Chen YJ, Cooper HS, et al. Colon Cancer, Version 1.2017, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* (2017) 15:370–98. doi: 10.6004/jnccn.2017.0036
8. Atkinson NSS, Ket S, Bassett P, Aponte D, De Aguiar S, Gupta N, et al. Narrow-Band Imaging for Detection of Neoplasia at Colonoscopy: A Meta-Analysis of Data From Individual Patients in Randomized Controlled Trials. *Gastroenterology* (2019) 157:462–71. doi: 10.1053/j.gastro.2019.04.014
9. Kumagai Y, Takubo K, Kawada K, Higashi M, Ishiguro T, Sobajima J, et al. A Newly Developed Continuous Zoom-Focus Endocytoscope. *Endoscopy* (2017) 49:176–80. doi: 10.1055/s-0042-11926
10. Wanders LK, East JE, Uitentuis SE, Leeflang MM, Dekker E. Diagnostic Performance of Narrowed Spectrum Endoscopy, Autofluorescence Imaging, and Confocal Laser Endomicroscopy for Optical Diagnosis of Colonic Polyps: A Meta-Analysis. *Lancet Oncol* (2013) 14:1337–47. doi: 10.1016/S1470-2045(13)70509-6
11. van den Broek FJ, Fockens P, van Eeden S, Reitsma JB, Hardwick JC, Stokkers PC, et al. Endoscopic Tri-Modal Imaging for Surveillance in Ulcerative Colitis: Randomised Comparison of High-Resolution Endoscopy and Autofluorescence Imaging for Neoplasia Detection; and Evaluation of Narrow-Band Imaging for Classification of Lesions. *Gut* (2008) 57:1083–9. doi: 10.1136/gut.2007.144097
12. Ichimasa K, Kudo SE, Mori Y, Wakamura K, Ikehara N, Kutsukawa M, et al. Double Staining With Crystal Violet and Methylene Blue is Appropriate for Colonic Endocytoscopy: An *In Vivo* Prospective Pilot Study. *Dig Endosc* (2014) 26:403–8. doi: 10.1111/den.12164
13. De Palma GD, Colavita I, Zambrano G, Giglio MC, Maione F, Luglio G, et al. Detection of Colonic Dysplasia in Patients With Ulcerative Colitis Using a Targeted Fluorescent Peptide and Confocal Laser Endomicroscopy: A Pilot Study. *PloS One* (2017) 12:e0180509. doi: 10.1371/journal.pone.0180509
14. Wang Y, Huang D, Chen KY, Cui M, Wang W, Huang X, et al. Fucosylation Deficiency in Mice Leads to Colitis and Adenocarcinoma. *Gastroenterology* (2017) 152(1):193–205 e10. doi: 10.1053/j.gastro.2016.09.004
15. Goto Y, Uematsu S, Kiyono H. Epithelial Glycosylation in Gut Homeostasis and Inflammation. *Nat Immunol* (2016) 17(11):1244–51. doi: 10.1038/ni.3587
16. Tian R, Zhao S, Liu G, Chen H, Ma L, You H, et al. Construction of Lanthanide-Doped Upconversion Nanoparticle-Uelx Europaeus Agglutinin-I Bioconjugates With Brightness Red Emission for Ultrasensitive *In Vivo* Imaging of Colorectal Tumor. *Biomaterials* (2019) 212:64–72. doi: 10.1016/j.biomaterials.2019.05.010
17. Isermann T, Şener ÖÇ, Stender A, Klemke L, Winkler N, Neesse A, et al. Suppression of HSF1 Activity by Wildtype P53 Creates a Driving Force for P53 Loss-of-Heterozygosity. *Nat Commun* (2021) 12(1):4019. doi: 10.1038/s41467-021-24064-1
18. Park SY, Lee SH, Kawasaki N, Itoh S, Kang K, Hee Ryu S, et al. Alpha1-3/4 Fucosylation at Asn 241 of Beta-Haptoglobin is a Novel Marker for Colon Cancer: A Combinatorial Approach for Development of Glycan Biomarkers. *Int J Cancer* (2012) 130(10):2366–76. doi: 10.1002/ijc.26288
19. Mishra A, Behura A, Mawatwal S, Kumar A, Naik L, Mohanty SS, et al. Structure-Function and Application of Plant Lectins in Disease Biology and Immunity. *Food Chem Toxicol* (2019) 134:110827. doi: 10.1016/j.fct.2019.110827
20. di Pietro M, Bird-Lieberman EL, Liu X, Nuckcheddy-Grant T, Bertani H, O'Donovan M, et al. Autofluorescence-Directed Confocal Endomicroscopy in Combination With a Three-Biomarker Panel Can Inform Management Decisions in Barrett's Esophagus. *Am J Gastroenterol* (2015) 110(11):1549–58. doi: 10.1038/ajg.2015.295
21. Du Y, Fan K, Zhang H, Li L, Wang P, He J, et al. Endoscopic Molecular Imaging of Early Gastric Cancer Using Fluorescently Labeled Human H-Ferritin Nanoparticle. *Nanomedicine* (2018) 14:2259–70. doi: 10.1016/j.nano.2018.07.007
22. Shahid MW, Buchner AM, Raimondo M, Woodward TA, Krishna M, Wallace MB. Accuracy of Real-Time vs. Blinded Offline Diagnosis of Neoplastic Colorectal Polyps Using Probe-Based Confocal Laser Endomicroscopy: A Pilot Study. *Endoscopy* (2012) 44(7):343–8. doi: 10.1055/s-0031-1291589
23. Andre B, Vercauteren T, Buchner AM, Krishna M, Ayache N, Wallace MB. Software for Automated Classification of Probe-Based Confocal Laser Endomicroscopy Videos of Colorectal Polyps. *World J Gastroenterol* (2012) 18(39):5560–9. doi: 10.3748/wjg.v18.i39.5560

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Wang, Tian, Jiang, Pang, Shi, Fan, Wang, Jiang, Hu, Xiao and Lin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 21 December 2021
doi: 10.3389/fonc.2021.801876

Deep Learning Enables Prostate MRI Segmentation: A Large Cohort Evaluation With Inter-Rater Variability Analysis

Yongkai Liu[1,2], Qi Miao[1,3], Chuthaporn Surawech[1,4], Haoxin Zheng[1,5], Dan Nguyen[6], Guang Yang[7], Steven S. Raman[1] and Kyunghyun Sung[1,2]*

[1] Department of Radiological Sciences, David Geffen School of Medicine, University of California, Los Angeles, CA, United States, [2] Physics and Biology in Medicine Interdisciplinary Program (IDP), David Geffen School of Medicine, University of California, Los Angeles, CA, United States, [3] Department of Radiology, The First Affiliated Hospital of China Medical University, Shenyang City, China, [4] Department of Radiology, Division of Diagnostic Radiology, Faculty of Medicine, Chulalongkorn University and King Chulalongkorn Memorial Hospital, Bangkok, Thailand, [5] Department of Computer Science, Henry Samueli School of Engineering and Applied Science, University of California, Los Angeles, CA, United States, [6] Medical Artificial Intelligence and Automation Laboratory, Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, TX, United States, [7] National Heart and Lung Institute, Imperial College London, London, United Kingdom

OPEN ACCESS

Edited by:
Ming Li,
Fudan University, China

Reviewed by:
Weiwei Zong,
Henry Ford Health System, United States
Enrico Checcucci,
Istituto di Candiolo (IRCCS), Italy

***Correspondence:**
Qi Miao
meganmiao89@gmail.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *25 October 2021*
Accepted: *23 November 2021*
Published: *21 December 2021*

Citation:
Liu Y, Miao Q, Surawech C, Zheng H, Nguyen D, Yang G, Raman SS and Sung K (2021) Deep Learning Enables Prostate MRI Segmentation: A Large Cohort Evaluation With Inter-Rater Variability Analysis. Front. Oncol. 11:801876. doi: 10.3389/fonc.2021.801876

Whole-prostate gland (WPG) segmentation plays a significant role in prostate volume measurement, treatment, and biopsy planning. This study evaluated a previously developed automatic WPG segmentation, deep attentive neural network (DANN), on a large, continuous patient cohort to test its feasibility in a clinical setting. With IRB approval and HIPAA compliance, the study cohort included 3,698 3T MRI scans acquired between 2016 and 2020. In total, 335 MRI scans were used to train the model, and 3,210 and 100 were used to conduct the qualitative and quantitative evaluation of the model. In addition, the DANN-enabled prostate volume estimation was evaluated by using 50 MRI scans in comparison with manual prostate volume estimation. For qualitative evaluation, visual grading was used to evaluate the performance of WPG segmentation by two abdominal radiologists, and DANN demonstrated either acceptable or excellent performance in over 96% of the testing cohort on the WPG or each prostate sub-portion (apex, midgland, or base). Two radiologists reached a substantial agreement on WPG and midgland segmentation (κ = 0.75 and 0.63) and moderate agreement on apex and base segmentation (κ = 0.56 and 0.60). For quantitative evaluation, DANN demonstrated a dice similarity coefficient of 0.93 ± 0.02, significantly higher than other baseline methods, such as DeepLab v3+ and UNet (both p values < 0.05). For the volume measurement, 96% of the evaluation cohort achieved differences between the DANN-enabled and manual volume measurement within 95% limits of agreement. In conclusion, the study showed that the DANN achieved sufficient and consistent WPG segmentation on a large, continuous study cohort, demonstrating its great potential to serve as a tool to measure prostate volume.

Keywords: prostate segmentation, deep attentive neural network, large cohort evaluation, qualitative evaluation, quantitative evaluation, volume measurement

INTRODUCTION

Whole-prostate gland (WPG) segmentation plays an important role in prostate volume measurement, biopsy, and surgical planning (1). Magnetic resonance imaging (MRI)-targeted transrectal ultrasound fusion (MRI-fusion) biopsy has shown increased detection of clinically significant PCa and reduced identification of clinically insignificant PCa (2), where the WPG segmentation is critical to enabling the MRI-fusion biopsy (3). Also, prostate volume measurement *via* WPG segmentation can be used to quantify the progression of benign prostatic hyperplasia (1) and to assist surgical planning (4).

Manual segmentation of WPG is time-consuming and laborious and commonly suffers from inter-rater variability (5), making it inadequate for large-scale applications (6). Deep learning (DL) (7–10) has increasingly been utilized for the automatic segmentation of WPG. Zhu et al. (11) proposed a deeply supervised convolutional neural network (CNN) using convolutional information to segment the prostate from MR images. Cheng et al. (8) developed a DL model with holistically nested networks for prostate segmentation on MRI. Jia et al. (12) proposed an atlas registration and ensemble deep CNN-based prostate segmentation. In addition, attentive DL (13) models were introduced to enhance DL by paying attention to the particular regions of interest in an adaptive way and thus have achieved better segmentation performance than other DL-based models. However, to the best of our knowledge, the evaluation of these methods was currently limited by a relatively small sample size, ranging from tens to hundreds of MRI scans. It is relatively expensive to create large, continuous samples with manual segmentation of WPG, which limits the ability to test the DL models in a clinical setting.

In this paper, we evaluated a previously developed DL-based automatic segmentation model, deep attentive neural network (DANN) (13), using a large, continuous cohort of prostate 3T MRI scans acquired between 2016 and 2020. The WPG segmentation by DANN was evaluated both quantitatively and qualitatively. The quantitative evaluation was performed by using an independent testing set with manual segmentation as a ground truth on a small dataset (n = 100). The dice similarity coefficient (DSC) (14) was used to measure the segmentation performance, compared with other baseline DL methods. For qualitative evaluation, the segmentation performance was evaluated by two abdominal radiologists independently *via* visual grading since the ground-truth manual segmentation was not available for the large cohort (n = 3,210). Inter-rater agreement between the two radiologists was evaluated to check the consistency of the visual grading. We further investigated the segmentation on different anatomical locations (i.e., apex, midgland, and base) as a secondary analysis. Finally, we conducted the volume measurement using DANN-based segmentation on a small cohort (n = 50) (DANN-enabled volume measurement) and compared it with the manual volume measurement.

MATERIALS AND METHODS

MRI Datasets

With approval from the institutional review board (IRB), this retrospective study was carried out in compliance with the United States Health Insurance Portability and Accountability Act (HIPAA) of 1996. After excluding MRI scans with severe artifacts and patients with prior surgery history and Foley catheter, a total of 3,695 MRI scans, acquired on 3 T scanners (Skyra, Prisma, and Vida, Siemens Healthineers, Erlangen, Germany), from January of 2016 to August of 2020, were included in the study. Axial and coronal T2-weighted (T2W) Turbo spin-echo (TSE) images were used. **Table 1** shows the characteristics of the T2W MRI scan in the study.

Out of 3,695 3T MRI scans, 335 MRI scans (9%) were used as a training set, and the remaining 3,360 (91%) MRI scans were used as a testing set. Training and testing datasets were randomly chosen from the whole dataset. The testing set included a qualitative evaluation set (n = 3,210), a quantitative evaluation set (n = 100), and a volume measurement evaluation set (n = 50). **Table 2** shows the data characteristics for each dataset. Training, quantitative, and volume measurement evaluation sets required manual prostate contours as the segmentation reference standard. The manual annotation was prepared by an abdominal radiologist (MQ) with more than 5 years of experience in the interpretation of prostate MRI. In the training set, prostate contours were drawn on all axial T2W images from all MRI scans and on four mid-coronal T2W images (8th to 11th out of twenty slices) from a subset of 100 MRI scans. In the quantitative and volume measurement evaluation sets, prostate contours were drawn on all axial T2W images.

TABLE 1 | T2-weighted TSE MRI sequence parameters in the study.

View	Axial	Coronal
Matrix size	320 × 320	320 × 320
Flip angle	160°	147°
Resolution	0.625 × 0.625 × 3.6	0.625 × 0.625 × 3.6
Field of view (mm^2)	200 × 200	200 ×200
Repetition time (ms)	3,000–7,480	2,880–7,200
Echo time (ms)	97–112	97–109
Number of slices	20	20
Scan time (s)	200	200

ms, millisecond; s, second; mm, millimeter.

TABLE 2 | Data characteristics in the training, qualitative, and quantitative evaluation.

		Training dataset	Qualitative evaluation dataset	Quantitative evaluation dataset	Volume evaluation dataset
Number of MRI scans		335	3,210	100	50
Number of patients with endorectal coil		3	84	0	0
MRI scans with different vendors	Skyra	295	2,806	93	45
	Prisma	10	145	4	3
	Vida	30	259	3	2

DL-Based Whole Prostate Gland Segmentation Model

Figure 1 shows the overall workflow of the automatic WPG segmentation with DANN (13). We added the segmentation on the coronal plane to assist the selection of axial slices, reducing the inference time of segmentation on the axial plane. During the testing, the workflow went through the following steps. First, $DANN_{cor}$, responsible for segmenting coronal slices, was adopted to segment the prostate on the two-middle coronal images (9th and 10th slices out of twenty slices) for each MRI scan in the entire testing set. The segmented coronal images were used to automatically select the axial T2W images that contained the prostate gland. This would provide proper through-plane coverage of the prostate in the axial slices. Next, $DANN_{ax}$ was used to perform the WPG segmentation on the selected axial T2W images for each MRI scan in all the testing sets.

Both $DANN_{ax}$ and $DANN_{cor}$ were trained independently using the training set (n = 335). First, a subset of the training data (n = 100) was used for training of $DANN_{cor}$, and four-middle coronal slices (8th to 11th slices out of twenty slices) were used to make use of as many samples as possible. Once the initial training of $DANN_{cor}$ was finished, two middle coronal slices were used as input to $DANN_{cor}$ for the rest of the training data. The segmented coronal slices by $DANN_{cor}$ were used to select certain axial slices, and $DANN_{ax}$ was trained using all the selected axial slices in the entire training set. Training and inferencing were conducted on a desktop computer with a 64-Linux system with 4 Titan Xp GPU of 32 GB GDDR5 RAM. All the networks were trained with stochastic gradient descent as the optimizer, with binary cross-entropy as the loss function. PyTorch was used to implement all the DL networks. The models were initially trained using 80% of the training dataset, and the rest of the training dataset was used as the validation dataset. After the optimal hypermeters were found, we retrained the models using the whole training dataset. The learning rate was initially set to 2.5e-3. All the networks were trained for 100 epochs with batch size 12.

Evaluation of Segmentation Performance

Qualitative Evaluation of Segmentation Performance

We adopted the visual grading, similar to (15), to qualitatively evaluate the WPG segmentation. Two abdominal radiologists (MQ and CS; each has over five years of experience in prostate MRI interpretation) assigned a visual grade, ranging from 1 to 3, to evaluate the segmentation performance, focusing on the whole prostate and sub-portions of the prostate (e.g., apex, midgland, and base). 1, 2, and 3 indicate unacceptable, acceptable, and excellent segmentation performance, respectively. **Table 3** shows the details when assigning the visual grade. Typical examples associated with each visual grade are shown in **Figure 2**. The readers independently ranked the segmentation quality. In addition, inter-rater reliability was assessed. To further

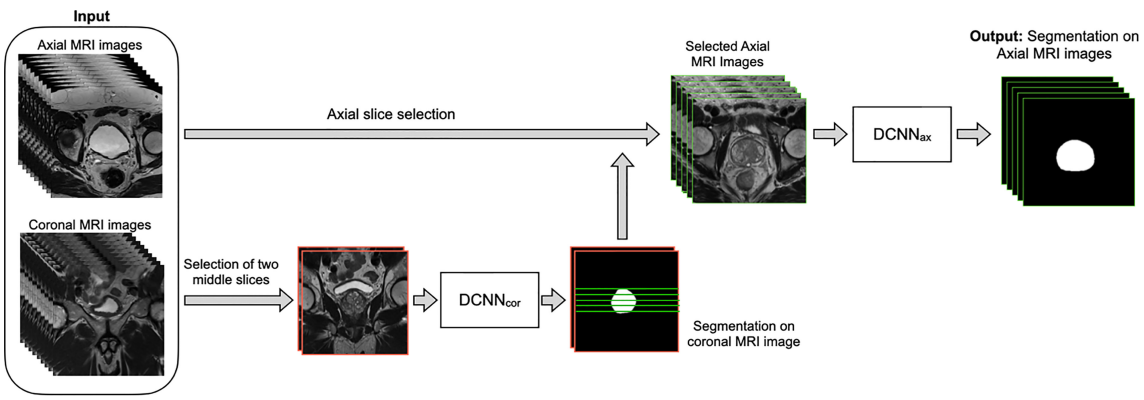


FIGURE 1 | The overall workflow of the automatic WPG segmentation with DANN. Both axial and coronal T2W images were used as input, where the coronal images were used to assist the selection of certain axial images containing the prostate gland. $DANN_{cor}$ was firstly performed on the two middle coronal images, indicated by images with the red border. Next, green lines selected by the prostate segmentation on the coronal images were used to determine the selection of axial slices (images with green borders). Once the axial images were selected, $DANN_{ax}$ was performed on the axial MRI slices for the segmentation of WPG.

TABLE 3 | Description of each visual grade for qualitative segmentation evaluation.

Score	Visual scoring description
3	The segmentation is excellent. The vast majority (>90%) of the prostate region has been correctly segmented, and the percentage of prostate slices with the failure segmentation is less than 10%.
2	The segmentation is acceptable. Most of the region (>70%) is correctly segmented, and the percentage of prostate slices that the method fails to segment is less than 30%.
1	The segmentation is unacceptable. More than 30% of the prostate region has been not correctly segmented or wrongly segmented, or the percentage of prostate slices that the method fails to segment is larger than 30%.

investigate the segmentation at sub-portions of the prostate, we performed the sub-analysis for MRI scans without excellent segmentation performance agreed by both radiologists. Also, the segmentation performance for MRI scans with and without endorectal coil (ERC) was compared.

Quantitative Evaluation of Segmentation Performance

3D DSC (16) was used to quantitatively evaluate and compare the segmentation performance in the quantitative evaluation set (n = 100). The manual segmentations (M) were prepared by the

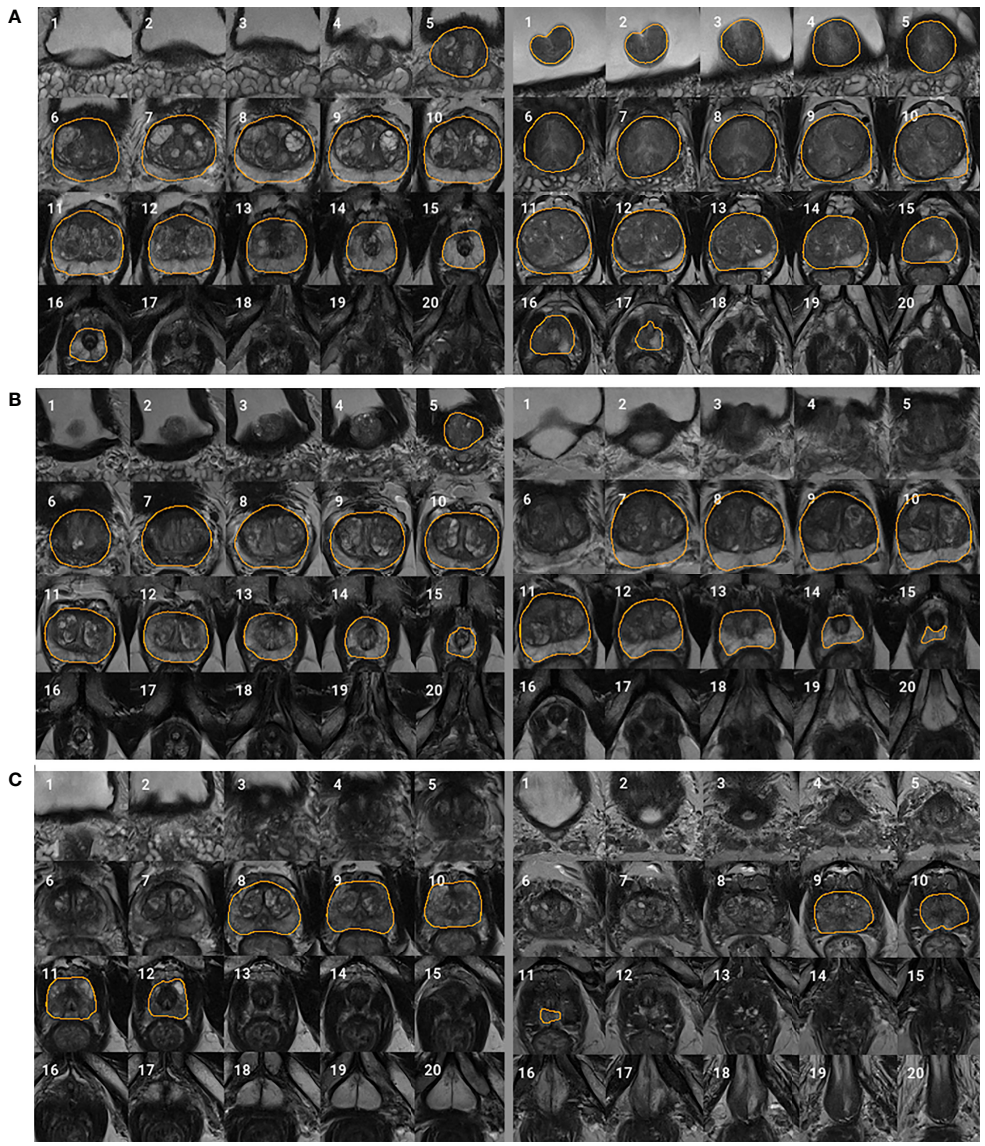


FIGURE 2 | Typical examples for each visual grade. Rows **(A–C)** represent two segmentation examples with visual grades 3 (excellent), 2 (acceptable), and 1 (unacceptable), respectively. Slices 1–20 represent MRI slices from superior to inferior. Regions encircled by organ boundary are the prostate whole gland.

radiologist on all axial slices as ground truths. DSC measures the overlapping between M and method-based (N) segmentation of the WPG volume and can be formulated as:

$$DSC = \frac{2|M \cap M|}{|M| \cup |N|}, \quad (1)$$

where ∩ and ∪ indicate the intersection and union, respectively.

Evaluation of Volume Measurement

We further evaluated the performance of DANN-enabled volume measurements. After the radiologist manually drew the WPG contour on all axial slices, Pyradiomics (17) was used to calculate the prostate volume in the volume measurement evaluation set (n = 50). The prostate volume from the DANN-based segmentation was compared with the manual volume measurement. The Bland–Altman plot (18) was used to analyze the agreement between manual and DANN-enabled WPG volume measurements.

Statistical Analysis

Mean and standard deviation were used to describe the distribution of DSC. The quantitative segmentation performance difference between the DANN and the baselines was compared using a paired sample t-test (19). p values < 0.05 were considered statistically significant. Inter-rater reliability between two radiologists was measured by using the κ statistic (20). The relationship between the value of κ and inter-rater reliability is listed as below, $\kappa < 0$: pool agreement; $0 < \kappa < 0.2$: slight agreement; $0.21 < \kappa < 0.4$: fair agreement; $0.41 < \kappa < 0.6$: moderate agreement; $0.61 < \kappa < 0.8$: substantial agreement; $0.81 < \kappa < 1.0$: almost perfect agreement.

RESULTS

Qualitative Evaluation of WPG Segmentation

Figure 3 shows the proportion of acceptable or excellent segmentation quality in all MRI scans on the qualitative evaluation set at the whole prostate, or each sub-portion (apex, midgland, or base) of the prostate. The DANN method exhibited an acceptable or excellent segmentation performance in more than 96% of the MRI scans on the whole prostate or each sub-portion of the prostate. The segmentation at the midgland portion had achieved the best segmentation performance, while it performed the worst at the base portion.

Qualitative Evaluation and Inter-Rater Variability for WPG Segmentation

For WPG segmentation, 97.9% (n = 3,141) and 93.2% (n = 2,992) of the MRI scans were graded as having acceptable or excellent segmentation performance. **Table 4** includes the confusion matrix to show the inter-rater variability of the visual grading. Overall, two readers reached a substantial consensus on the visual grading in 95.8% of the patients (κ = 0.74). When readers differed on the grading, the discrepancy in grading was less than one. A percentage of 94.6% of segmentation results were unanimously considered as acceptable or excellent. Moreover, 91.5% of the MRI scans (n = 2,861) were graded as having excellent segmentation performance according to the two radiologists. Unacceptable segmentation performance occurred only in 1.2% of the MRI scans (n = 39), agreed by the two radiologists.

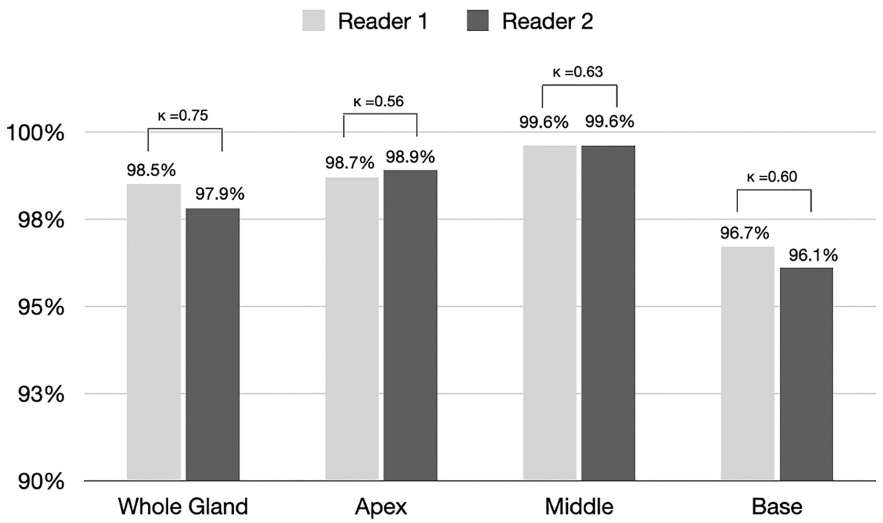


FIGURE 3 | The proportion of segmentation with acceptable or excellent performance evaluated by radiologists 1 and 2 among all MRI scans (n = 3210). Kappa statistics between the two readers were also provided in the figure.

TABLE 4 | Confusion matrices between the visual grades assigned by two readers.

All		Reader 2			Kappa (κ)
Reader 1	Visual grade	1	2	3	Substantial agreement
	1	47 (1.5)	1 (0.0)	0 (0.0)	(κ =0.75)
	2	22 (0.7)	99 (3.1)	49 (1.5)	
	3	0 (0.0)	63 (2.0)	2,929 (91.3)	

Kappa coefficient (κ) is used to measure the inter-rater variability between the two readers.

Sub-Analysis of MRI Scans Without Excellent WPG Segmentation

We conducted the sub-analysis related to each sub-portion of the prostate (apex, midgland, or base) when the WPG segmentation was not excellent. The MRI scans with excellent segmentation agreed by two readers were excluded (n = 2,929), and the rest of the MRI scans were used for the analysis (n = 281). **Figure 4** shows the confusion matrices of each sub-portion of the prostate on the rest of the MRI scans. A percentage of 46.3% of the MRI scans achieved the acceptable (or better) segmentation quality at the base slices, while 94.3% and 83.3% of the MRI scans achieved the acceptable (or better) segmentation quality at the midgland and apex slices.

Comparison Between MRI Scans With and Without ERC

We compared the WPG segmentation quality for the MRI scans with and without ERC (21). **Figure 5** shows the confusion matrices of the visual grades of segmentation on MRI scans with and without ERC. There were substantial agreements (κ = 0.64 and 0.85) between the two radiologists on WPG segmentation of MRI scans with and without ERC. When considering the inter-rater agreement of WPG segmentation, DANN demonstrated acceptable WPG performance in more than 95.5% of MRI scans with ERC compared to 84.3% of those without ERC. MRI scans with ERC had a larger proportion of

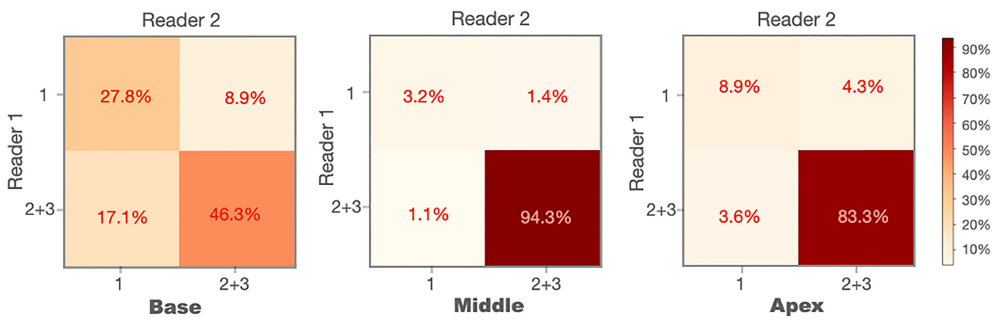


FIGURE 4 | Confusion matrices of the prostate base, midgland, and apex for the cases without excellent segmentation (n = 281).

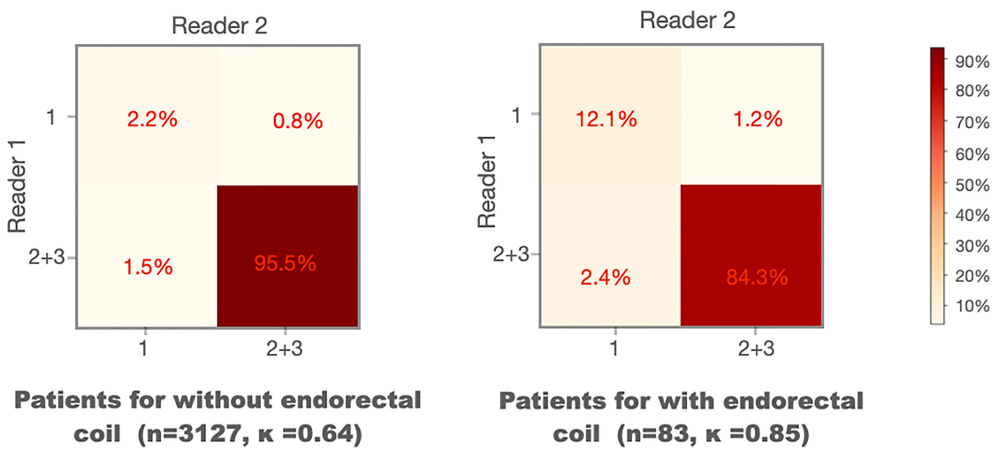


FIGURE 5 | Confusion matrices of the visual grades of segmentation on MRI scans with and without endorectal coils. Kappa coefficient (κ) is used to measure the inter-rater variability between the two readers.

unacceptable WPG segmentation compared to those without ERC (12.1% vs. 2.2%).

Quantitative Evaluation of WPG Segmentation

The quantitative performance of the DANN was compared to the other two baseline methods, including DeepLab v3+ (22) and UNet (23). **Table 5** shows the comparisons of DSCs between DANN and the baseline methods. The DANN achieved a DSC of 0.93, which was higher than those of DeepLab v3+ and UNet with significant differences (both p values < 0.05).

Evaluation of Volume Measurement

Figure 6 shows the agreement between manual and DANN-enabled volume measurements in the Bland–Altman plot. The mean difference between the two-volume measurements was calculated as an estimated bias. Standard deviation (SD) of the differences, and 95% limits of agreement (average difference ± 1.96 SD) were calculated to assess the random fluctuations around this mean. A total of 48 out of 50 cases (96%) had the volume measurement differences within 95% limits of agreement, indicating that the manual and DANN-enabled volume measurements can be potentially used interchangeably.

TABLE 5 | Quantitative DSC comparisons with baseline methods.

Methods	DSC
Proposed method	0.93 ± 0.02
DeepLab v3+	0.92 ± 0.02 p < 0.05
UNet	0.91 ± 0.03 p < 0.05

DISCUSSION

A deep attentive neural network (13), DANN, for the automatic WPG segmentation was evaluated on a large, continuous patient cohort. In the qualitative evaluation, DANN demonstrated that the segmentation quality is either acceptable or excellent in most cases. Two radiologists exhibited a substantial agreement for the qualitative evaluation. In the quantitative evaluation, DANN exhibited a significantly higher DSC than the baseline methods, such as UNet and DeepLab v3+. Also, 96% of the testing cohort had volume measurement differences within 95% limits of agreement.

We found that DANN demonstrated worse segmentation performance at the prostate base than at the apex and midgland slices. This may be due to the fact that the anatomical structure of the prostate base is relatively more complex than other prostate portions. The prostate base is in continuity with the bladder and seminal vesicles, and thus the boundary may contain partial volume effects and mild movement artifacts.

We observed that the segmentation performance was somewhat limited when MRI scans were acquired with an ERC. We believe that this may be because there were only three MRI scans with ERC in the training dataset. A large training data with ERC may allow the model to learn representative features related to the prostate MRI with ERC. In addition, images often exhibit large intensity variation compared to the MRI scans without ERC as ERC is close to the prostate. This may require including an even larger training dataset to account for these intensity variations than those without ERC.

We refined DANN by adding the coronal segmentation to assist the selection of axial slices for WPG segmentation. With assistance from the coronal segmentation, the axial model conducted the segmentation only on the selected axial slices instead of applying it to all axial slices, which reduces the inference time. **Table 6** contains the inference time between the segmentation with and without coronal segmentation. The total inference time in a combination of

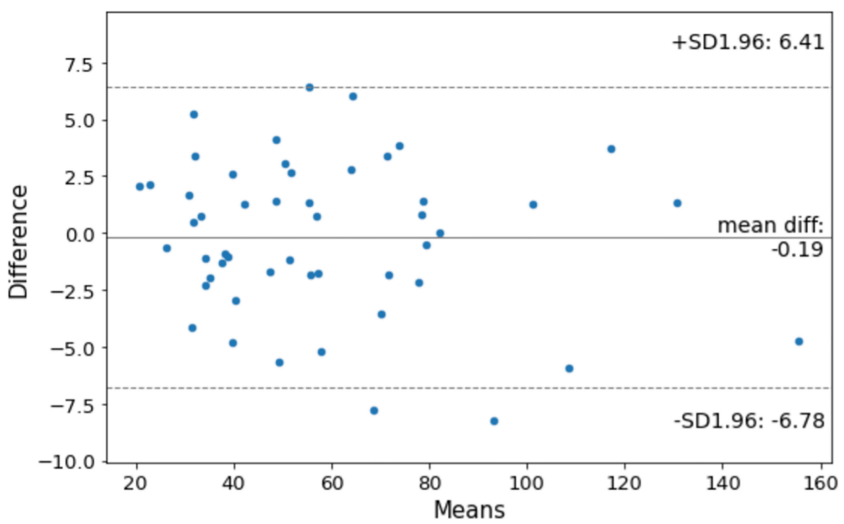


FIGURE 6 | Bland–Altman plot to show the agreement between manual and DANN-enabled WPG volume measurements.

TABLE 6 | Inference time estimation and DSCs obtained with and without coronal segmentation assistance.

	Without coronal segmentation assistance	With coronal segmentation assistance
Overall inference time estimation in the qualitative evaluation	16.4 min (67,775)	12.6 min (45,713)
DSCs obtained in the quantitative evaluation	0.93	0.93

() indicates the total amount of MRI slices the method needed to segment.

coronal and axial slices was 25% less than the inference time without assisting the selection of axial slices (12.6 vs. 16.4 min). In addition, we observed that DSC was not different when adding the coronal segmentation in the quantitative evaluation.

Compared with quantitative evaluation, qualitative evaluation includes unique characteristics and benefits. The DSC-based evaluation often overlooks the segmentation performance on small regions when they were combined with larger regions. Prostate at apex or base slices is relatively smaller than the one in the middle, and therefore, the quantitative evaluation may not be sensitive enough to illustrate limitations at these locations when 3D DSC is used for the evaluation. Also, the DSC-based evaluation is not directly associated with clinical implications, while qualitative evaluation categorized the segmentation results based on the quality to which segmentation can be acceptable clinically.

Our study still has a few limitations: 1) the MRI scans in this study were acquired from three 3T MRI scanners at a single medical center. Prostate MRI sequence parameters are generally well-standardized by the Prostate Imaging–Reporting and Data System (PI-RADS) guidelines (24), but future studies would include testing DANN at multiple institutions. 2) The inter-rater variability was tested between two radiologists. We will include more radiologists to evaluate comprehensive inter-rater variability. 3) Large GPU memory was required during the training and testing since DANN included the spatial attention mechanism that caused considerable computational complexity. In the future, we will explore the criss-cross attention module (25) that uses the contextual information of all the pixels on the criss-cross path for each pixel, which has shown the potential to reduce the GPU memory.

CONCLUSION

Our study showed that the proposed deep learning-based prostate segmentation (DANN) could generate segmentation of the prostate with sufficient quality in a consistent manner when a large, continuous cohort of prostate MRI scans was used for evaluation. The qualitative evaluation conducted by two abdominal radiologists showed that 95% of the segmentation results were either acceptable or excellent with a great inter-rater agreement. In the quantitative evaluation, DANN was superior to the state-of-art deep learning methods, and the difference between manual and DANN-enabled volume measurements was subtle in most cases. The method has a great potential to serve as a tool to assist prostate volume measurements, and biopsy and surgical planning in a clinically relevant setting.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this study will be made available by the authors in accordance with institutional management and sharing policy.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of UCLA. The ethics committee waived the requirement of written informed consent for participation.

AUTHOR CONTRIBUTIONS

Conceptualization, YL and KS. Methodology, YL. Software, YL. Validation, QM and CS. Formal analysis, YL and KS. Investigation, YL and KS. Resources, KS and SR. Data curation, YL. Writing—original draft preparation, YL and KS. Writing—review and editing, YL, KS, SR, GY, QM, CS, and DN. Visualization, YL. Supervision, KS and SR. Project administration, KS. Funding acquisition, KS and GY. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported in part by the National Institutes of Health R01-CA248506 and funds from the Integrated Diagnostics Program, Departments of Radiological Sciences and Pathology, David Geffen School of Medicine, UCLA. This study was also supported in part by the British Heart Foundation (Project Number: TG/18/5/34111, PG/16/78/32402), the European Research Council Innovative Medicines Initiative (DRAGON, H2020-JTI-IMI2 101005122), the AI for Health Imaging Award (CHAIMELEON, H2020-SC1-FA-DTS-2019-1 952172), and the UK Research and Innovation Future Leaders Fellowship (MR/V023799/1).

REFERENCES

1. Garvey B, Türkbey B, Truong H, Bernardo M, Periaswamy S, Choyke PL. Clinical Value of Prostate Segmentation and Volume Determination on MRI in Benign Prostatic Hyperplasia. *Diagn Interv Radiol* (2014) 20:229. doi: 10.5152/dir.2014.13322
2. Ahmed HU, Bosaily AE-S, Brown LC, Gabe R, Kaplan R, Parmar MK, et al. Diagnostic Accuracy of Multi-Parametric MRI and TRUS Biopsy in Prostate

Cancer (PROMIS): A Paired Validating Confirmatory Study. *Lancet* (2017) 389:815–22. doi: 10.1016/S0140-6736(16)32401-1
3. Kasivisvanathan V, Rannikko AS, Borghi M, Panebianco V, Mynderse LA, Vaarala MH, et al. MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis. *N Engl J Med* (2018) 378:1767–77. doi: 10.1056/NEJMoa1801993
4. Oelke M, Bachmann A, Descazeaud A, Emberton M, Gravas S, Michel MC, et al. EAU Guidelines on the Treatment and Follow-Up of Non-Neurogenic Male Lower Urinary Tract Symptoms Including Benign Prostatic Obstruction. *Eur Urol* (2013) 64:118–40. doi: 10.1016/j.eururo.2013.03.004
5. Wenger E, Mårtensson J, Noack H, Bodammer NC, Kühn S, Schaefer S, et al. Comparing Manual and Automatic Segmentation of Hippocampal Volumes: Reliability and Validity Issues in Younger and Older Brains. *Hum Brain Mapp* (2014) 35:4236–48. doi: 10.1002/hbm.22473
6. Yuan Y, Li B, Meng MQ-H. Bleeding Frame and Region Detection in the Wireless Capsule Endoscopy Video. *IEEE J BioMed Heal Inf* (2015) 20:624–30. doi: 10.1109/JBHI.2015.2399502
7. Jin Y, Yang G, Fang Y, Li R, Xu X, Liu Y, et al. 3d PBV-Net: An Automated Prostate MRI Data Segmentation Method. *Comput Biol Med* (2021) 128:104160. doi: 10.1016/j.compbiomed.2020.104160
8. Cheng R, Roth HR, Lay NS, Lu L, Turkbey B, Gandler W, et al. Automatic Magnetic Resonance Prostate Segmentation by Deep Learning With Holistically Nested Networks. *J Med Imaging* (2017) 4:41302. doi: 10.1117/1.JMI.4.4.041302
9. Checcucci E, Autorino R, Cacciamani GE, Amparore D, De Cillis S, Piana A, et al. Artificial Intelligence and Neural Networks in Urology: Current Clinical Applications. *Minerva Urol e Nefrol Ital J Urol Nephrol* (2019) 72:49–57. doi: 10.23736/S0393-2249.19.03613-0
10. Checcucci E, De Cillis S, Granato S, Chang P, Afyouni AS, Okhunov Z. Applications of Neural Networks in Urology: A Systematic Review. *Curr Opin Urol* (2020) 30:788–807. doi: 10.1097/MOU.0000000000000814
11. Zhu Q, Du B, Turkbey B, Choyke PL, Yan P. Deeply-Supervised CNN for Prostate Segmentation. *Int Jt Conf Neural Netw* (2017) 178–84. doi: 10.1109/IJCNN.2017.7965852
12. Jia H, Xia Y, Song Y, Cai W, Fulham M, Feng DD. Atlas Registration and Ensemble Deep Convolutional Neural Network-Based Prostate Segmentation Using Magnetic Resonance Imaging. *Neurocomputing* (2018) 275:1358–69. doi: 10.1016/j.neucom.2017.09.084
13. Liu Y, Yang G, Hosseiny M, Azadikhah A, Mirak SA, Miao Q, et al. Exploring Uncertainty Measures in Bayesian Deep Attentive Neural Networks for Prostate Zonal Segmentation. *IEEE Access* (2020) 8:151817–28. doi: 10.1109/ACCESS.2020.3017168
14. Dice LR. Measures of the Amount of Ecologic Association Between Species. *Ecology* (1945) 26:297–302. doi: 10.2307/1932409
15. Kirişli HA, Schaap M, Klein S, Papadopoulou S-L, Bonardi M, Chen C-H, et al. Evaluation of a Multi-Atlas Based Method for Segmentation of Cardiac CTA Data: A Large-Scale, Multicenter, and Multivendor Study. *Med Phys* (2010) 37:6279–91. doi: 10.1118/1.3512795
16. Liu Y, Yang G, Mirak SA, Hosseiny M, Azadikhah A, Zhong X, et al. Automatic Prostate Zonal Segmentation Using Fully Convolutional Network With Feature Pyramid Attention. *IEEE Access* (2019) 7:163626–32. doi: 10.1109/ACCESS.2019.2952534
17. Kitzing YX, Prando A, Varol C, Karczmar GS, Maclean F, Oto A. Benign Conditions That Mimic Prostate Carcinoma: MR Imaging Features With Histopathologic Correlation. *Radiographics* (2016) 36:162–75. doi: 10.1148/rg.2016150030
18. Giavarina D. Understanding Bland Altman Analysis. *Biochem Med* (2015) 25:141–51. doi: 10.11613/BM.2015.015
19. Semenick D. Tests and Measurements: The T-Test. *Strength \& Cond J* (1990) 12:36–7. doi: 10.1519/0744-0049(1990)012<0036:TTT>2.3.CO;2
20. McHugh ML. Interrater Reliability: The Kappa Statistic. *Biochem Med* (2012) 22:276–82. doi: 10.11613/BM.2012.031
21. Turkbey B, Merino MJ, Gallardo EC, Shah V, Aras O, Bernardo M, et al. Comparison of Endorectal Coil and Nonendorectal Coil T2W and Diffusion-Weighted MRI at 3 Tesla for Localizing Prostate Cancer: Correlation With Whole-Mount Histopathology. *J Magn Reson Imaging* (2014) 39:1443–8. doi: 10.1002/jmri.24317
22. Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-Decoder With Atrous Separable Convolution for Semantic Image Segmentation. In: *Lect. Notes Comput Sci* (2018). pp 833–51. Munich: European Conference on Computer Vision.
23. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Int Conf Med image Comput Comput Interv* (2015) 234–41. doi: 10.1007/978-3-319-24574-4_28
24. Turkbey B, Rosenkrantz AB, Haider MA, Padhani AR, Villeirs G, Macura KJ, et al. Prostate Imaging Reporting and Data System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2. *Eur Urol* (2019) 76:340–51. doi: 10.1016/j.eururo.2019.02.033
25. Huang Z, Wang X, Huang L, Huang C, Wei Y, Liu W. Ccnet: Criss-Cross Attention for Semantic Segmentation. *Proc IEEE/CVF Int Conf Comput Vis* (2019) 603–12. doi: 10.1109/ICCV.2019.00069

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Liu, Miao, Surawech, Zheng, Nguyen, Yang, Raman and Sung. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 26 January 2022
doi: 10.3389/fonc.2021.818329

Low-Dose 68 Ga-PSMA Prostate PET/MRI Imaging Using Deep Learning Based on MRI Priors

Fuquan Deng[1,2,3†], Xiaoyuan Li[4†], Fengjiao Yang[4], Hongwei Sun[5], Jianmin Yuan[6], Qiang He[6], Weifeng Xu[2], Yongfeng Yang[1,3], Dong Liang[1,3], Xin Liu[1,3], Greta S. P. Mok[7], Hairong Zheng[1,3] and Zhanli Hu[1,3]*

[1] Lauterbur Research Center for Biomedical Imaging, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, [2] Computer Department, North China Electric Power University, Baoding, China, [3] Chinese Academy of Sciences Key Laboratory of Health Informatics, Shenzhen, China, [4] Department of Nuclear Medicine, Nanjing First Hospital, Nanjing Medical University, Nanjing, China, [5] United Imaging Research Institute of Intelligent Imaging, Beijing, China, [6] Central Research Institute, United Imaging Healthcare Group, Shanghai, China, [7] Biomedical Imaging Laboratory (BIG), Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Avenida da Universidade, Macau SAR, China

OPEN ACCESS

Edited by:
Guang Yang,
Imperial College London,
United Kingdom

Reviewed by:
Yi Zhang,
Sichuan University, China
Shouping Zhu,
Xidian University, China
T. Niu,
Georgia Institute of Technology,
United States

***Correspondence:**
Zhanli Hu
zl.hu@siat.ac.cn

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 19 November 2021
Accepted: 27 December 2021
Published: 26 January 2022

Citation:
Deng F, Li X, Yang F, Sun H, Yuan J, He Q, Xu W, Yang Y, Liang D, Liu X, Mok GSP, Zheng H and Hu Z (2022) Low-Dose 68 Ga-PSMA Prostate PET/MRI Imaging Using Deep Learning Based on MRI Priors. Front. Oncol. 11:818329. doi: 10.3389/fonc.2021.818329

Background: 68 Ga-prostate-specific membrane antigen (PSMA) PET/MRI has become an effective imaging method for prostate cancer. The purpose of this study was to use deep learning methods to perform low-dose image restoration on PSMA PET/MRI and to evaluate the effect of synthesis on the images and the medical diagnosis of patients at risk of prostate cancer.

Methods: We reviewed the 68 Ga-PSMA PET/MRI data of 41 patients. The low-dose PET (LDPET) images of these patients were restored to full-dose PET (FDPET) images through a deep learning method based on MRI priors. The synthesized images were evaluated according to quantitative scores from nuclear medicine doctors and multiple imaging indicators, such as peak-signal noise ratio (PSNR), structural similarity (SSIM), normalization mean square error (NMSE), and relative contrast-to-noise ratio (RCNR).

Results: The clinical quantitative scores of the FDPET images synthesized from 25%- and 50%-dose images based on MRI priors were 3.84±0.36 and 4.03±0.17, respectively, which were higher than the scores of the target images. Correspondingly, the PSNR, SSIM, NMSE, and RCNR values of the FDPET images synthesized from 50%-dose PET images based on MRI priors were 39.88±3.83, 0.896±0.092, 0.012±0.007, and 0.996±0.080, respectively.

Conclusion: According to a combination of quantitative scores from nuclear medicine doctors and evaluations with multiple image indicators, the synthesis of FDPET images based on MRI priors using and 50%-dose PET images did not affect the clinical diagnosis of prostate cancer. Prostate cancer patients can undergo 68 Ga-PSMA prostate PET/MRI scans with radiation doses reduced by up to 50% through the use of deep learning methods to synthesize FDPET images.

Keywords: PET/MRI, prostate, low-dose restoration, deep learning, discrete wavelet transform

INTRODUCTION

Prostate cancer is one of the most common cancers worldwide, with approximately 1.41 million new cases reported in 2020, the third most common among 36 cancers (1). In recent years, studies have demonstrated that 68 Ga-prostate-specific membrane antigen (PSMA) PET/MRI provides accurate staging of primary prostate cancer with a high detection rate (2–6). In terms of evaluating recurrent prostate cancer, this imaging technique also has a high detection rate, even for patients with extremely low levels of prostate specific antigen (PSA; <0.5 ng/ml). Additionally, it plays an important role in tumor detection, preliminary staging, treatment response assessment, and treatment planning (7, 8).

However, this technique also has some limitations, including scanning time, the cost of the associated radiopharmaceuticals, and the radiation delivered by PET imaging (9). The economic factors and radiation risks have different kinds of impact on the patients. The purpose of reducing the dose of radiopharmaceuticals is related to the potential risks of ionizing radiation. To reduce the risk of radiation exposure that those involved in the scan may face, especially pediatric patients and volunteers, or when a variety of different tracers are used for follow-up or to monitor treatment response, fewer radiopharmaceuticals should be used to perform PET imaging. The reduction in the number of radiopharmaceuticals will reduce the quality of the PET images, thereby affecting quantitative analysis and clinical diagnosis.

In recent years, deep learning has entered various fields of medical imaging. Hu Chen et al. used learned experts' assessment-based reconstruction network to reconstruct CT directly from sinogram data (10). Maosong Ran et al. proposed a Parallel Dual-Domain Convolutional Neural Network for Compressed Sensing MRI to deal with the k-space and spatial data simultaneously (11) and a parameter-dependent framework to process the CT data with different scanning geometries and dose level in a unified network (12). Wenjun Xia et al. simultaneously leverage the spatial convolution to extract the local pixel-level features from the images and incorporate the graph convolution to analyze the nonlocal topological features in manifold space for low-dose CT reconstruction (13). Chenyu Shen et al. leveraged deep regularization by denoising from a Bayesian perspective to reconstruct PET images from a single corrupted sinogram without any supervised or auxiliary information (14). Researchers have proposed a variety of methods to ensure that the synthesized FDPET images have the same image quality as the clinical diagnostic images (15–17). In particular, deep learning has shown great potential in recovering LDPET images. Yan Wang et al. used a 3D conditional generative adversarial network (GAN) to synthesize FDPET images from head LDPET images (18). Wenzhuo Lu et al. used a fully optimized 3D U-net to effectively reduce the noise in LDPET images from the lungs while minimizing the deviation in the lung nodules (19). Yang Lei et al. proposed using CycleGAN to estimate prostate FDPET images from prostate LDPET images (20). Long Zhou et al. also used CycleGAN, denoising low-dose fluorodeoxyglucose (FDG) PET images and subsequently performing quantitative analysis on the images (21).

Deep learning also has a variety of exploratory research in prostate PET images. Pablo Borrelli et al. used convolutional neural network to detect lymph node metastases by PET/CT predicting prostate cancer-specific survival (22). Eirini Polymeri et al. evaluated a novel three-dimensional deep learning-based technique on PET/CT images for automated assessment of cancer in the prostate gland and its agreement with manual assessment (23). Sangwon Han et al. evaluated the performance of deep learning classifers for bone scans of prostate cancer patients (24). Dejan Kostyszyn et al. examined the capabilities of convolutional neural network for intraprostatic GTV contouring in 68Ga- and 18F-PSMA-PET (25). Sobhan Moazem et al. used UNet to predict treatment response in prostate cancer patients based on multimodal PET/CT for clinical decision support (26). Andrii Pozaruk et al. developed a novel augmented deep learning method based on GANs for accurate attenuation correction in the simultaneous PET/MR scanner (27).

In this study, we retrospectively analyzed the 68 Ga-PSMA PET/MRI data of 41 patients in Nanjing First Hospital of China. The PET images were reconstructed at acquisition times of 2.5%, 5%, 25%, 50%, and 100% of the standard acquisition time. A discrete-wavelet-transform convolutional neural network (DWTN) was used to restore the LDPET images to the original, FDPET images with or without the use of MRI priors, respectively, to explore the extent to which this method can reduce the required radiotracer dose.

MATERIALS AND METHODS

PET/MRI Data Acquisition

In this study, we used clinical images obtained from 68 Ga-PSMA PET/MRI examinations performed at Nanjing First Hospital of China from January 2021 to July 2021. The data were obtained from 41 male patients who might present with signs of future prostate cancer. The mean age of the patients was 67 ± 6 years, and the mean weight was 73 ± 10 kg. The research protocol was approved by the institutional ethics committee, and all patients were provided written informed content. Sixty minutes after the patient had been injected with 68 Ga-PSMA (in the range of 111-185×10^6 MBq), scanning data were collected from the PET/MRI scanner (United Imaging Healthcare, uPMR 790). The acquisition time of emission images was 600 seconds, and the PET images were reconstructed at acquisition times of 600, 300, 150, 30, and 15 seconds.

All PET images were reconstructed using ordered subset expectation maximization (OSEM) algorithm and a set series of parameters, for example, 3D iterative time-of-flight (TOF) and point-spread function (PSF) reconstruction, 2 iterations, 20 subsets, matrix 192×192, slice thickness 2.5 mm, and correction methods such as decay correction, attenuation correction, scatter correction, dead time correction, random correction, and detector normalization correction. The images reconstructed at the acquisition times above correspond to 100%-, 50%-, 25%-, 5%-,

and 2.5%-dose (low-dose) PET images, respectively. The 100% does PET images were used as ground-truth and the remaining dose PET images were used as input images. The water sequence decomposed from the T1-weighted MRI images was used as the prior images for generating the LDPET images.

Since integrated prostate PET/MRI was used for scanning, the PET and MRI scans were coaxial. The PET image matrix size is 192×192, and the MRI image matrix size is 552×387. To ensure that the image resolution was not lower than that of the original image, we used bicubic interpolation to resize the two modal images as a 512×512 matrix. Since the image matrix sizes of the two modalities are now the same, the images of the two modalities do not need to be registered. In total, the 41 patients had 4100 sets of images. These sets are divided into training data set and test data set. We used 90% of the data set to train the model and the remaining 10% to verify the effect of the images generated by the model. To avoid overfitting due to the small size of the training data set, we increased the number of images in the training data set by flipping the images down, left and right, quadrupling the size of the training data set. This process helped improve the generalizability of the deep learning models.

Discrete-Wavelet-Transform Convolutional Neural Network

The discrete-wavelet-transform convolutional neural network (DWTN) proposed in this study is an improvement of the densely self-guided wavelet network (28), which is suitable for LDPET image restoration tasks based on MRI priors. The structure of DWTN was shown in **Figure 1**. The multilayer self-guided architecture makes better use of multiscale image information; low-resolution feature information from the top layer is gradually fused with higher-resolution feature information to improve the network's ability to extract multiscale feature information from images. Wavelet transform is used instead of ordinary upsampling and downsampling and PixelShuffle and PixelUnshuffle to generate multiscale image information. Before the convolution process, the image is

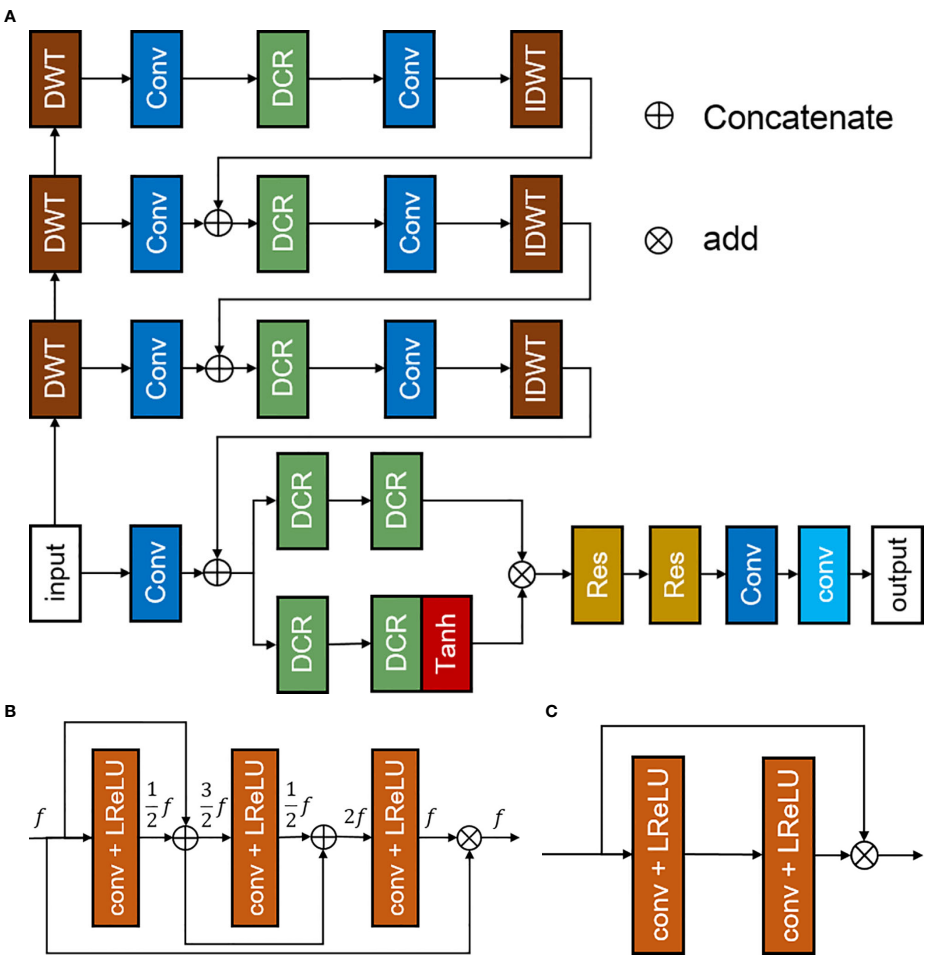


FIGURE 1 | Discrete-wavelet-transform neural network (DWTN), including **(A)** the structure of the DWTN network, **(B)** the structure of the densely connected residual (DCR) block, and **(C)** the structure of the residual block.

converted into horizontal, vertical, and diagonal detail images and thumbnails through discrete wavelet transform (29). At the full-resolution layer, the main branch and attention branch provide stability and process the feature images. At each layer, we add densely connected residual blocks to improve the convergence of the network. The top layer of the DWTN extracts large-scale image feature information in the lowest resolution space. The top-level network contains two convolution layers, a leaky ReLU layer, and a densely connected residual (DCR) block (30). As shown in **Figure 1** (B), the DCR block consists of three convolutional layers, each of which is followed by a leaky ReLU. Each feature image is connected by dense connections so that our model can use the previous feature information to solve the gradient disappearance problem. The middle two layers are similar to the top layer. For the full-resolution level, we add multiple DCR blocks after merging the multiscale feature information to enhance the feature extraction capability of the DWTN. For the attention branch, we add a tanh activation function after the two DCR blocks. The main branch and the attention branch are processed and added, and then through multiple residual blocks and convolution blocks, the added image feature information is extracted. The structure of residual block was shown in **Figure 1** (C). Finally, through a convolution block without an activation function, the details of the image are preserved. The last convolution layer uses a 1×1 convolution kernel with a step count of 1, and the remaining convolution layers all use a 3×3 convolution kernel with a step count of 1.

To improve the generalization of the network, we used perceptual loss, mean-square-error loss and kernel loss to constrain the network (31). Among them, perceptual loss enabled the network to learn the characteristics of the overall images and converge faster than mean-absolute-error loss. The perceptual loss used pre-trained VGG19 for extracting image features. The mean-square-error loss was to calculate the loss function at the pixel level to ensure the network to generate more details of the images. The kernel loss reduced the weight of the hidden layer inside network, thereby promoting the convergence of DWTN. The loss function formula was as follows:

$$\mathrm{Loss}_{total} = mse \times \mathrm{Loss}_{MSE} + vgg \times \mathrm{Loss}_{per} + kl \times \mathrm{Loss}_{ker}, \quad (1)$$

where mse, vgg and kl are 0.5, 0.5 and 0.0001, respectively. The formulas of each loss function are as follows:

$$\mathrm{Loss}_{MSE} = \mathbb{E}\,(\hat{y} - y)^2, \quad (2)$$

$$\mathrm{Loss}_{per} = \mathbb{E}(\mathrm{VGG}(\hat{y}) - \mathrm{VGG}(\mathrm{y}))^2, \quad (3)$$

$$\mathrm{Loss}_{ker} = \mathbb{E}(\hat{y})^2, \quad (4)$$

where $\hat{y}$ and y represent the generated images and ground-truth, respectively. VGG represents the processing of VGG model.

When initializing the training parameters, we used the ADAM optimizer, and the remaining parameters are set to their default values. We update the learning rate every 25 epochs and set the learning rate decay rate to 0.5. The weights of all hidden layers are initialized with Gaussian random numbers. The model was implemented on an NVIDIA GeForce RTX 2080Ti GPU with 11 GB of memory and run under the Microsoft Windows 10 operating system. During training, we used a batch size of 4 for 100 epochs.

Evaluation Metrics

Clinical quantitative evaluation. To evaluate the quality of the PET images, we evaluated the original LDPET images, the synthesized FDPET images, and the synthesized FDPET images based on the MRI priors. There were 13 sets of images, including 32 PET images in each group. The PET images were evaluated using a 5-point method by two nuclear medicine physicians from Nanjing First Hospital of China (32, 33).

Image quantitative analysis. To evaluate the image quality between the synthesized FDPET images and the original FDPET images, we used the PSNR, SSIM, RCNR, and NMSE as objective indicators.

The PSNR is a quantitative index for evaluating images and noise, and the SSIM is an index for evaluating the similarity of two image features; both indices offer comprehensive evaluations of two images.

$$\mathrm{PSNR} = 10 \times \log_{10}\left(\frac{(2^n - 1)^2}{MSE}\right), \quad (5)$$

Where MSE is the mean square error between the compared images and the ground-truth.

$$\mathrm{SSIM} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (6)$$

where μ_x is the mean of compared images and μ_y is the mean of ground-truth, respectively. σ_x and σ_y are variance of compared images and ground-truth, respectively. σ_{xy} is the covariance between compared images and ground-truth. c_1 and c_2 are 0.01 and 0.03, respectively.

The NMSE is an indicator of the quantitative analysis of two images at the pixel level.

$$\mathrm{NMSE} = \frac{\Sigma_{i \in V}(X_i - Y_i)^2}{\Sigma_{i \in V}(Y_i)^2}, \quad (7)$$

Where X_i and Y_i represent the pixel value of the compared image and groud-truth, respectively.

The contrast-to-noise ratio is an objective index used to evaluate the quality of medical images; the RCNR is a dimensionless image index based on CNR that is used to compare the contrast of two images.

$$\mathrm{CNR} = \frac{X - X_{background}}{X} \quad (8)$$

$$\mathrm{RCNR} = \frac{CNR_x}{CNR_y}, \quad (9)$$

Where X and $X_{background}$ represent the mean of the matrix and background matrix, respectively.

RESULTS

Compared with the LDPET images, the FDPET images synthesized by the deep learning method demonstrated significantly improved image quality. PET images with a dose of less than 5% showed irregular spots, and their contours, shapes, and contrast were different from those of the target images. In the images synthesized from those with a dose of less than 5%, the spots were eliminated, and the shape features and contrast were relatively consistent with those of the target images. In PET images with a dose greater than 5%, the contours, shapes, and contrast were similar to those the target images, but subtle differences could be observed. In the synthesized PET images, the shapes, contours, and contrast were consistent with those of the target images.

After training the model, we calculated the mean and standard deviation of the RCNR, PSNR, SSIM, NMSE among the original LDPET original image, the synthesized FDPET images, the prior-synthesized FDPET images, and the target images of all doses in the test set. **Table 1** shows the mean and standard deviation values of the image indicators. To more intuitively visualize the differences in these image indicators, **Figure 2** shows the histograms of the structural similarity and NMSE indicators. The processed LDPET images had significantly better image quality than the original LDPET images. The images based on MRI prior synthesis showed better image quality at the global-feature and pixel levels.

Figure 3 shows one patient's original 2.5%-, 5%-, 25%-, and 50%-dose PET images and their corresponding synthesized FDPET images and prior-synthesized FDPET images for the prostate. The average RCNR of the LDPET images and MRI prior-based synthesized LDPET images was close to 1. Moreover, the pelvic contour details of the prior-based synthesized images for doses of 25% and above are visible.

We transferred the synthesized images and MRI prior synthesized images for each LDPET image to DICOM format, subtracted each image from the original FDPET image matrix, and finally divided the difference matrix by the maximum value of the original image. The resulting image matrixes are shown in **Figure 4**. The error between the 25%-dose synthesized images and the prior-synthesized images are within 25%, and the error between the 50%-dose synthesized images and the prior-synthesized images are within 10%.

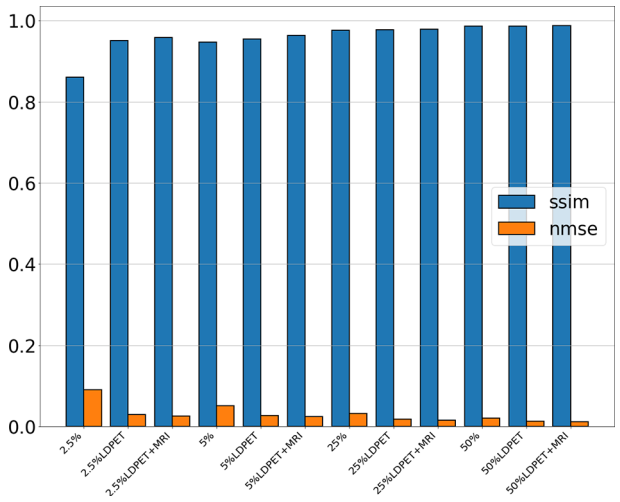


FIGURE 2 | NMSE and SSIM of the original LDPET images, synthetic FDPET images, and MR prior-synthesized FDPET images at all doses, where blue represents SSIM, orange represents NMSE, and different degrees of color represent images of different doses.

Six patients were selected from the test set, all of whom had images of prostate or pelvic lesions, as shown in **Figure 5**. The figure shows the diffusion weighted (DW) image, apparent diffusion coefficient (ADC) image, T2-weighted image, FDPET image, and various synthesized images. The DW image, ADC image, and T2-weighted image in the MRI sequence are important references for clinical diagnosis, and the PET images and MRI images are complementary.

In the clinical quantification phase, we selected 32 images with lesions in the pelvis or prostate from the test set for scoring. **Table 2** shows the mean and variance of the scored from the two nuclear medicine doctors for PET images of different doses, processed PET images, and prior-based PET images. The average score of the FDPET images MRI prior-synthesized from 25%-dose PET images is 0.1 points lower than the average score of the target images. In contrast, the average score of the FDPET images synthesized from 25% dose PET images without the prior is 0.3 points lower than the average score of the target images. In addition, to improve the

TABLE 1 | Objective indicators for LDPET images, synthesized FDPET images, MR prior-synthesized FDPET images, and target images.

Image	PSNR	SSIM	NMSE	RCNR
2.5%	25.56±4.99	0.745±0.112	0.065±0.045	0.835±0.267
2.5%LDPET	32.51±4.89	0.820±0.079	0.029±0.021	1.046±0.213
2.5%LDPET+MRI	33.34±4.47	0.846±0.060	0.026±0.018	1.118±0.218
5%	26.99±5.56	0.756±0.113	0.057±0.037	0.860±0.227
5%LDPET	33.25±4.57	0.814±0.137	0.026±0.018	1.032±0.216
5%LDPET+MRI	33.78±4.25	0.817±0.141	0.024±0.017	0.964±0.248
25%	29.58±7.05	0.832±0.114	0.047±0.036	0.901±0.195
25%LDPET	36.90±4.40	0.893±0.090	0.017±0.012	1.035±0.123
25%LDPET+MRI	37.86±4.16	0.916±0.063	0.015±0.012	1.004±0.126
50%	32.02±8.66	0.865±0.123	0.040±0.036	0.909±0.183
50%LDPET	39.48±3.90	**0.919±0.067**	0.012±0.008	1.009±0.079
50%LDPET+MRI	**39.88±3.83**	0.896±0.092	**0.012±0.007**	**0.996±0.080**

In bold: The best performance in this indicator.

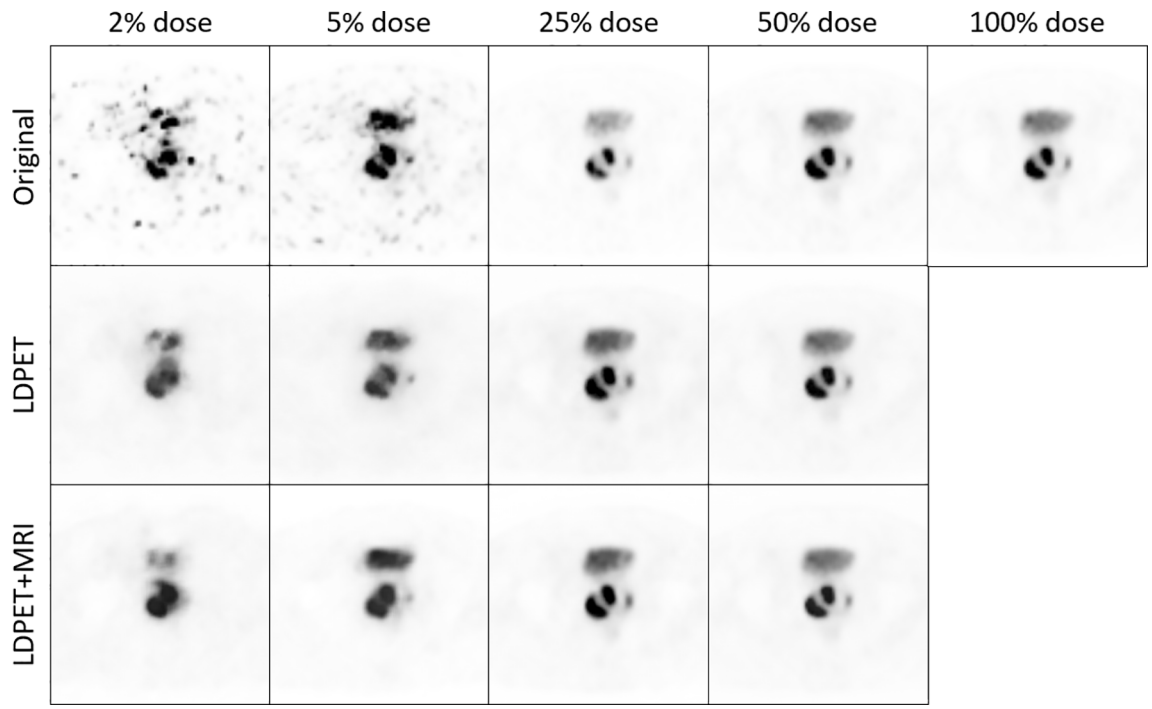


FIGURE 3 | The original LDPET images, synthesized FDPET images, MR prior-synthesized FDPET images of all doses, and their ROIs.

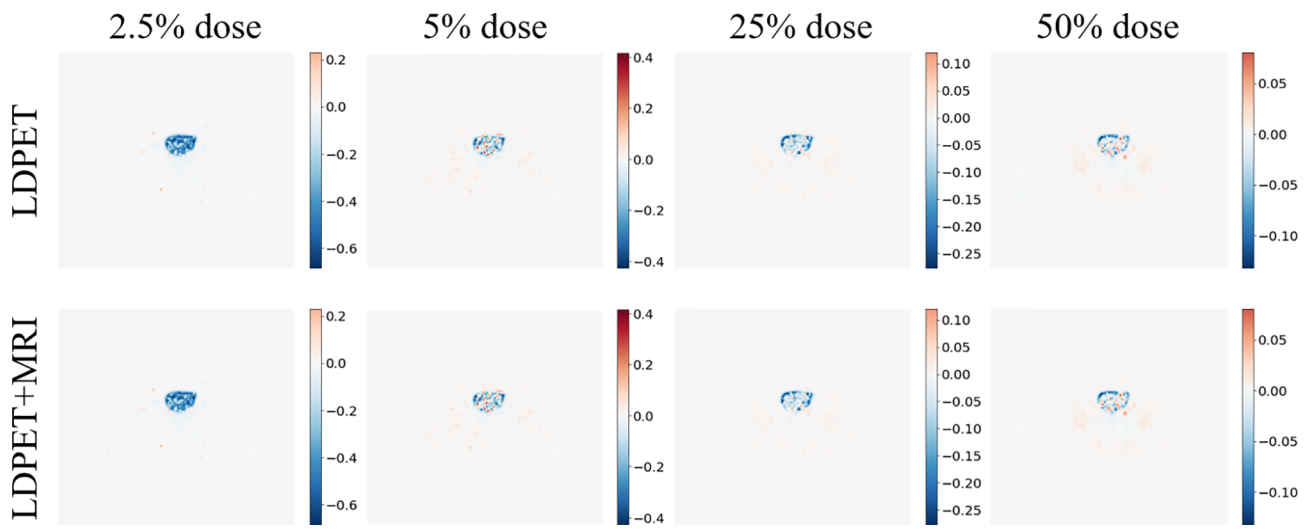


FIGURE 4 | Synthesized FDPET image, MR prior for all doses, combined FDPET image and target subtraction difference map and original image.

credibility of the analysis, we combined the scores of the two doctors. When the two nuclear medicine doctors had different scores for the same image, then the lower score of the two was taken. **Figure 6** shows the distribution of this score. When MRI images were used as the prior, more than 80% of the FDPET images synthesized from 25%-dose PET images had scores of 4, while the rest had scores of 3. When no priors were used, the scores of the FDPET images synthesized from the 25%-dose images indicated that the FDPET images were between good and poor quality. Regardless of whether MRI images were used as priors, the scores of the FDPET image synthesized from the 50%-dose PET images were greater than or equal to 4. At each dose, using MRI images as

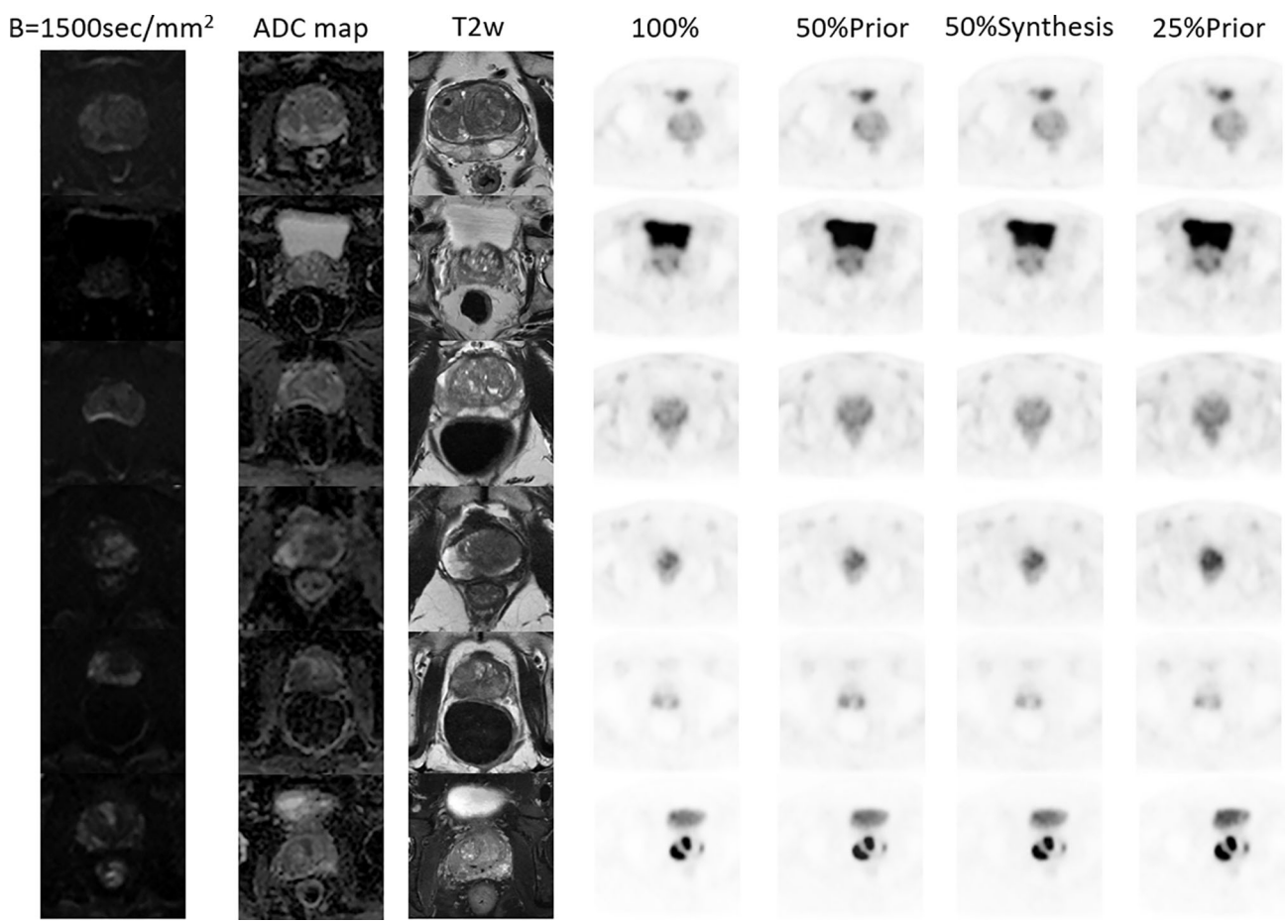


FIGURE 5 | MR and PET images of 6 patients with prostate or pelvic lesions from the test set. The MR sequences included DW-, ADC- and T2-weighted images. The PET images consist of the original FDPET images, FDPET images synthesized from 50%-dose images with MR priors, FDPET images synthesized from 50%-dose images without MR priors, and FDPET images synthesized from 25%-dose images with MR priors.

TABLE 2 | Mean clinical quantitative scores from nuclear medicine doctors on LDPET images, synthesized FDPET images, MR prior-synthesized FDPET images, and target images.

	Original	LDPET	LDPET+MRI
2.5%	1.00±0.00	2.44±0.56	2.69±0.46
5%	1.69±0.46	2.94±0.43	3.28±0.57
25%	2.47±0.50	3.62±0.48	3.84±0.36
50%	3.03±0.39	4.03±0.17	4.03±0.17
100%	3.94±0.24	–	–

priors in synthesizing the images was better than using single LDPET images to synthesize FDPET images.

DISCUSSION

In the LDPET image denoising method, MRI images were used as prior image to provide rich tissue and anatomical information for PET image synthesis, thereby improving image quality, contours and details. However, the use of MRI as priors can also lead to registration problems. Due to the characteristics of the MRI images and the PSMA PET images, accurate registration, regardless of whether rigid registration or flexible registration is used, can be difficult. Data registration is not a problem, however, because the integrated PET/MRI device uses coaxial scanning, and the data it collects can be directly applied to LDPET restoration. Theoretically, data collected by integrated PET/MRI are more suitable for MRI prior-based LDPET estimation than data collected by sequential PET/MRI, insert PET/MRI, or multiple devices.

When using a single LDPET image for denoising, synthesized prostate FDPET images have poor overall contour and edge details and cannot be restored well. When including MRI images as priors for LDPET image denoising, the edges of PET images with a dose of less than 5% is blurred, and the surrounding contours are not clear, which can affect the diagnosis. In PET images with a dose higher than 5%, the shape and edges of the key parts of the prostate are clear, the contours of the surrounding organs are distinct, the contrast is relatively close to that of the target image, and the clinical quantitative score is the same as that of the target

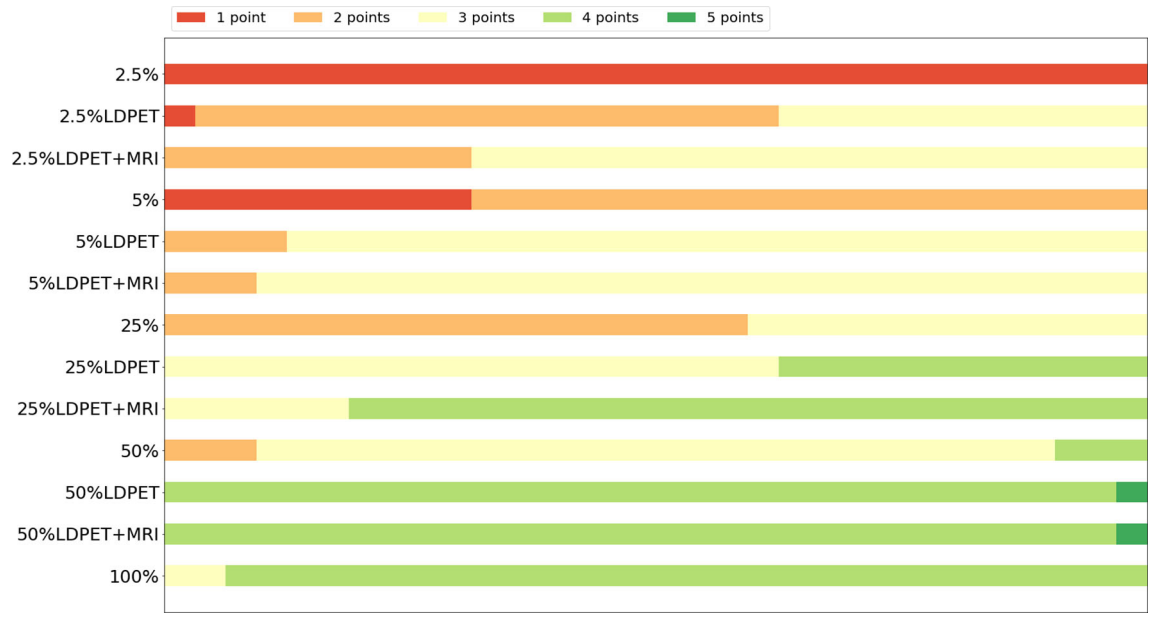


FIGURE 6 | Distribution of clinical quantitative scores from nuclear medicine doctors on LDPET images, synthesized FDPET images, MR prior-synthesized FDPET images, and target images.

image. However, **Figure 3** shows that the 25%-dose PET synthesized images based on the MRI priors and the ground-truth have a large error in the local area of the prostate. Thus, the 50%-dose PET images synthesized based on the MRI priors showed sufficient quality to meet the requirements for clinical analysis.

However, our method still has certain limitations. First, the proposed network uses the Haar wavelet transform for up- and downsampling. Other potential wavelet transforms include Gaussian, Morlet, Shanno, and other transformation methods, and whether the Haar wavelet transform is the most suitable for MRI prior-based LDPET estimation is not yet known. Second, the proposed network has a significant effect on LDPET images with good overall characteristics but has a poor effect on LDPET images with inconspicuous overall characteristics. The recovery effect of the convolutional neural network on PET images with a dose of 5% and below needs to be improved; for example, the use of generative adversarial neural networks for image recovery and PET/MRI examinations could reduce the dose even further.

CONCLUSION

In conclusion, we used a convolutional neural network combining discrete wavelet transform and convolution methods to estimate FDPET images from PSMA LDPET image collected by PET/MRI. After clinical quantitative analysis and objective image index analysis, the deep learning method we proposed was shown to be capable of synthesizing a FDPET image from 50%-dose PSMA PET images collected by PET/MRI, indicating that the dose can be reduced by 50%.

DATA AVAILABILITY STATEMENT

Further inquiries can be directed to the corresponding author. Requests to access the datasets should be directed to Zhanli Hu, zl.hu@siat.ac.cn.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by (Permit S-152/2020) Department of Nuclear Medicine, Nanjing First Hospital, Nanjing Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

FD was responsible for programming and revising the manuscript. X-yL and FY performed data acquisition and analyzed the data. HS, JY, QH, YY, DL, XL, GM, HZ and WX made a substantial contribution to the study and helped revise the manuscript critically for important intellectual content. ZH made substantial contributions to the conception and design of the study. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (32022042, 81871441), the Shenzhen Excellent Technological Innovation Talent Training Project of

China (RCJC20200714114436080), the Natural Science Foundation of Guangdong Province in China (2020A1515010733), Chinese Academy of Sciences Key Laboratory of Health Informatics in China (2011DP173015).

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
2. Eiber M, Weirich G, Holzapfel K, Souvatzoglou M, Haller B, Rauscher I, et al. Simultaneous 68ga-PSMA HBED-CC PET/MRI Improves the Localization of Primary Prostate Cancer. *Eur Urol* (2016) 70(5):829–36. doi: 10.1016/j.eururo.2015.12.053
3. Hicks RM, Simko JP, Westphalen AC, Nguyen HG, Greene KL, Zhang L, et al. Diagnostic Accuracy of 68Ga-PSMA-11 PET/MRI Compared With Multiparametric MRI in the Detection of Prostate Cancer. *Radiology* (2018) 289(3):730–7. doi: 10.1148/radiol.2018180788
4. Park SY, Zacharias C, Harrison C, Fan RE, Kunder C, Hatami N, et al. Gallium 68 PSMA-11 PET/MR Imaging in Patients With Intermediate- or High-Risk Prostate Cancer. *Radiology* (2018) 288(2):495–505. doi: 10.1148/radiol.2018172232
5. Murthy V, Sonni I, Jariwala N, Juarez R, Reiter RE, Raman SS, et al. The Role of PSMA PET/CT and PET/MRI in the Initial Staging of Prostate Cancer. *Eur Urol Focus* (2021) 7(2):258–66. doi: 10.1016/j.euf.2021.01.016
6. Grubmüller B, Baltzer P, Hartenbach S, D'Andrea D, Helbich TH, Haug AR, et al. PSMA Ligand PET/MRI for Primary Prostate Cancer: Staging Performance and Clinical Impact. *Clin Cancer Res* (2018) 24(24):6300. doi: 10.1158/1078-0432.CCR-18-0768
7. Kranzbühler B, Nagel H, Becker AS, Müller J, Huellner M, Stolzmann P, et al. Clinical Performance of 68Ga-PSMA-11 PET/MRI for the Detection of Recurrent Prostate Cancer Following Radical Prostatectomy. *Eur J Nucl Med Mol Imaging* (2018) 45(1):20–30. doi: 10.1007/s00259-017-3850-x
8. Wang R, Shen G, Yang R, Ma X, Tian R. 68ga-PSMA PET/MRI for the Diagnosis of Primary and Biochemically Recurrent Prostate Cancer: A Meta-Analysis. *Eur J Radiol* (2020) 130:109131. doi: 10.1016/j.ejrad.2020.109131
9. Vandenberghe S, Marsden PK. PET-MRI: A Review of Challenges and Solutions in the Development of Integrated Multimodality Imaging. *Phys Med Biol* (2015) 60(4):R115. doi: 10.1088/0031-9155/60/4/R115
10. Chen H, Zhang Y, Chen Y, Zhang J, Zhang W, Sun H, et al. LEARN: Learned Experts' Assessment-Based Reconstruction Network for Sparse-Data CT. *IEEE Trans Med Imaging* (2018) 37(6):1333–47. doi: 10.1109/TMI.2018.2805692
11. Ran M, Xia W, Huang Y, Lu Z, Bao P, Liu Y, et al. Md-Recon-Net: A Parallel Dual-Domain Convolutional Neural Network for Compressed Sensing Mri. *IEEE Trans Radiat Plasma Med Sci* (2020) 5(1):120–35.
12. Xia W, Lu Z, Huang Y, Liu Y, Chen H, Zhou J, et al. CT Reconstruction With PDF: Parameter-Dependent Framework for Data From Multiple Geometries and Dose Levels. *IEEE Trans Med Imaging* (2021) 40(11):3065–76. doi: 10.1109/ISBI48211.2021.9433906
13. Xia W, Lu Z, Huang Y, Shi Z, Liu Y, Chen H, et al. MAGIC: Manifold and Graph Integrative Convolutional Network for Low-Dose CT Reconstruction. *IEEE Trans Med Imaging* (2021) 40(12):3459–72. doi: 10.1109/TMI.2021.3088344
14. Shen C, Xia W, Ye H, Hou M, Chen H, Liu Y, et al. Unsupervised PET Reconstruction From a Bayesian Perspective. *arXiv preprint arXiv* (2021), 2110.15568.
15. An L, Zhang P, Adeli E, Wang Y, Ma G, Shi F, et al. Multi-Level Canonical Correlation Analysis for Standard-Dose PET Image Estimation. *IEEE Trans Image Process* (2016) 25(7):3303–15. doi: 10.1109/TIP.2016.2567072
16. Wang Y, Zhang P, An L, Ma G, Kang J, Shi F, et al. Predicting Standard-Dose PET Image From Low-Dose PET and Multimodal MR Images Using Mapping-Based Sparse Representation. *Phys Med Biol* (2016) 61(2):791–812. doi: 10.1088/0031-9155/61/2/791
17. Xiang L, Qiao Y, Nie D, An L, Lin W, Wang Q, et al. Deep Auto-Context Convolutional Neural Networks for Standard-Dose PET Image Estimation From Low-Dose PET/MRI. *Neurocomputing* (2017) 267:406–16. doi: 10.1016/j.neucom.2017.06.048
18. Wang Y, Yu B, Wang L, Zu C, Lalush DS, Lin W, et al. 3D Conditional Generative Adversarial Networks for High-Quality PET Image Estimation at Low Dose. *NeuroImage* (2018) 174:550–62. doi: 10.1016/j.neuroimage.2018.03.045
19. Cui J, Gong K, Guo N, Wu C, Meng X, Kim K, et al. PET Image Denoising Using Unsupervised Deep Learning. *Eur J Nucl Med Mol Imaging* (2019) 46:2780–9. doi: 10.1007/s00259-019-04468-4
20. Yang L, Xue D, Tonghe W, Kristin H, Tian L, Walter JC, et al. Estimating Standard-Dose PET From Low-Dose PET With Deep Learning. *Proc.SPIE* (2020) 11313. doi: 10.1117/12.2548461
21. Zhou L, Schaefferkoetter JD, Tham IWK, Huang G, Yan J. Supervised Learning With Cyclegan for Low-Dose FDG PET Image Denoising. *Med Image Anal* (2020) 65:101770. doi: 10.1016/j.media.2020.101770
22. Borrelli P, Larsson M, Ulén J, Enqvist O, Trägårdh E, Poulsen MH, et al. Artificial Intelligence-Based Detection of Lymph Node Metastases by PET/CT Predicts Prostate Cancer-Specific Survival. *Clin Physiol Funct Imaging* (2021) 41(1):62–7. doi: 10.1111/cpf.12666
23. Polymeri E, Sadik M, Kaboteh R, Borrelli P, Enqvist O, Ulén J, et al. Deep Learning-Based Quantification of PET/CT Prostate Gland Uptake: Association With Overall Survival. *Clin Physiol Funct Imaging* (2020) 40 (2):106–13. doi: 10.1111/cpf.12611
24. Han S, Oh JS, Lee JJ. Diagnostic Performance of Deep Learning Models for Detecting Bone Metastasis on Whole-Body Bone Scan in Prostate Cancer. *Eur J Nucl Med Mol Imaging* (2021). doi: 10.1007/s00259-021-05481-2
25. Kostyszyn D, Fechter T, Bartl N, Grosu AL, Gratzke C, Sigle A, et al. Intraprostatic Tumor Segmentation on PSMA PET Images in Patients With Primary Prostate Cancer With a Convolutional Neural Network. *J Nucl Med* (2021) 62(6):823. doi: 10.2967/jnumed.120.254623
26. Moazemi S, Essler M, Schultz T, Bundschuh RA, et al. Predicting Treatment Response in Prostate Cancer Patients Based on Multimodal PET/CT for Clinical Decision Support. In: T Cham, Syeda-Mahmood, editors. *Multimodal Learning for Clinical Decision Support*. Springer International Publishing (2021). p. 22–35.
27. Pozaruk A, Pawar K, Li S, Carey A, Cheng J, Sudarshan V, et al. Augmented Deep Learning Model for Improved Quantitative Accuracy of MR-Based PET Attenuation Correction in PSMA PET-MRI Prostate Imaging. *Eur J Nucl Med Mol Imaging* (2021) 48(1):9–20. doi: 10.1007/s00259-020-04816-9
28. Liu W, Yan Q, Zhao Y. Densely Self-Guided Wavelet Network for Image Denoising. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (2020). p. 432–3.
29. G R, Lee GR, Wasilewski F, Wohlfahrt K, O'Leary A. PyWavelets: A Python Package for Wavelet Analysis. *J Open Source Software* (2019) 4:1237. doi: 10.21105/joss.01237
30. Park B, Yu S, Jeong J. Densely Connected Hierarchical Network for Image Denoising. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (2019).
31. Johnson J, Alahi A, Fei-Fei L. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In: CB Leibe, J Matas, N Sebe, M Welling, editors. *Computer Vision – ECCV 2016*. Springer International Publishing (2016). p. 694–711.
32. Ouyang J, Chen KT, Gong E, Pauly J, Zaharchuk G. Ultra-Low-Dose PET Reconstruction Using Generative Adversarial Network With Feature Matching and Task-Specific Perceptual Loss. *Med Phys* (2019) 46(8):3555–64. doi: 10.1002/mp.13626
33. Chen KT, Gong E, de Carvalho Macruz F, Xu B, Boumis J, Khalighi A, et al. Ultra–Low-Dose 18f-Florbetaben Amyloid PET Imaging Using Deep Learning With Multi-Contrast MRI Inputs. *Radiology* (2018) 290(3):649–56. doi: 10.1148/radiol.2018180940

Conflict of Interest: HS, JY, and QH are employees of the United Imaging Healthcare group.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Deng, Li, Yang, Sun, Yuan, He, Xu, Yang, Liang, Liu, Mok, Zheng and Hu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 27 January 2022
doi: 10.3389/fonc.2021.820102

Feasibility of Shear Wave Elastography Imaging for Evaluating the Biological Behavior of Breast Cancer

Chaoxu Liu[†], *Jin Zhou*[†], *Cai Chang*[*] *and Wenxiang Zhi*[*]

Department of Ultrasonography, Fudan University Shanghai Cancer Center, Shanghai Medical College, Fudan University, Shanghai, China

OPEN ACCESS

Edited by:
Guang Yang,
Imperial College London,
United Kingdom

Reviewed by:
Rong Wu,
Shanghai General Hospital, China
Xin-Wu Cui,
Huazhong University of Science and Technology, China

***Correspondence:**
Wenxiang Zhi
zwenx1123@163.com
Cai Chang
changc61@163.com

[†]*These authors have contributed equally to this work and share first authorship*

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *22 November 2021*
Accepted: *30 December 2021*
Published: *27 January 2022*

Citation:
Liu C, Zhou J, Chang C and Zhi W (2022) Feasibility of Shear Wave Elastography Imaging for Evaluating the Biological Behavior of Breast Cancer. Front. Oncol. 11:820102. doi: 10.3389/fonc.2021.820102

Objective: To explore the feasibility of shear wave elastography (SWE) parameters for assessing the biological behavior of breast cancer.

Materials and Methods: In this prospective study, 224 breast cancer lesions in 216 female patients were examined by B-mode ultrasound and shear wave elastography in sequence. The maximum size (S_{max}) of the lesion was measured by B-mode ultrasound, and then shear wave elastography was performed on this section to obtain relevant parameters, including maximum elasticity (E_{max}), mean elasticity (E_{mean}), standard deviation of elasticity (SD), and the area ratio of shear wave elastography to B-mode ultrasound (AR). The relationship between SWE parameters and pathological type, histopathological classification, histological grade, lymphovascular invasion status (LVI), axillary lymph node status (ALN), and immunohistochemistry of breast cancer lesions was performed according to postoperative pathology.

Results: In the univariate analysis, the pathological type and histopathological classification of breast cancer were not significantly associated with SWE parameters; with an increase in the histological grade of invasive ductal carcinoma (IDC), SD ($p = 0.016$) and S_{max} ($p = 0.000$) values increased. In the ALN-positive group, S_{max} ($p = 0.004$) was significantly greater than in the ALN-negative group; S_{max} ($p = 0.003$), E_{max} ($p = 0.034$), and SD ($p = 0.045$) were significantly higher in the LVI-positive group than in the LVI-negative group; SD ($p = 0.043$, $p = 0.047$) and S_{max} ($p = 0.000$, $p = 0.000$) were significantly lower in the ER^+ and PR^+ groups than in the ER^- and PR^- groups, respectively; AR ($p = 0.032$) was significantly higher in the ER^+ groups than in the ER^- groups, and S_{max} ($p = 0.002$) of the $HER2^+$ group showed higher values than that of the $HER2^-$ group; S_{max} ($p = 0.000$), SD ($p = 0.006$), and E_{max} ($p = 0.004$) of the Ki-67 high-expression group showed significantly higher values than those of the Ki-67 low-expression group. In the multivariate analysis, Ki-67 was an independent factor of S_{max} ($p = 0.005$), E_{max} ($p = 0.004$), and SD ($p = 0.006$); ER was an independent influencing factor of S_{max} ($p = 0.000$) and AR ($p = 0.032$). LVI independently influences S_{max} ($p = 0.006$).

Conclusions: The SWE parameters E_{max}, SD, and AR can be used to evaluate the biological behavior of breast cancer.

Keywords: ultrasound, shear wave elastography, breast cancer, prognosis, biological behavior

INTRODUCTION

Breast cancer is one of the most common cancers, and the mortality rate of breast cancer ranks first among women (1). Predicting the prognosis of breast cancer preoperatively can aid in understanding the disease course and guiding treatment. The pathological type, histopathological classification, histological grade, axillary lymph node status (ALN), lymphovascular invasion status (LVI), and immunohistochemical factors such as estrogen receptor (ER) and progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and the Ki-67 proliferation index affect the prognosis of breast cancer (2–5), and these methods of evaluating the breast cancer prognosis are obtained invasively *via* biopsy. Therefore, noninvasive methods to assess the prognosis of breast cancer are urgently needed. Shear wave elastography (SWE) imaging is a recently emerging elastography technology, which can qualitatively and quantitatively differentiate between malignant and benign foci, and its clinical value has been widely recognized (6–9). As quantitative parameters of SWE, the maximum elasticity (E_{max}), standard deviation of elasticity (SD), and mean elasticity (E_{mean}) can quantitatively reflect the tissue hardness of the lesion, and the relationship between these parameters and patient clinicopathological characteristics has attracted considerable clinical attention (10). Determining whether these parameters can be used to quantitatively evaluate the biological behavior of breast cancer has been a major research focus in recent years (11, 12).

Kim's research results showed that tumors with higher E_{max} and E_{mean} values were found to exhibit poor pathological differentiation (13). Huang et al. concluded that SD can reflect the heterogeneity of tumors and can be used to distinguish benign from malignant lesions. The SD of malignant tumors was found to be relatively high (14). Our research revealed that in addition to the higher E_{max} values of malignant foci, the elastic areas of some foci were larger than the B-mode areas. Does the hardness of the tissues surrounding the lesion change earlier than the morphology of the lesion? In Leong's study, it was mentioned that the area ratio of shear wave elastography to B-mode ultrasound (AR) is helpful for identifying benign and malignant breast foci (15), but the clinicopathological characteristics of AR in different malignant foci were not discussed. In view of the fact that there are few studies in this area, we aim to explore the distribution characteristics and influencing factors of AR in different breast cancer types in our research. We also intend to explore the relationships between the SWE parameters and the clinicopathological characteristics of breast cancer patients to clarify the feasibility of using SWE parameters to predict the biological behavior of breast cancer preoperatively.

MATERIALS AND METHODS

Patients

Our research obtained approval from the Shanghai Cancer Center Institutional Review Board, and all patients in this prospective study provided written informed consent. This work was conducted from January 2019 to December 2019 at Fudan University Shanghai Cancer Center (FUSCC). We prospectively recruited 354 breast patients with 362 breast foci for B-mode ultrasound (US) and SWE during this period. The maximum diameter of the collected breast lesions was less than or equal to 30 mm. Compared with postoperative and biopsy pathologies, 116 benign breast lesions were excluded. Another 22 breast patients were excluded for the reasons as follows: 7 lesions in 7 patients showed artifacts because of tumor prominence from the skin surface; 10 breast lesions in 10 patients were located beside or behind the nipple, which causes inaccuracy in elastography imaging; and 5 patients were lost to follow-up. Therefore, a total of 224 breast cancer foci from 216 patients were ultimately included.

B-Mode Ultrasound and Shear Wave Elastography Examinations

B-mode US and SWE were performed by a Supersonic Aixplorer color Doppler ultrasound diagnostic apparatus (Supersonic Imagine, Aix en Provence, France) and which was equipped with an SL 15-4 MHz linear array transducer, and all imaging was collected by 2 sonographers with 5 and 10 years of working experience in breast US. In this study, the largest section of the lesion was selected to measure S_{max} in B-mode US examination. Then, the mode was switched to SWE for the largest section of the lesion to obtain the SWE image. The sampling frame was placed to contain the entire section of the lesion. The transducer remained relatively fixed for approximately 5 s without applying pressure to the lesion until the SWE image was stable and no artifacts were observed, and then the SWE image was obtained. Regions of interest (ROIs) were acquired respectively in shear wave elastography and B-mode ultrasound by drawing the outline of the border of the lesion to exactly include the lesion without breast tissue around the lesion. E_{max} value, E_{mean} value, and SD value of the entire lesion in the SWE image were generated automatically by the US system, as well as the area of the lesion in the B-mode US image and the area of the lesion in the SWE image, and then AR was calculated by dividing the area of SWE by the area of B-mode US. Each of the lesions was continuously obtained three times for three images, and then average values of three images were calculated to obtain the final results of the SWE parameters.

Pathologic Analysis

The pathological evaluation parameters included ER and PR status, HER2 status, Ki-67 proliferation index, LVI status, ALN

status, pathological type, histopathological classification, and histological grade. Histological grading was determined by the Nottingham grading system (16). The modified American Society of Clinical Oncology-College of American Physicians scoring guidelines were used (17, 18). Positive ER and PR status was ruled as ≥1% of tumor cells with positive nuclear staining. On the basis of standard criteria, the immunohistochemical staining of HER2 was scored as 0, 1+, 2+, and 3+. A score of 3+ was classified as HER2 positive, and a score of 0 and 1+ was classified as negative. There was another situation where tumors had a score of 2+, fluorescence *in situ* hybridization testing (FISH) was performed to identify the amplification of HER2, and a positive FISH amplification was defined to be HER2 positive. The cells with Ki-67 nuclear-stained were defined to be positively stained. Based on this criterion, the Ki-67 proliferation index was defined as the percentage of cells that were Ki-67-positive in greater than or equal to 500 tumor cells in the hot spot of each slide. High Ki-67 index expression was ruled as ≥14%, and low Ki-67 index expression was ruled as <14% (19). ALN positivity refers to axillary lymph node metastasis. LVI positivity was defined as the existence of cancer cells in the lymphatic and/or blood vessels in or around the tumor. Two senior physicians pathologically diagnosed the lesions, and a consensus was reached through discussion when disagreements occurred.

Statistical Analysis

The Statistical Package for the Social Sciences for Windows (SPSS, Chicago, IL, USA, version 23.0) was used to perform all statistical analyses. A simple linear regression model was used to analyze the relationship between the SWE parameters and clinicopathological parameters. Stepwise multiple linear regression analysis was used to determine the pathological parameters independently related to the SWE parameters. A p-value < 0.05 was considered statistically significant.

RESULTS

Clinical Characteristics

The average age of the 216 breast cancer patients was 53.7 ± 10.6 years (range 28–79 years), the average size of the 224 foci was 17.6 ± 5.7 mm (range 5.0–30.0 mm), and the average values of the SWE parameters were as follows: E_{max} 190.4 ± 84.1 kPa (range 12.0–300.0 kPa), E_{mean} 61.8 ± 29.3 kPa (range 4.0–169.5 kPa), SD 38.7 ± 20.5 kPa (range 2.6–87.9 kPa), and AR 1.97 ± 0.79 (range 0.82–8.05), and the AR values of all lesions were higher than 1 except for one lesion, which was 0.82 (**Figures 1**–**3**). The B-mode US and SWE parameters of the lesions are listed in (**Figure 4**).

A total of 224 breast cancer foci were observed among 216 breast cancer patients. The pathological types of cancers included 188 lesions of invasive ductal carcinomas (IDC), 21 lesions of ductal carcinomas *in situ* (DCIS), 7 lesions of invasive lobular breast carcinomas (ILC), 1 lesion of lobular carcinoma *in situ*, 3 lesions of intraductal papillary carcinomas, 1 lesion of solid intraductal papillary carcinoma with interstitial infiltration, 2 lesions of mucinous adenocarcinomas, and 1 lesion of invasive carcinoma with mucus secretion.

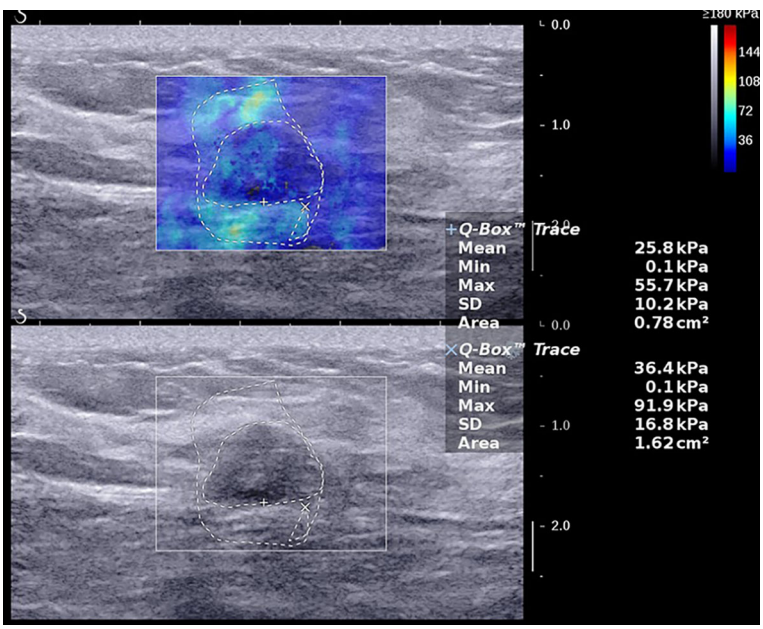


FIGURE 1 | A 47-year-old female with a 14.0-mm lesion in the right breast. A freehand ROI was drawn manually to measure areas by drawing the outline of the border of the lesion respectively in shear wave elastography and B-mode ultrasound. The SWE values of the lesion were as follows: E_{max}, 137.9 kPa; E_{mean}, 44.4 kPa; SD, 27.2 kPa; AR, 2.29. The lesion was invasive ductal carcinoma with the histological grade II, showing negative lymphovascular invasion and negative axillary lymph node status, ER^+ and PR^+, $HER2^-$, and low expression of Ki-67, 10%.

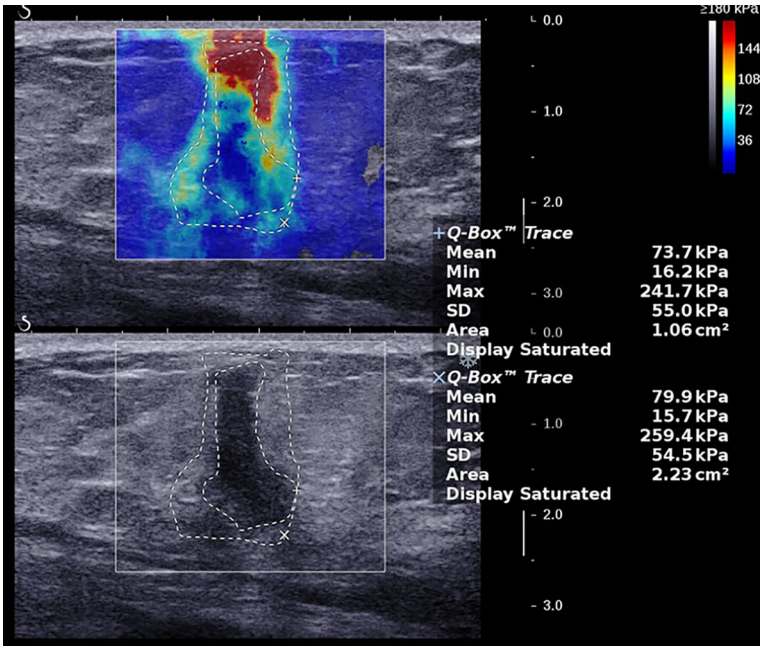


FIGURE 2 | A 57-year-old female with a 15.0-mm lesion in the left breast. The SWE parameter values of the cancer were as follows: E_{max}, 261.5 kPa; E_{mean}, 82.4 kPa; SD, 54.2 kPa; AR, 1.95. The lesion was invasive ductal carcinoma with the histological grade II, showing negative lymphovascular invasion and negative axillary lymph node status, ER^- and PR^-, $HER2^+$, and high expression of Ki-67, 15%.

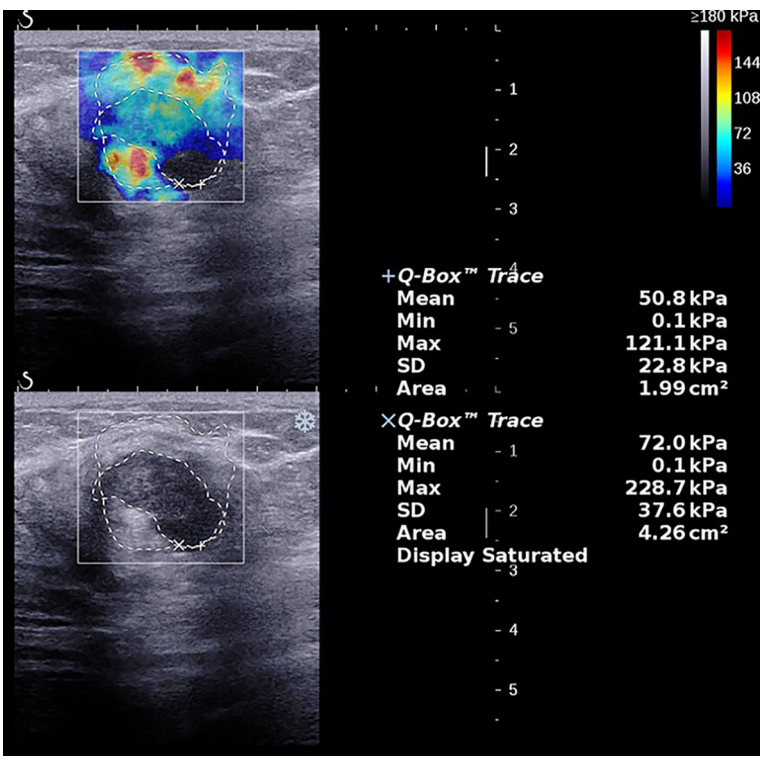


FIGURE 3 | A 62-year-old female with a 28.0-mm lesion in the left breast. The SWE parameter values of the cancer were as follows: E_{max}, 236.0 kPa; E_{mean}, 69.3 kPa; SD, 41.5 kPa; AR, 2.02. The lesion was invasive ductal carcinoma with the histological grade III, showing negative lymphovascular invasion and negative axillary lymph node status, ER^- and PR^-, $HER2^-$, and high expression of Ki-67, 60%.

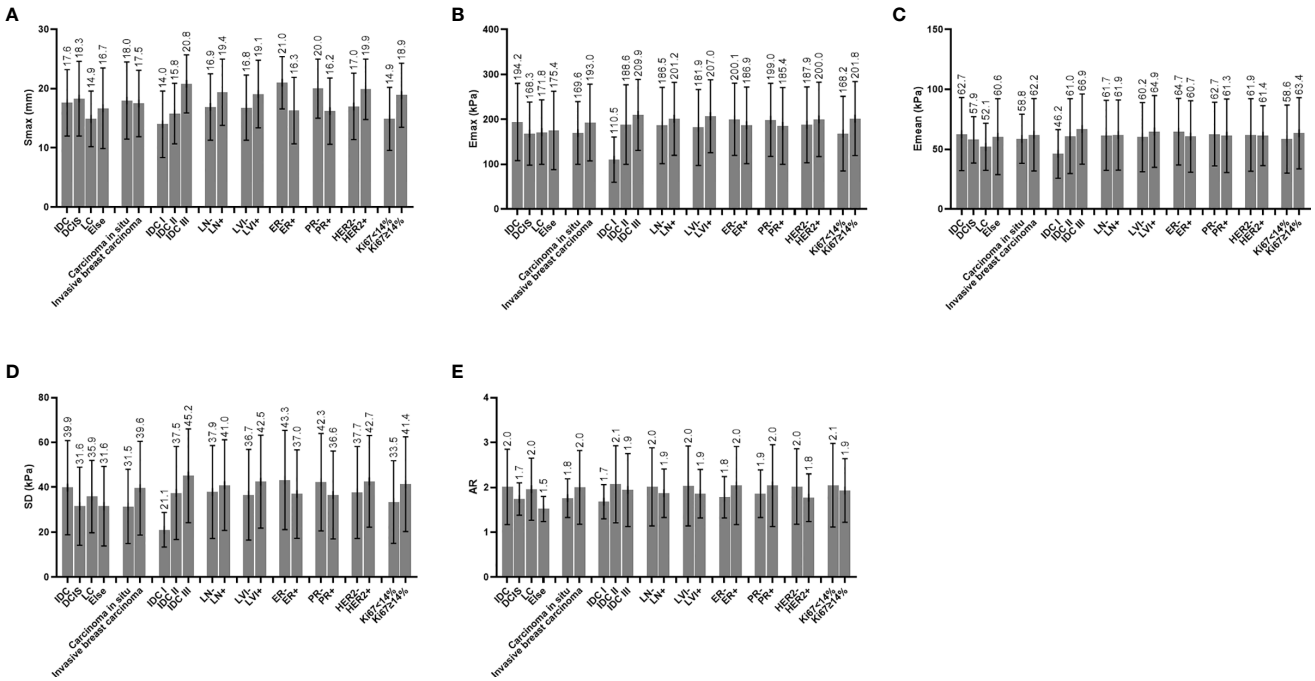


FIGURE 4 | **(A)** Distribution characteristics of S_{max} between different groups. **(B)** Distribution characteristics of E_{max} between different groups. **(C)** Distribution characteristics of E_{mean} between different groups. **(D)** Distribution characteristics of SD between different groups. **(E)** Distribution characteristics of AR between different groups; IDC, invasive ductal carcinoma; DCIS, ductal carcinoma in situ; LC, lobular carcinoma; ALN, axillary lymph node status; LVI, lymphovascular invasion; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; Ki67, Ki67 index; S_{max}, maximal size; E_{max}, maximal elasticity; E_{mean}, mean elasticity; SD, standard deviation of elasticity; AR, area ratio of shear wave elastography to B-mode ultrasound.

Simple Linear Regression Analysis of the Correlation between SWE Parameters and Clinicopathological Parameters

Pathological Type, Histopathological Classification, and Histological Grade

In this study, pathological type had no correlation with SWE parameters (**Table 1**). On histopathological classification, there were 199 cases of invasive breast cancer and 25 cases of cancer *in situ*. The E_{max} ($p = 0.190$) and SD ($p = 0.062$) values of the invasive breast cancer group were higher than those of the cancer *in situ* group, but no significant differences were observed (**Table 1**).

Since most of the breast cancer cases in this research were IDC and there were fewer other pathological types of breast cancers, IDC cases were chosen to be analyzed for the relationship between

TABLE 1 | Univariate analysis of correlation of pathological type and histopathological classification with B-mode ultrasound and shear wave elastography parameters.

Variable		S_{max}		E_{max}		E_{mean}		SD		AR	
	n	*p*	*β* (95% CI)	*p*	*β* (95% CI)	*p*	*β* (95% CI)	*p*	*β* (95% CI)	*p*	*β* (95% CI)
Pathological type											
IDC	188	0.583	−0.716 (−3.287 to 1.854)	0.182	25.933 (−12.246 to 64.113)	0.484	4.739 (−8.577 to 18.055)	0.080	8.280 (−1.009 to 17.569)	0.128	0.278 (−0.080 to 0.636)
DCIS	21	\		\		\		\		\	
LC	8	0.143	−3.458 (−8.100 to 1.183)	0.919	3.544 (−65.400 to 72.488)	0.634	−5.826 (−29.871 to 18.220)	0.611	4.330 (−12.444 to 21.103)	0.500	0.222 (−0.425 to 0.869)
Else	7	0.514	−1.619 (−6.495 to 3.257)	0.846	7.148 (−65.274 to 79.570)	0.834	2.690 (−22.568 to 27.949)	0.999	−0.010 (−17.629 to 17.610)	0.535	−0.214 (−0.894 to 0.465)
Histopathological classification											
cancer *in situ*	25	\		\		\		\		\	
invasive breast cancer	199	0.707	−0.542 (−2.824 to 1.919)	0.190	23.432 (−11.676 to 58.540)	0.589	3.363 (−8.893 to 15.619)	0.062	8.121 (−0.419 to 16.661)	0.161	0.236 (−0.095 to 0.567)

IDC, invasive ductal carcinoma; DCIS, ductal carcinoma in situ; LC, lobular carcinoma; S_{max}, maximal size; E_{max}, maximal elasticity; E_{mean}, mean elasticity; SD, standard deviation of elasticity; AR, area ratio of shear wave elastography to B-mode ultrasound.

$p < 0.05$ was considered statistically significant.

histological grade and SWE parameters. In this study, there were 188 cases of IDC (including 1 case with missing histological grading information), including 4 cases of IDC-I, 115 cases of IDC-II, and 68 cases of IDC-III. Although the E_{max} (p = 0.072), E_{mean} (p = 0.339), and SD (p = 0.119) values were not significantly different between IDC-I and IDC-II, the average values of these parameters in the IDC-II group was greater than those in the IDC-I group. S_{max} (p = 0.000) and SD (p = 0.016) of the IDC-III group showed significantly greater values than those of the IDC-II group. The S_{max} (p = 0.010), E_{max} (p = 0.024), and SD (p = 0.024) in the group of IDC-III showed significantly higher values than those in the group of IDC-I (**Table 2**).

Axillary Lymph Node Status and Lymphovascular Invasion Status

There were 60 lesions in the ALN-positive group and 164 lesions in the ALN-negative group, and there were 76 lesions in the LVI-positive group and 148 lesions in the LVI-negative group. The S_{max} (p = 0.004) of the lesions showed lower values in the group of ALN-negative than that in the group of ALN-positive. E_{max} (p = 0.034), SD (p = 0.045), and S_{max} (p = 0.003) in the LVI-positive group were significantly higher than those in the LVI-negative group. Although the distributions were not significantly different, AR in the ALN-positive group (p = 0.237) was lower than in the ALN-negative group and AR in the LVI-positive group (p = 0.143) was lower than in the LVI-negative group (**Table 3**).

Immunohistochemical Factors

SD and S_{max} were significantly correlated with ER and PR status. Compared with the ER^- and the PR^- lesions, the ER^+ and the PR^+ lesions had lower SD (p = 0.043, p = 0.047) and S_{max} (p = 0.000, p = 0.000) values. AR (p = 0.032) was significantly higher in the ER^+ lesions than in the ER^- lesions. AR (p = 0.106)

TABLE 2 | Univariate analysis of correlation of histological grading with B-mode ultrasound and shear wave elastography parameters.

Variable		S_{max}		E_{max}		E_{mean}		SD		AR	
	n	*p*	*β* (95% CI)	*p*	*β* (95% CI)	*p*	*β* (95% CI)	*p*	*β* (95% CI)	*p*	*β* (95% CI)
Histological grading of IDC											
I	4	0.010	−6.809 (−11.945 to −1.673)	0.024	−99.363 (−185.467 to −13.260)	0.189	−20.634 (−51.483 to 10.215)	0.024	−24.090 (−45.032 to −3.147)	0.555	−0.255 (−1.107 to 0.597)
II	115	0.000	−4.983 (−6.510 to −3.456)	0.103	−21.281 (−46.882 to 4.320)	0.212	−5.820 (−14.992 to 3.353)	0.016	−7.662 (−13.889 to −1.435)	0.277	0.140 (−0.113 to 0.393)
III	68	\		\		\		\		\	

IDC, invasive ductal carcinoma; S_{max}, maximal size; E_{max}, maximal elasticity; E_{mean}, mean elasticity; SD, standard deviation of elasticity; AR, area ratio of shear wave elastography to B-mode ultrasound.
$p < 0.05$ was considered statistically significant.

TABLE 3 | Univariate analysis of correlation of ALN, LVI, and immunohistochemical biomarkers with B-mode ultrasound and shear wave elastography parameters.

	Variable	*N*	S_{max}		E_{max}		E_{mean}		SD		AR	
			p	*β* (95% CI)	*p*	*β* (95% CI)	*p*	*β* (95% CI)	*p*	*β* (95% CI)	*p*	*β* (95% CI)
ALN	ALN^-	164	\		\		\		\		\	
	ALN^+	60	0.004	2.448 (0.792 to 4.103)	0.248	14.699 (−10.286 to −10.286)	0.970	0.169 (−8.551 to 8.889)	0.318	3.102 (−3.004 to 9.209)	0.237	−0.142 (−0.377 to 0.094)
LVI	LVI^-	148	\		\		\		\		\	
	LVI^+	76	0.003	2.342 (0.795 to 3.889)	0.034	25.086 (1.882 to 48.289)	0.252	4.735 (−3.397 to 12.867)	0.045	5.807 (0.135 to 11.480)	0.143	−0.164 (−0.384 to 0.056)
ER	ER^-	59	\		\		\		\		\	
	ER^+	165	0.000	−4.719 (−6.296 to −3.142)	0.301	−13.229 (−38.363 to 11.906)	0.373	−3.967 (−12.718 to 4.784)	0.043	−6.295 (−12.391 to −0.198)	0.032	0.257 (0.022 to 0.492)
PR	PR^-	82	\		\		\		\		\	
	PR^+	142	0.000	−3.794 (−5.262 to −2.327)	0.245	−13.574 (−36.541 to 9.394)	0.720	−1.458 (−9.472 to 6.556)	0.047	−5.646 (−11.223 to −0.070)	0.106	0.178 (−0.038 to 0.394)
HER2	$HER2^-$	178	\		\		\		\		\	
	$HER2^+$	46	0.002	2.936 (1.128 to 4.744)	0.384	12.140 (−15.285 to 39.566)	0.927	−0.443 (−10.003 to 9.116)	0.141	5.003 (−1.674 to 11.679)	0.055	−0.251 (−0.508 to 0.006)
Ki67	<14%	76	\		\		\		\		\	
	≥14%	148	0.000	3.971 (2.483 to 5.459)	0.004	33.541 (10.525 to 56.557)	0.247	4.785 (−3.347 to 12.916)	0.006	7.900 (2.272 to 13.528)	0.259	−0.127 (−0.347 to 0.094)

ALN, axillary lymph node status; LVI, lymphovascular invasion; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; Ki67, Ki67 index; S_{max}, maximal size; E_{max}, maximal elasticity; E_{mean}, mean elasticity; SD, standard deviation of elasticity; AR, area ratio of shear wave elastography to B-mode ultrasound.
$p < 0.05$ was considered statistically significant.

in the PR^+ lesions was also higher than in the PR^- lesions but there were no significant differences. The S_{max} ($p = 0.002$) of the $HER2^+$ foci had significantly higher values than that of the $HER2^-$ foci. The E_{max} ($p = 0.004$), SD ($p = 0.006$), and S_{max} ($p = 0.000$) values of the lesions in the Ki-67 high-expression group were significantly higher than those in the Ki-67 low-expression group. The AR ($p = 0.055$) of the $HER2^+$ lesions was lower than those of the $HER2^-$ lesions, and AR ($p = 0.259$) in the Ki-67 high-expression group was also lower than that in the Ki-67 low-expression group, but both no significant differences were observed (**Table 3**).

Multivariate Analysis

It was shown in the multiple linear regression analysis that the Ki-67 index was an independent influencing factor of E_{max} ($p = 0.004$), SD ($p = 0.006$), and S_{max} ($p = 0.005$); ER status and LVI were independent influencing factors of Smax ($p = 0.000$, $p = 0.006$), and ER status independently influenced AR ($p = 0.032$) (**Table 4**).

DISCUSSION

This study explored the relationship between SWE parameters and clinicopathological characteristics. The results showed that in the univariate analysis, there was no significant difference in the distribution of SWE parameters among different pathological types of breast cancers, which is consistent with the conclusions reported by Ganau et al. (20), who indicated that pathological type is not the main factor affecting SWE parameters. Invasive breast cancer and cancer *in situ* have different biological behaviors and different clinical treatment plans. Therefore, we divided breast cancers into invasive breast cancer group and cancer *in situ* group. Compared with the cancers *in situ*, the invasive breast cancers showed higher values of E_{max} ($p = 0.190$) and SD ($p = 0.062$). As the aggressiveness of the lesion increases, both E_{max} and SD may increase.

The histological grade of IDC affects the clinical characteristics of the tissue. Our research clarified that lesions become harder and more heterogeneous with an increasing histological grade. The SD ($p = 0.016$) and S_{max} ($p = 0.000$) in the group of IDC-III were significantly higher than those in the group of IDC-II, and the S_{max} ($p = 0.010$), E_{max} ($p = 0.024$), and SD ($p = 0.024$) in the group of IDC-III showed higher values than those in the group of IDC-I. Even though the E_{max} ($p = 0.072$), E_{mean} ($p = 0.339$), and SD ($p = 0.119$) in the group of IDC-II were greater than those in the group of IDC-I, the SWE parameter distribution between the two groups did not show significant differences. This may be due to the small number of cases in the IDC-I group. The study by Evans et al. obtained similar results, showing that a higher histological grade was positively associated with a higher E_{mean} (21).

Our results showed that in the ALN-positive group, S_{max} ($p = 0.004$) showed greater values than that in the ALN-negative group, indicating that larger masses have a higher risk of ALN metastasis. There were no significant differences in the distribution of SWE parameters between the ALN-positive group and the ALN-negative group; E_{mean} ($p = 0.970$) values in the ALN-positive group and the ALN-negative group were similar; however, E_{max} ($p = 0.248$) in the ALN-positive group was greater than that in the ALN-negative group. The result is consistent with the report by Xue et al., in which they concluded that the E_{max} values in the breast cancer ALN-negative group were lower than that in the ALN-positive group ($p = 0.110$) but without significant differences (22). ALN can reflect prognosis, and it has been reported that E_{mean} or E_{max} of breast cancer in the ALN-positive group has higher Young's modulus value (23, 24).

LVI positivity means that tumor cells have infiltrated the lymphatic and/or blood vessels in the area around the tumor, which is one of the critical steps of metastasis and is related to a poor prognosis (25, 26). The research by Son et al. indicated that the state of LVI is associated with E_{mean} and E_{max}, with both values being higher in the LVI-positive group (27). In our research, the E_{max} ($p = 0.034$), SD ($p = 0.045$), and S_{max} ($p = 0.003$) of the LVI-positive group showed significantly higher values than those of the LVI-negative group, indicating that larger lesions become harder and more heterogeneous and are more likely to be LVI-positive. Therefore, preoperative E_{max} and SD values can infer whether a lesion has progressed to be LVI-positive. However, our research did not find a correlation between LVI and E_{mean}. In the multiple linear regression analysis, the LVI did not show an independent correlation

TABLE 4 | Multivariate linear regression analysis for clinicopathological factors.

	Variable	Smax		Emax		SD		AR	
		p	*β* (95% CI)	*p*	*β* (95% CI)	*p*	*β* (95% CI)	*p*	*β* (95% CI)
LVI	LVI^-	\							
	LVI^+	0.006	2.208 (0.586 to 3.469)						
ER	ER^-	\						\	
	ER^+	0.000	−3.838 (−5.491 to −2.184)					0.032	0.257 (0.022 to 0.492)
Ki67	<14%	\		\		\			
	≥14%	0.005	2.239 (0.677 to 3.801)	0.004	33.541 (10.525 to 56.557)	0.006	7.900 (2.272 to 13.528)		

LVI, lymphovascular invasion; ER, estrogen receptor; Ki67, Ki67 index; S_{max}, maximal size; E_{max}, maximal elasticity; SD, standard deviation of elasticity; AR, area ratio of shear wave elastography to B-mode ultrasound.

$p < 0.05$ was considered statistically significant.

with E_{max} and SD, which may be due to the correlation between those two parameters and Ki-67. LVI-positive is often correlated with a high Ki-67 index (28).

ER, PR, HER2, and Ki-67 are markers used to distinguish breast cancer subtypes, and their expression status affects patients' clinical treatment plan and prognosis (29). Our research results showed that the SD (p = 0.043, p = 0.047) and S_{max} (p = 0.000, p = 0.000) values of the ER^+ and the PR^+ lesions were smaller than those of the ER^- and the PR^- lesions. AR (p = 0.032) was significant higher in the ER^+ lesions than in the ER^- lesions. We also found that AR (p = 0.106) was higher in the PR^+ lesions, but there were no significant differences. Existing research results have indicated that ER^- and PR^- breast cancers tend to a worse prognosis than ER^+ or PR^+ breast cancers, while ER^+ and PR^+ tumors are well differentiated and less aggressive to a large extent (30). Therefore, based on the patient's preoperative S_{max}, SD, and AR values, we can infer the expression status of ER and PR. ER status independently affected S_{max} (p = 0.000) and AR (p = 0.032) in the multiple linear regression analysis, but failed to show an independent association with SD. This can be attributed to the relationship between SD and Ki-67, as ER^- tumors tend to have a higher Ki-67 index (31). In addition, our research shows that HER2 status is correlated with S_{max} (p = 0.002). The S_{max} value in the $HER2^+$ group was larger than that in the $HER2^-$ group, which may be a result from the high degree of malignancy and faster growth rate of the $HER2^+$ lesions (32). Our study also found that the E_{max} and SD values of the $HER2^+$ lesions were higher than those of the $HER2^-$ lesions, but no significant differences were observed. At present, all the evidence from clinical studies have indicated that ER^+, PR^+, and $HER2^-$ tumors have a good prognosis and show an excellent response to hormone therapy (33–37). This conclusion coincides with our research results, which show that ER^+, PR^+, and $HER2^-$ tumors have smaller sizes, lower E_{max} and SD values, and higher AR values in most cases, which indicates a better prognosis.

Ki-67 is a nuclear antigen that reflects the proliferation activity of tumor cells, and its expression level is related to the degree of tumor malignancy. The higher the Ki-67 index, the more rapid the rate of tumor proliferation is, and the stronger the invasion and metastasis abilities are (38–40). In the univariate analysis, the E_{max} (p = 0.004), SD (p = 0.006) and S_{max} (p = 0.000) values of the lesions in the Ki-67 high-expression group showed significantly higher values than those in the Ki-67 low-expression group. In the multiple regression analysis, the Ki-67 index was also an independent influencing factor of E_{max} (p = 0.007), SD (p = 0.007), and S_{max} (p = 0.006). In this study, the higher SD values in the Ki-67 high-expression group may be triggered by the following mechanism: the more vigorous the cell karyokinesis is during tumor growth, the higher the Ki-67 index will be, which, in turn, increases the SD value, reflecting the heterogeneity in tumor cells. Therefore, the SD value can be used to noninvasively assess the biological behavior of cancer cells *in vivo*. E_{max} represents the hardness of the lesion, which is affected by the structure of necrotic tissue and fibrous tissue. In this study, tumors with a higher Ki-67 value had a larger E_{max} value, which is similar to the results of Nishimura et al. (41). This may explain why ischemic and hypoxic necrosis occurs inside lesions with high rates of tumor proliferation, which subsequently leads to fibrosis and hyperplasia, thereby increasing the hardness of the tumor (42, 43). In the Ki-67 high-expression group, AR (p = 0.259) was lower than in the Ki-67 low-expression group, although no significant differences were observed. In B-mode US images, infiltrating growth of the tumor does not change the acoustic impedance, but the shear wave elastic velocity of the tumor is changed, so the shear wave elastic area of the tumor could be larger than the B-mode image area, which may explain why the AR value is higher in malignant tumors than in benign breast foci. The AR values of all lesions in our study were higher than 1 except for one lesion, which was 0.82. We speculated that the worse the prognosis of the tumor, the more likely visible the surrounding infiltrated parts of the tumor on B-mode US, then the area of B-mode US will be larger, but the area of SWE may not increase proportionally, so AR values may be lower. On the contrary, the better the prognosis of breast cancer, the higher its AR value. So, this may explain why AR was higher in ALN-negative, LVI-negative, ER^+, PR^+, $HER2^-$, and Ki-67 low-expression groups, which means better prognosis in our research.

This study has some shortcomings that should be noted: the number of pathological types of breast cancer cases showed a large difference in their distribution, with IDC being the majority of cases, which is related to the different incidences of different breast cancer types. We hope to expand the sample size in the future to further explore the distribution of SWE parameters in different pathological types of breast cancers.

CONCLUSION

The shear wave elastography parameters E_{max}, SD, and AR can be used to evaluate the biological behavior of breast cancer.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Fudan University Shanghai Cancer Center Institutional Review Board. The patients/participants

provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CL, JZ, and WZ offered the conception and design of the study. CC offered administrative support. CL and WZ collected the clinical and image data. CL and JZ conducted data analysis and interpretation. CL and JZ wrote the manuscript. WZ and CC reviewed and re-edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation of China (81801701 and 81830058).

ACKNOWLEDGMENTS

We thank the doctors from Ultrasound Department and Breast Surgery of FUSCC for their help with the acquisition of cases and images.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2021) 71:209–49. doi: 10.3322/caac.21660
2. Yu F, Quan F, Xu J, Zhang Y, Xie Y, Zhang J, et al. Breast Cancer Prognosis Signature: Linking Risk Stratification to Disease Subtypes. *Brief Bioinform* (2019) 20:2130–40. doi: 10.1093/bib/bby073
3. Duffy MJ, Harbeck N, Nap M, Molina R, Nicolini A, Senkus E, et al. Clinical Use of Biomarkers in Breast Cancer: Updated Guidelines From the European Group on Tumor Markers (EGTM). *Eur J Cancer* (2017) 75:284–98. doi: 10.1016/j.ejca.2017.01.017
4. Giuliano AE, Connolly JL, Edge SB, Mittendorf EA, Rugo HS, Solin LJ, et al. Breast Cancer-Major Changes in the American Joint Committee on Cancer Eighth Edition Cancer Staging Manual. *CA Cancer J Clin* (2017) 67:290–303. doi: 10.3322/caac.21393
5. Hanker AB, Sudhan DR, Arteaga CL. Overcoming Endocrine Resistance in Breast Cancer. *Cancer Cell* (2020) 37:496–513. doi: 10.1016/j.ccell.2020.03.009
6. Brem RF, Lenihan MJ, Lieberman J, Torrente J. Screening Breast Ultrasound: Past, Present, and Future. *AJR Am J Roentgenol* (2015) 204:234–40. doi: 10.2214/AJR.13.12072
7. Diana A, Carlino F, Franzese E, Oikonomidou O, Criscitiello C, De Vita F, et al. Early Triple Negative Breast Cancer: Conventional Treatment and Emerging Therapeutic Landscapes. *Cancers (Basel)* (2020) 12:819. doi: 10.3390/cancers12040819
8. Choi HJ, Ko KH, Jung HK. Shear Wave Elastography for Surgically Verified Breast Papillary Lesions: Is it Effective for Differentiation Between Benign and Malignant Lesions? *J Ultrasound Med* (2017) 36:2007–14. doi: 10.1002/jum.14293
9. Youk JH, Gweon HM, Son EJ. Shear-Wave Elastography in Breast Ultrasonography: The State of the Art. *Ultrasonography* (2017) 36:300–9. doi: 10.14366/usg.17024
10. Song EJ, Sohn YM, Seo M. Tumor Stiffness Measured by Quantitative and Qualitative Shear Wave Elastography of Breast Cancer. *Br J Radiol* (2018) 91:20170830. doi: 10.1259/bjr.20170830
11. Huang Y, Liu Y, Wang Y, Zheng X, Han J, Li Q, et al. Quantitative Analysis of Shear Wave Elastic Heterogeneity for Prediction of Lymphovascular Invasion in Breast Cancer. *Br J Radiol* (2021) 94:20210682. doi: 10.1259/bjr.20210682
12. Cantisani V, David E, Barr RG, Radzina M, de Soccio V, Elia D, et al. US-Elastography for Breast Lesion Characterization: Prospective Comparison of US BIRADS, Strain Elastography and Shear Wave Elastography. *Ultraschall Med* (2021) 42:533–40. doi: 10.1055/a-1134-4937
13. Kim H, Lee J, Kang BJ, Kim SH. What Shear Wave Elastography Parameter Best Differentiates Breast Cancer and Predicts Its Histologic Aggressiveness? *Ultrasonography* (2021) 40:265–73. doi: 10.14366/usg.20007
14. Evans A, Whelehan P, Thomson K, McLean D, Brauer K, Purdie C, et al. Quantitative Shear Wave Ultrasound Elastography: Initial Experience in Solid Breast Masses. *Breast Cancer Res* (2010) 12:R104. doi: 10.1186/bcr2787
15. Leong LC, Sim LS, Lee YS, Ng FC, Wan CM, Fook-Chong SM, et al. A Prospective Study to Compare the Diagnostic Performance of Breast Elastography Versus Conventional Breast Ultrasound. *Clin Radiol* (2010) 65:887–94. doi: 10.1016/j.crad.2010.06.008
16. Elston CW, Ellis IO. Pathological Prognostic Factors in Breast Cancer. I. The Value of Histological Grade in Breast Cancer: Experience From a Large Study With Long-Term Follow-Up. *Histopathology* (1991) 19:403–10. doi: 10.1111/j.1365-2559.1991.tb00229.x
17. Hammond ME, Hayes DF, Wolff AC. Clinical Notice for American Society of Clinical Oncology-College of American Pathologists Guideline Recommendations on ER/PgR and HER2 Testing in Breast Cancer. *J Clin Oncol* (2011) 29:e458. doi: 10.1200/JCO.2011.35.2245
18. Wolff AC, Hammond M, Allison KH, Harvey BE, Mangu PB, Bartlett J, et al. Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. *J Clin Oncol* (2018) 36:2105–22. doi: 10.1200/JCO.2018.77.8738
19. Penault-Llorca F, Radosevic-Robin N. Ki67 Assessment in Breast Cancer: An Update. *Pathology* (2017) 49:166–71. doi: 10.1016/j.pathol.2016.11.006
20. Ganau S, Andreu FJ, Escribano F, Martin A, Tortajada L, Villajos M, et al. Shear-Wave Elastography and Immunohistochemical Profiles in Invasive Breast Cancer: Evaluation of Maximum and Mean Elasticity Values. *Eur J Radiol* (2015) 84:617–22. doi: 10.1016/j.ejrad.2014.12.020
21. Evans A, Whelehan P, Thomson K, McLean D, Brauer K, Purdie C, et al. Invasive Breast Cancer: Relationship Between Shear-Wave Elastographic Findings and Histologic Prognostic Factors. *Radiology* (2012) 263:673–7. doi: 10.1148/radiol.12111317
22. Xue X, Li J, Wan W, Shi X, Zheng Y. Kindlin-2 Could Influence Breast Nodule Elasticity and Improve Lymph Node Metastasis in Invasive Breast Cancer. *Sci Rep* (2017) 7:6753. doi: 10.1038/s41598-017-07075-1
23. Evans A, Rauchhaus P, Whelehan P, Thomson K, Purdie CA, Jordan LB, et al. Does Shear Wave Ultrasound Independently Predict Axillary Lymph Node Metastasis in Women With Invasive Breast Cancer? *Breast Cancer Res Treat* (2014) 143:153–7. doi: 10.1007/s10549-013-2747-z
24. Au FW, Ghai S, Lu FI, Moshonov H, Crystal P. Quantitative Shear Wave Elastography: Correlation With Prognostic Histologic Features and Immunohistochemical Biomarkers of Breast Cancer. *Acad Radiol* (2015) 22:269–77. doi: 10.1016/j.acra.2014.10.007
25. Song YJ, Shin SH, Cho JS, Park MH, Yoon JH, Jegal YJ. The Role of Lymphovascular Invasion as a Prognostic Factor in Patients With Lymph Node-Positive Operable Invasive Breast Cancer. *J Breast Cancer* (2011) 14:198–203. doi: 10.4048/jbc.2011.14.3.198
26. Ma Q, Dieterich LC, Detmar M. Multiple Roles of Lymphatic Vessels in Tumor Progression. *Curr Opin Immunol* (2018) 53:7–12. doi: 10.1016/j.coi.2018.03.018
27. Son MJ, Kim S, Jung HK, Ko KH, Koh JE, Park AY. Can Ultrasonographic Vascular and Elastographic Features of Invasive Ductal Breast Carcinoma Predict Histologic Aggressiveness? *Acad Radiol* (2020) 27:487–96. doi: 10.1016/j.acra.2019.06.009
28. Asaoka M, Patnaik SK, Zhang F, Ishikawa T, Takabe K. Lymphovascular Invasion in Breast Cancer is Associated With Gene Expression Signatures of Cell Proliferation But Not Lymphangiogenesis or Immune Response. *Breast Cancer Res Treat* (2020) 181:309–22. doi: 10.1007/s10549-020-05630-5
29. Hammerl D, Smid M, Timmermans AM, Sleijfer S, Martens J, Debets R. Breast Cancer Genomics and Immuno-Oncological Markers to Guide Immune Therapies. *Semin Cancer Biol* (2018) 52:178–88. doi: 10.1016/j.semcancer.2017.11.003

30. Barzaman K, Karami J, Zarei Z, Hosseinzadeh A, Kazemi MH, Moradi-Kalbolandi S, et al. Breast Cancer: Biology, Biomarkers, and Treatments. *Int Immunopharmacol* (2020) 84:106535. doi: 10.1016/j.intimp.2020.106535
31. Li Y, Zhang X, Qiu J, Pang T, Huang L, Zeng Q. Comparisons of P53, KI67 and BRCA1 Expressions in Patients With Different Molecular Subtypes of Breast Cancer and Their Relationships With Pathology and Prognosis. *J Buon* (2019) 24:2361–8.
32. Chung I, Reichelt M, Shao L, Akita RW, Koeppen H, Rangell L, et al. High Cell-Surface Density of HER2 Deforms Cell Membranes. *Nat Commun* (2016) 7:12742. doi: 10.1038/ncomms12742
33. Dai X, Xiang L, Li T, Bai Z. Cancer Hallmarks, Biomarkers and Breast Cancer Molecular Subtypes. *J Cancer* (2016) 7:1281–94. doi: 10.7150/jca.13141
34. Nicolini A, Ferrari P, Duffy MJ. Prognostic and Predictive Biomarkers in Breast Cancer: Past, Present and Future. *Semin Cancer Biol* (2018) 52:56–73. doi: 10.1016/j.semcancer.2017.08.010
35. Lee O, Sullivan ME, Xu Y, Rogers C, Muzzio M, Helenowski I, et al. Selective Progesterone Receptor Modulators in Early-Stage Breast Cancer: A Randomized, Placebo-Controlled Phase II Window-Of-Opportunity Trial Using Telapristone Acetate. *Clin Cancer Res* (2020) 26:25–34. doi: 10.1158/1078-0432.CCR-19-0443
36. Bartlett J, Sgroi DC, Treuner K, Zhang Y, Ahmed I, Piper T, et al. Breast Cancer Index and Prediction of Benefit From Extended Endocrine Therapy in Breast Cancer Patients Treated in the Adjuvant Tamoxifen-To Offer More? (ATTom) Trial. *Ann Oncol* (2019) 30:1776–83. doi: 10.1093/annonc/mdz289
37. Smyth LM, Piha-Paul SA, Won HH, Schram AM, Saura C, Loi S, et al. Efficacy and Determinants of Response to HER Kinase Inhibition in HER2-Mutant Metastatic Breast Cancer. *Cancer Discov* (2020) 10:198–213. doi: 10.1158/2159-8290.CD-19-0966
38. Nielsen TO, Leung S, Rimm DL, Dodson A, Acs B, Badve S, et al. Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group. *J Natl Cancer Inst* (2021) 113:808–19. doi: 10.1093/jnci/djaa201
39. Thangarajah F, Enninga I, Malter W, Hamacher S, Markiefka B, Richters L, et al. A Retrospective Analysis of Ki-67 Index and Its Prognostic Significance in Over 800 Primary Breast Cancer Cases. *Anticancer Res* (2017) 37:1957–64. doi: 10.21873/anticanres.11536
40. Liang Q, Ma D, Gao RF, Yu KD. Effect of Ki-67 Expression Levels and Histological Grade on Breast Cancer Early Relapse in Patients With Different Immunohistochemical-Based Subtypes. *Sci Rep* (2020) 10:7648. doi: 10.1038/s41598-020-64523-1
41. Nishimura R, Osako T, Okumura Y, Hayashi M, Toyozumi Y, Arima N. Ki-67 as a Prognostic Marker According to Breast Cancer Subtype and a Predictor of Recurrence Time in Primary Breast Cancer. *Exp Ther Med* (2010) 1:747–54. doi: 10.3892/etm.2010.133
42. Hasebe T, Mukai K, Tsuda H, Ochiai A. New Prognostic Histological Parameter of Invasive Ductal Carcinoma of the Breast: Clinicopathological Significance of Fibrotic Focus. *Pathol Int* (2000) 50:263–72. doi: 10.1046/j.1440-1827.2000.01035.x
43. Zhang Y, Ou Y, Yu D, Yong X, Wang X, Zhu B, et al. Clinicopathological Study of Centrally Necrotizing Carcinoma of the Breast. *BMC Cancer* (2015) 15:282. doi: 10.1186/s12885-015-1305-y

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Liu, Zhou, Chang and Zhi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 28 January 2022
doi: 10.3389/fonc.2021.842169

Application of Intravoxel Incoherent Motion Diffusion-Weighted Imaging in Predicting and Monitoring Early Efficacy of Anti-Angiogenic Therapy in the C6 Glioma Rat Model

Weishu Hou[1†], Yangyang Xue[1,2†], Yinfeng Qian[1], Hongli Pan[1], Man Xu[1], Yujun Shen[3], Xiaohu Li[1] and Yongqiang Yu[1*]*

[1] Department of Radiology, First Affiliated Hospital of Anhui Medical University, Hefei, China, [2] Department of Nuclear Medicine, The First Affiliated Hospital of Anhui Medical University, Hefei, China, [3] Department of Basic Medical Sciences, School of Basic Medical Sciences, Anhui Medical University, Hefei, China

OPEN ACCESS

Edited by:
Guang Yang,
Imperial College London,
United Kingdom

Reviewed by:
Zixing Huang,
Sichuan University, China
Zeyu Xiao,
Jinan University, China

***Correspondence:**
Xiaohu Li
zhaorenlxh@126.com
Yongqiang Yu
ayfyyuyongqiang@126.com

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 23 December 2021
Accepted: 31 December 2021
Published: 28 January 2022

Citation:
Hou W, Xue Y, Qian Y, Pan H, Xu M, Shen Y, Li X and Yu Y (2022) Application of Intravoxel Incoherent Motion Diffusion-Weighted Imaging in Predicting and Monitoring Early Efficacy of Anti-Angiogenic Therapy in the C6 Glioma Rat Model. Front. Oncol. 11:842169. doi: 10.3389/fonc.2021.842169

Objective: To investigate the feasibility of intravoxel incoherent motion (IVIM) diffusion-weighted imaging (DWI) in evaluating early effects of anti-angiogenic therapy in the C6 glioma rat model.

Methods: Twenty-six rats of the C6 glioma model were randomly divided into a treatment group (received bevacizumab) and a control group (physiological saline). IVIM-DWI was performed on days 0, 1, 3, 5, and 7 after anti-angiogenic therapy and tumor growth and IVIM-DWI parameters were dynamically observed. Hematoxylin and eosin, CD34 microvessel density (MVD), proliferation of cell nuclear antigen (PCNA), and Hif-α staining were performed on day 7. One-way ANOVA was used to compare intra-group differences and an independent-samples *t*-test was used to compare inter-group differences of MRI parameters. Correlations between IVIM-DWI parameters, tumor size, and pathological results were analyzed.

Results: The relative change in tumor volume (ΔVolume) in the two groups differed significantly on days 5 and 7 ($p = 0.038$ and $p < 0.001$). The perfusion-related parameters D*- and f-values decreased in the treatment group and demonstrated significant differences compared with the control group on days 3, 5, and 7 ($p = 0.033$, $p < 0.001$, and $p < 0.001$, respectively). The diffusion-related parameters ADC and D-values increased in the treatment group and were found to be significantly differently different from the control group on days 5 and 7 (both $p < 0.001$). The initial D-value showed a negative correlation with ΔVolume ($\gamma = -0.744$, $p < 0.001$), whereas the initial D*-value and relative change of D-value had a positive correlation with ΔVolume ($\gamma = 0.718$, $p < 0.001$ and $\gamma = 0.800$, $p < 0.001$, respectively). MVD was strongly positively correlated with D*-value ($r = 0.886$, $p = 0.019$), PCNA was negatively correlated with ADC- and D-values ($r = -0.848$, $p = 0.033$; and $r = -0.928$ $p = 0.008$, respectively), and Hif-1α was strongly negatively correlated with D*-value ($r = -0.879$, $p = 0.010$).

Conclusion: IVIM-DWI was sensitive and accurate in predicting and monitoring the effects of early anti-angiogenesis therapy in a C6 glioma rat model.

Keywords: intravoxel incoherent motion diffusion-weighted imaging (IVIM-DWI), glioma, angiogenesis inhibitors, therapy, vascular normalization

INTRODUCTION

High-grade glioma is the most aggressive primary malignant tumor of the central nervous system in adults and has a high expression of vascular endothelial growth factor (VEGF) (1). VEGF plays an essential role in regulating angiogenesis of glioma and, thus, anti-angiogenesis therapy targeting the VEGF signaling pathway has been recognized as an effective targeted therapy method for malignant glioma (2, 3). Recently the recombinant humanized monoclonal antibody bevacizumab (Avastin®, Bev), an anti-angiogenesis medicine, has been widely used as a treatment, either alone or in combination with traditional chemotherapy, for recurrent glioblastoma and newly diagnosed high-grade glioma (4–6).

However, to date, how to monitor the effects of anti-angiogenesis therapy sensitively, dynamically, and non-invasively is still a major challenge of glioma therapy. Diffusion-weighted imaging (DWI) is a functional magnetic resonance imaging (MRI) sequence used to evaluate or predict treatment outcome of malignant tumors; however, the apparent diffusion coefficient (ADC) has been shown to be influenced by both the diffusion of water molecules and microcirculation due to the mono-exponential model of DWI (7, 8). Recently, an intravoxel incoherent motion DWI (IVIM-DWI) has been developed that can separate microcirculation from restricted Brownian self-diffusion, and has been applied to evaluate the microcirculation of tumors after anti-angiogenic therapy several types of animal models (9–11). However, few studies have focused on the early and dynamic changes of IVIM-DWI parameters during therapy and the correlation between IVIM-DWI parameters and histological assessment.

Therefore, the purpose of this study was to observe the early dynamics of IVIM-DWI parameters during anti-angiogenic therapy to explore the accuracy of IVIM-DWI in anti-angiogenic therapy outcome *via* determining the angiogenesis, cell proliferation, and hypoxia of cells in the C6 glioma rat model.

MATERIALS AND METHODS

Ethics Statement

All treated procedures were approved by the Ethics Committee for Animal Experimentation of Anhui Medical University of Anhui Province, China (Approval no. SCXK-Wan-2017–001) and were conducted in strict accordance with the Guidelines of the National Institutes of Health for the Care and Use of Laboratory Animals.

Cell Culture and Animal Model Establishment

C6 glioma cells were purchased from the Institute of Biochemistry and Cell Biology (Shanghai, PR, China). The cells were kept in the incubator at 37°C in 5% carbon dioxide, and cultured in high glucose Dulbecco's Modified Eagle's Medium (Abcam, Cambridge, UK). After C6 cells had reached approximately 90% confluency, they were digested with 0.05% trypsin and washed twice with phosphate buffered saline (Abcam, Cambridge, UK). Finally, the C6 cell suspension, at a concentration of 1×10^7 cells/ml, was prepared to establish the glioma orthotopic rat model.

Male Sprague–Dawley rats (purchased from Animal Center of Anhui Province, No. SCXK-Wan-2017–001) aged 6–8 weeks were anesthetized by intraperitoneal injection with 10% chloral hydrate (0.3 ml/100 g), and the glioma orthotopic rat model was established using a stereotactic apparatus (Kopf, Cayunga, CA, USA). The heads of the rats were shaved, disinfected with 0.1% povidone-iodine, and fixed on the stereotactic apparatus. An HP-4 dental drill bit was used to drill a hole in the skull at the following three coordinates: 1 mm to the anterior arcuate suture, 3 mm to the right of the sagittalis suture, and at a depth of 5 mm. Finally, 10 μl of the C6 cell suspension was injected into the caudate nucleus with a 0.5-ml microsyringe at a rate of 1 μl/min.

Therapy of the Glioma Model

Rat models that were established successfully were randomly divided into a treatment group and control group. Rats in the treatment group were administered with Bev (Roche, Shanghai, China) diluted with 0.9% physiological saline, while rats in the control group were administered with vehicle (0.9% physiological saline). The two groups received intraperitoneal injection of Bev or vehicle, respectively, at a dose of 5 mg/kg once daily for 7 days.

MRI Scanning

Prior to MRI, all rats were anesthetized with 3% isoflurane and continuous anesthesia with 1%–2% isoflurane using an anesthesia machine for small animals (Suzhou Zhongzhi Medical Technologies, China), and their core temperature maintained at 37°C. MRI scans were then performed using a 3.0-T MRI system (Discovery 750w, GE Medical System, Milwaukee, WI, USA) with a custom-built eight-channel receiver coil for animals (GE Medical System). Both the treated and control groups underwent an initial MRI scan before treatment (day 0) and then subsequent MRI sequences on days 1, 3, 5 and 7 after treatment with Bev or vehicle. MRI sequences included axial fast spin-echo T1-weighted imaging (T2W1) with the following parameters: repetition time/echo time (TR/TE) =

400 ms/9.5 ms; slice thickness = 2.0 mm; field of view (FOV) = 60 mm × 48 mm; and number of excitations (NEX) = 4.0. For axial spin echo T2-weighted imaging, the parameters included the following: TR/TE = 2,000 ms/46 ms; slice thickness = 2.0 mm; FOV = 60 mm × 48 mm; and NEX = 3.0. For IVIM DWI scans, single-shot echo-planar imaging DWI sequence was performed with nine b-values of 0, 10, 50, 100, 200, 400, 600, 800, and 1000 s/mm^2. Other parameters included TR/TE = 3000 ms/102.4 ms; flip angle, 90°; matrix = 64 × 64; FOV = 90 mm × 90 mm; section thickness = 2.0 mm; and NEX= 4.0, with a total scanning time of 6 min and 15 s.

Imaging Analysis

All MRI data were transferred to a post-processing workstation AW4.6 (GE Healthcare, USA), and tumor volume and IVIM-DWI parameter measurements were performed independently by one radiologist with 12 years' experience in MRI diagnosis.

For the calculation of the glioma volume, the region of interest (ROI) was drawn along the edge of tumor on T2WI images to obtain the surface area of the tumor in each slice (mm^2), after which the surface area of all the slices was summed and multiplied by the slice thickness (2 mm) to determine the total tumor volume (mm^3).

For the quantitative measurement of IVIM-DWI parameters, MRI data were analyzed using MADC software (GE Healthcare, USA) in the AW4.6 post-processing workstation. First, the IVIM-DWI parameter (apparent diffusion coefficient, ADC; diffusion coefficient, D; pseudodiffusion coefficient, D*; and perfusion fraction, f) maps were generated automatically using MADC software. Next, due to the low signal-to-noise ratio of IVIM-DWI image, the ROI was set in the section that demonstrated the largest tumor diameter according to both axial T2WI image and DWI image with a *b*-value of 1000 s/mm^2. The ROI (12–20 mm^3) was drawn manually along the edge of the tumor in order to cover as much of the solid part as possible and exclude necrotic and hemorrhagic areas. Finally, the average values of IVIM-DWI quantitative parameters were calculated automatically.

Histological Assessment

After the last MRI scan on the 7th day of treatment, all rats were sacrificed by cervical dislocation. Then, the brain tissue was removed, fixed with 4% paraformaldehyde, dehydrated with 70% alcohol, embedded with paraffin, and dissected into axial sections. To ensure that the tumor sections corresponded to IVIM-DWI images as accurately as possible, sections were processed sequentially at a thickness of 4 μm, in the same orientation as the axial MRI plane. Finally, hematoxylin and eosin staining was performed according to standard procedures.

To evaluate neoangiogenesis and cell proliferation of hypoxic glioma cells, sections were further analyzed by immunohistochemical staining using CD34 (1:250; Abcam), PCNA (PCNA, 1:100; Santa Cruz Biotechnology, CA, USA), and Hif1-α (1:100; Abcam) monoclonal antibodies. A pathologist with 10 years of experience reviewed the microvessel density (MVD) of CD34-positive cells and PCNA and Hif1-α expression. MVD calculation was determined by identifying three "hotspots" with dense positive cells in a low-power field (40×) and the number of CD34-positive cells was counted manually under a high-power field (200×); MVD was then calculated as the average number of the three "hotspots". Expression of PCNA and Hif1-α was assessed using ImageJ software (National Institutes of Health).

Statistical Analysis

All statistical processes were performed by SPSS 19.0 (IBM, Armonk, NY, USA). The tumor volume and quantitative IVIM-DWI parameters were expressed as the mean ± standard deviation. A one-way analysis of variance (ANOVA) with post-hoc test (Bonferroni test) was used to compare IVIM-DWI parameters among the five time points in the treatment and control group. An independent-samples *t*-test was used to compare IVIM-DWI parameters between the treatment and control groups. Pearson's analysis was used to analyze the correlation between IVIM-DWI parameters and the average tumor growth rate. Pearson's or Spearman's correlation tests were used to analyze the correlation between IVIM-DWI parameters and histological assessment in the treatment group. $p < 0.05$ indicated a significant difference.

RESULTS

Effect of Bev on Tumor Volume

Overall, 26 out of 30 rats underwent a successful glioma orthotopic procedure; two rats died due to intracranial hemorrhage, and no tumors were found in an additional two rats. (**Figure 1**). There was no bleeding, weight loss, or other serious side effects during the treatment course of Bev or vehicle in our study. The tumor volume of the control and treatment groups before treatment (day 0) was 23.4 ± 5.58 mm^3 and 25.2 ± 4.33 mm^3, respectively, indicating no significant difference between the two groups (p = 0.76). After treatment with Bev, the average tumor volume of the two groups differed significantly on day 7 (90.7 ± 13.4 mm^3 vs. 75.3 ± 11.5 mm^3, p = 0.022, **Figure 2A**). The relative change in tumor size was calculated as $\Delta\text{Volume} = (\text{volume}_n - \text{volume}_0)/\text{volume}_0 \times 100\%$. The ΔVolume in the control group was 29.1% ± 5.4% on day 1, 85.1% ± 13.5% on day 3, 200.5% ± 45.4% on day 5, and 287.6% ± 58.4% on day 7, while that of the treatment group was 27.7% ± 4.7% on day 1, 42.1% ± 12.3% on day 3, 138.8% ± 25.7% on day 5, and 198.8% ± 42.3% on day 7. The ΔVolume in the two groups differed significantly on days 5 and day 7 (p = 0.038 and $p <$ 0.001, respectively, (**Figure 2B**).

IVIM-DWI Parameter Images

The gliomas were generally located in the right cerebral hemisphere with scattered hemorrhage, and tumors, as visualized on T2WI and IVIM-DWI (b = 1,000 s/mm^2) images, were verified withs HE staining (**Figure 3**). T2WI image showed that tumors grew slowly from days 0 to day 7 in the treatment group and had indicative hyperintensities on DWI

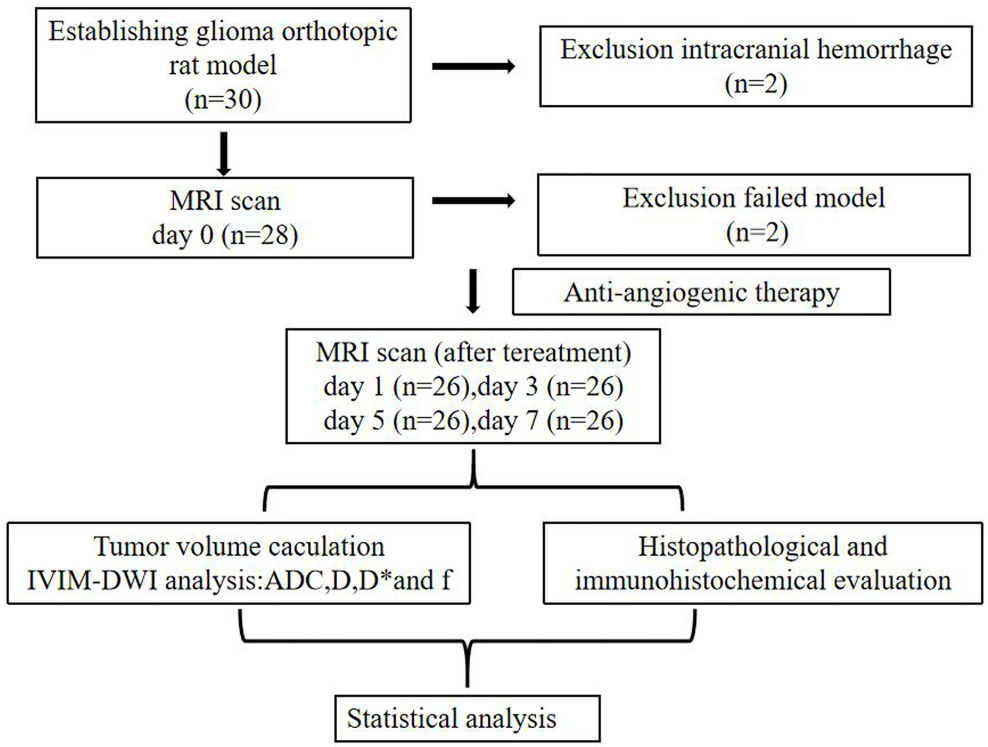


FIGURE 1 | The experiment flow chart of glioma model establishment, anti-angiogenic therapy, and MRI scan. ADC, apparent diffusion coefficient; D, diffusion coefficient; D*, pseudodiffusion coefficient; f, perfusion fraction.

images and hypointensities on ADC maps, with colors, ranging from blue to red, representing values ranging from low to high. Pseudocolor maps indicated that the D-values increased gradually while D*- and f-values decreased during the 7 days (**Figure 4**).

Intraobserver and Interobserver Agreement of IVIM-DWI Parameters

The reproducibility of ADC, D, and D* ranged from good to excellent and the reproducibility of f ranged from moderate to good. The interclass correlation coefficient of the IVIM-DWI parameters is summarized in **Table 1**.

IVIM-DWI Parameter Analyses

The dynamic changes of the IVIM-DWI parameters in the control and treatment groups at each time point are displayed in **Table 2**. ADC-value in the treatment group increased gradually while it decreased in the control group during the therapy course, with the two groups displaying significant differences among the five time points (treatment group: $F = 18.72$, $p < 0.001$; control group:

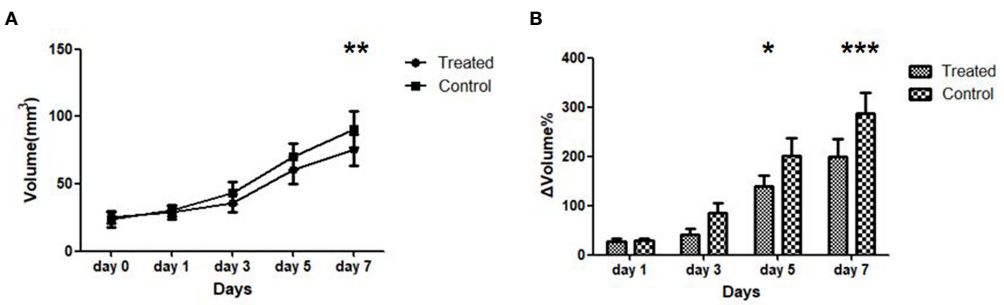


FIGURE 2 | Dynamic of tumor volume size **(A)** and average tumor growth rate ΔVolume **(B)** of the treated and control group during anti-angiogenic course. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Error bars denote standard errors.

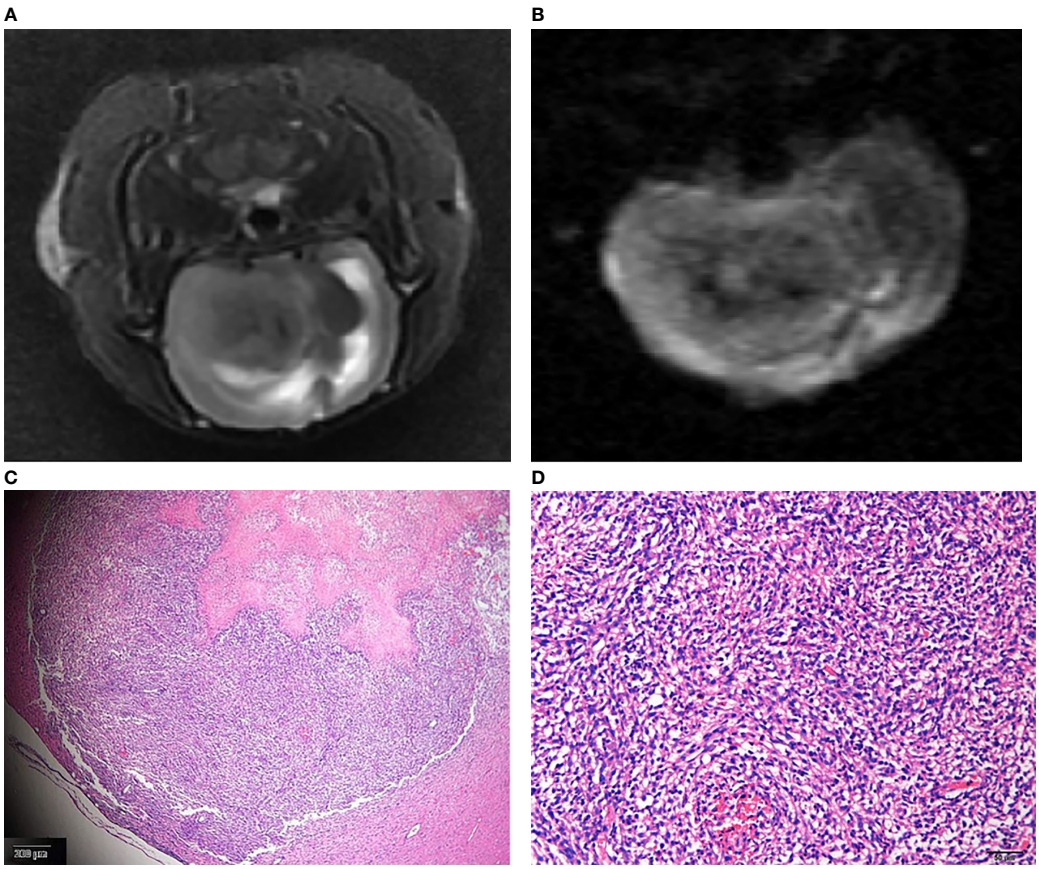


FIGURE 3 | The glioma on T2WI **(A)** and IVIM-DWI (b = 1,000 s/mm^2) **(B)** images. Corresponding HE staining showed atypical tumor cells (**C** original magnification ×4; **D** original magnification ×10).

$F = 7.98$, $p < 0.001$). Multiple comparisons between the two groups are shown in **Figures 5A, B**. For intergroup analyses, the baseline ADC-values were 0.585 ± 0.112 ×10^{-3} mm^2/s and 0.643 ± 0.290 × 10^{-3} mm^2/s in the control and treatment groups, respectively, which displayed no significant difference. However, the ADC-values in the treatment group were significantly higher than those in the control group from day 3 onwards (day 3: $p = 0.033$; day 5: $p < 0.001$; day 7: $p < 0.001$, **Figure 6A**).

The D-value in the treatment group increased daily while it significantly decreased in the control group. There was significant difference of D-value among the five time points in both treatment and control groups (treatment: $F = 33.17$, $p < 0.001$; control group: $F = 23.64$, $p < 0.001$). Comparison results between the two groups are shown in **Figures 5C, D**. For intergroup analyses, the baseline D-values were 0.274 ± 0.052 × 10^{-3} mm^2/s and 0.288 ± 0.050 ×10^{-3} mm^2/s in the control and treatment groups, respectively, which displayed no significant difference. Following daily treatment of Bev, the D-values in the treatment group were significantly higher compared with those in the control group on days 5 and day 7 (both $p < 0.001$, **Figure 6B**).

For the perfusion-related parameters, the D*-value of the treatment group decreased gradually and displayed significant differences among the five time points ($F = 11.08$, $p < 0.001$). By contrast, the D*-value of the control group increased markedly with time ($F = 4.23$, $p = 0.008$). The results of multiple comparisons between the two groups are shown in **Figures 5E, F**. After treatment, D*-values in the treatment group were significantly lower than those in the control group on days 3, 5, and 7 (day 3: $p = 0.005$; day 5: $p < 0.001$; day 7: $p < 0.001$, **Figure 6C**).

The f-value of the treatment group decreased slightly during the treatment course while that of the control group increased during the 7 days; however, there was no difference among the five time points in the two groups (treatment group: $F = 0.95$, $p = 0.45$; control group: $F = 2.08$, $p = 0.11$, **Figures 5G, H**). Only the f-value on the 7th day showed a significant difference between the two groups ($p = 0.008$, **Figure 6D**).

Correlation Between ΔVolume and IVIM Parameters

The initial D-value showed a moderate negative correlation with ΔVolume ($\gamma = -0.744$, $p < 0.001$), while the initial D*-value and relative change of D-value (ΔD) demonstrated moderate positive correlations with ΔVolume (D*: $\gamma = 0.718$, $p < 0.001$; ΔD: $\gamma = 0.800$, $p < 0.001$, **Figure 7**). No correlation between other initial or relative changes of IVIM parameters and tumor size were found.

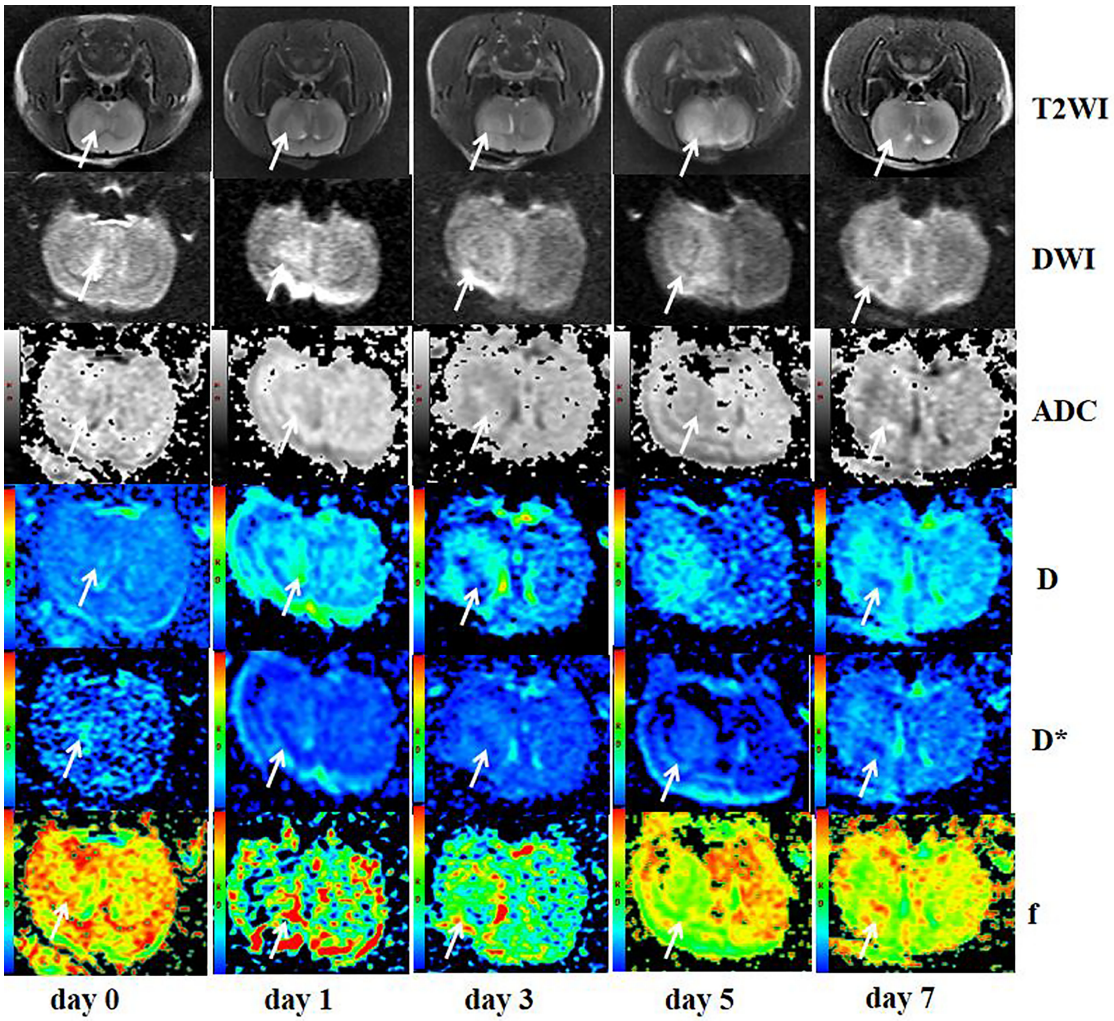


FIGURE 4 | Axial T2WI, DWI (b = 1000), ADC, and pseudocolor maps of D, D*, and f at different time points in the treated group. T2WI image showed the tumor grew slowly from day 0 to day 7 (arrows) and presented hyperintensity on DWI images and hypointensity on ADC maps (arrows). Color ranging from blue to red represented values ranging from low to high. D-values increased gradually (arrows), while D*- and f-values decreased during the 7 days (arrows).

TABLE 1 | The interobserver and intraobserver reproducibility of IVIM-DWI parameters.

	Interobserver	Intraobserver
ADC	0.944 (0.702–0.956)	0.903 (0.783–0.936)
D	0.778 (0.692–0.895)	0.895 (0.744–0.921)
D*	0.856 (0.603–0.881)	0.842 (0.689–0.911)
f	0.648 (0.497–0.785)	0.731 (0.558–0.802)

The 95% confidence intervals of interclass correlation coefficient (ICC) were in parentheses.

ADC, apparent diffusion coefficient; D, diffusion coefficient; D, pseudodiffusion coeffifficient; f, perfusion fraction.*

Histological Staining and Correlations With IVIM-DWI Parameters

CD34. Vascular endothelial cells were stained yellow and brown in CD34-stained sections. The CD34-positive vessels in the control group were dilated and distorted while they were tiny and regular in the treatment group, with a significant difference in the MVD between the two groups on day 7 ($p < 0.001$, **Table 3** and **Figure 8**). In addition, MVD was strongly correlated with D*-value ($r = 0.886$, $p = 0.019$, **Table 4**).

Regarding PCNA, the cell nuclei were stained brown in PCNA-stained sections. The score of PCNA-positive cells in the treatment group was markedly lower than that in the control group on day 7 ($p < 0.001$, **Table 2** and **Figure 8**). PCNA was

TABLE 2 | Comparison of IVIM-DWI parameters between the control and treated groups at different time points before and after antiangiogenic therapy.

	Day 0	Day 1	Day 3	Day 5	Day 7
ADC ($\times10^{-3}$ mm^2/s)					
Control	0.585 ± 0.112	0.577 ± 0.126	0.487 ± 0.130	0.450 ± 0.150	0.363 ± 0.140
Treated	0.643 ± 0.290	0.668 ± 0.271	0.733 ± 0.182	0.846 ± 0.278	0.968 ± 0.340
p	0.215	0.055	0.033	<0.001	<0.001
D ($\times10^{-3}$ mm^2/s)					
Control	0.274 ± 0.052	0.268 ± 0.040	0.192 ± 0.087	0.145 ± 0.058	0.086 ± 0.039
Treated	0.288 ± 0.050	0.265 ± 0.064	0.396 ± 0.150	0.499 ± 0.016	0.630 ± 0.290
p	0.730	0.798	0.052	<0.001	<0.001
D* ($\times10^{-3}$ mm^2/s)					
Control	5.797 ± 1.642	5.946 ± 1.620	7.233 ± 2.760	8.189 ± 2.830	9.061 ± 1.870
Treated	6.374 ± 1.730	5.758 ± 1.450	4.446 ± 1.709	3.475 ± 1.440	3.290 ± 1.102
p	0.505	0.810	0.005	<0.001	<0.001
f (%)					
Control	65.833 ± 11.502	59.837 ± 14.230	68.440 ± 22.003	69.164 ± 22.340	80.335 ± 28.037
Treated	69.525 ± 11.320	67.325 ± 12.530	59.355 ± 20.026	55.329 ± 24.703	44.858 ± 14.930
p	0.581	0.368	0.498	0.304	0.008

Comparisons of IVIM-DWI parameters between the control and treated groups at the same time point were performed by the independent-samples t-test.
ADC, apparent diffusion coefficient; D, diffusion coefficient; D, pseudodiffusion coeffifficient; f, perfusion fraction.*

strongly correlated with ADC-value and the D-value (ADC: $r = -0.848$, $p = 0.033$; D: $r = -0.928$, $p = 0.008$, **Table 4**).

Regarding the Hif-1α, both the nuclei and cytoplasm of cells were stained brown in Hif-1α-stained sections. Hif-1α-positive cells were mainly distributed at the edge of the tumor necrosis area. The Hif-1α-positive area in the treatment group was markedly smaller than that of the control group, and showed a significant difference ($p < 0.001$, **Table 2** and **Figure 8**). The score of Hif-1α was strongly correlated with D-value (r = −0.879, P = 0.010, **Table 4**).

DISCUSSION

Based on our previous study, the orthotopic C6 glioma model exhibited greater growth rate with less necrosis or hemorrhage area from day 14 to day 21 after C6 cell transplantation (12). Therefore, we started anti-angiogenesis therapy from the 14th day after establishing the glioma model and continued the therapy for 7 days to obtain the early dynamic changes in IVIM-DWI parameters.

Based on the dual-exponential model of IVIM-DWI, the perfusion-related parameters f and D* are influenced by the blood volume and velocity of the tumor (13). In our study, the f and D*-values demonstrated a pronounced reduction in the treatment group, which proved the effect of Bev-mediated inhibition of VEGF-dependent vascular growth. Besides, the positive correlation between the initial D* and ΔVolume indicated that tumors with high blood perfusion responded better to anti-angiogenic therapy. This result was essentially in agreement with previous studies that the D* and f-values of malignant tumors decreased significantly after anti-angiogenic therapy (14–16).

Our previous study of DCE-MRI data from the orthotopic C6 glioma model showed slightly increased K^{trans} and K_{ep} on day 1 after anti-angiogenic therapy, which may be the result of decreased leakage and increased circulation presented in the "normalization of tumor vasculature" theory (17, 18). However, the transient increase of perfusion-related MRI parameters was not found in the current IVIM-DWI research, suggesting that there was no direct correlation between IVIM-DWI perfusion-related parameters and DCE-MRI perfusion-related parameters. In fact, previous research has indicated that the perfusion parameters of DCE-MRI and IVIM-DWI were not completely consistent and possibly even contrasting. Indeed, IVIM-DWI provides information on the total blood transit through a voxel and is focused on the intravascular blood movement, while DCE-MRI provides information on the exchange of contrast agents between the intravascular and extravascular spaces (19, 20).

CD34-positive tubules are a pivotal feature of microvascular growth and were used to calculate the density of neovascularization in our study. Our study showed that D*-value had a significant positive correlation with MVD, which indicated that the perfusion-related IVIM parameters could reflect the change of blood flow accurately in glioma models following intervention of anti-angiogenic therapy. Besides, although previous research found that the f-value had more potential to predict the tumor response after antiangiogenic therapy, there was no explicit correlation between MVD and f in our present study, a result that was likely because of the obvious image noise of the f pseudo-color image (21).

The diffusion-related parameters ADC and D reflect the microscopic Brownian motion of water molecules in tumor tissue (22). The decreased tumor blood perfusion following Bev treatment resulted in a decrease in cell proliferation, and resulted in a significant increase in ADC and D-values from day 5. Together, the data indicated that the change of diffusion-related parameters occurred later than that of perfusion-related parameters. The degree of proliferation can be reflected by PCNA because it is strongly associated with DNA replication in tumor cells (23). The decrease of tumor cells after anti-angiogenic therapy was also confirmed by the significant differences in PCNA scores between the control and treatment groups on day 7. The negative correlation between diffusion-

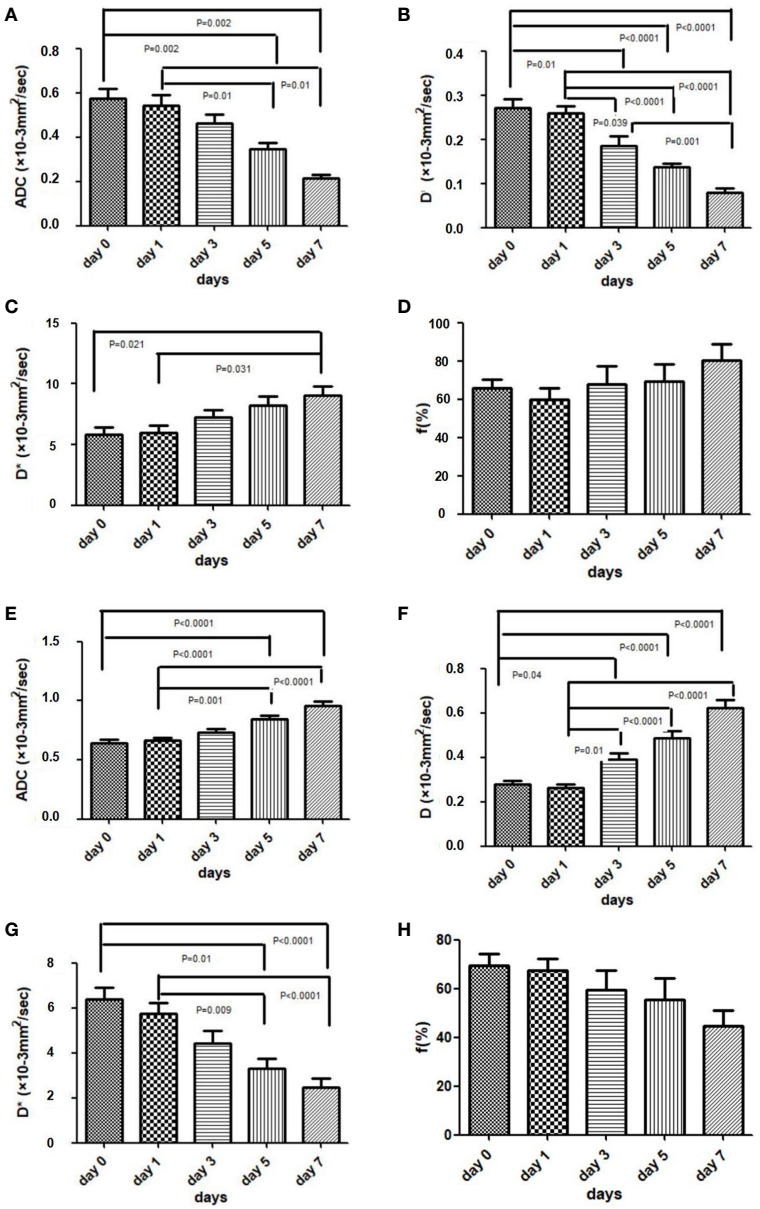


FIGURE 5 | Box plots comparing the IVIM-DWI parameters in the control groups **(A–D)** and treated group **(E–H)** at different time points. The four IVIM-DWI parameters between the two groups all showed opposite increased or decreased trends. ANOVA followed by the Bonferroni test was used in multiple comparison of different time points. Error bars denote standard errors.

related parameters and PCNA in the treatment group indicated that ADC and D-values could accurately reflect the cell density after anti-angiogenic therapy. The correlation between ADC-value and PCNA was weaker than that between D-value and PCNA, probably because the ADC-value was influenced by both microscopic water diffusion and blood perfusion of the tumor.

Hypoxia reduces the responsiveness of tumors to conventional chemotherapy and radiotherapy, and is an important factor of anti-angiogenic therapy failure for high-grade glioma (24). Previous research has confirmed that hypoxia is related to distorted neovascularization and proliferative cell density of tumors (25). In our study, we used Hif-1α as the pathological gold standard to evaluate the hypoxia of high-grade glioma and found that there was positive correlation between D-value and Hif-1α. The increased D-value indicated the reduced tumor cell density, which resulted in the decreased oxygen consumption of tumor cells. Our results confirmed that the diffusion-related parameter D-value could reflect the hypoxia of glioma after anti-angiogenic therapy. Unfortunately, we did not find any correlation between

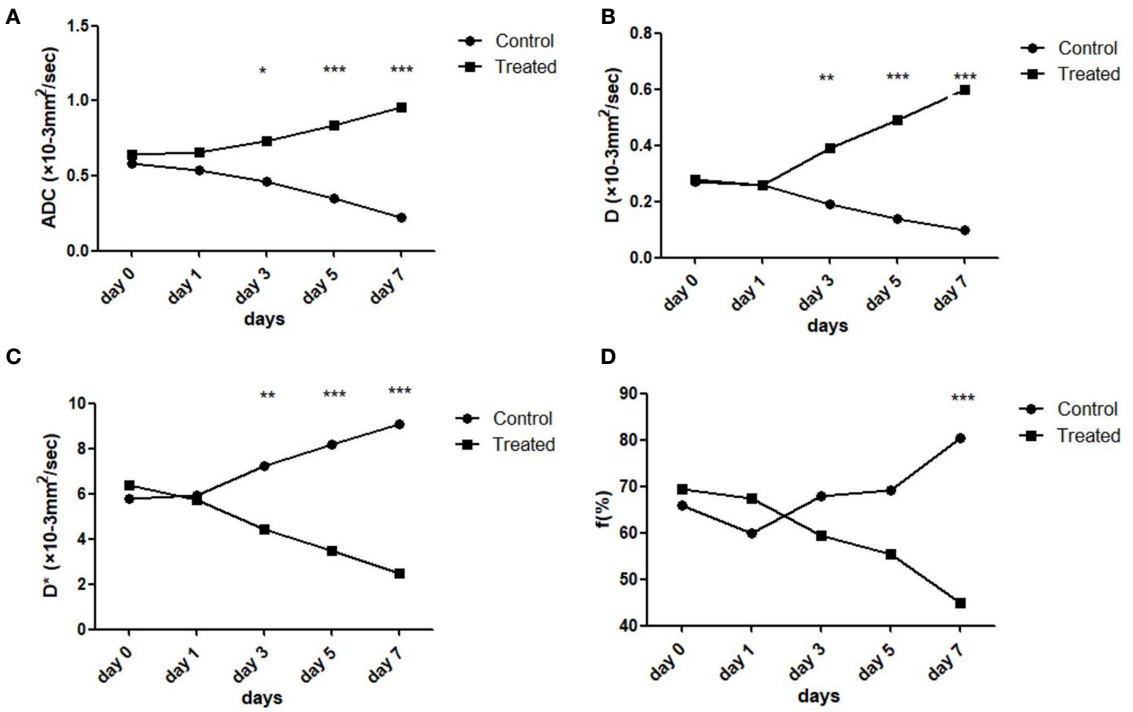


FIGURE 6 | Line graphs comparing the IVIM-DWI parameters of the treated and control group on different time points. The four IVIM-DWI parameters between the two groups all showed opposite trends during therapy **(A–D)**. The independent-samples *t*-test was used. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

perfusion-related parameters of IVIM-DWI and Hif-1α, which may be related to the complex changes in blood vessels after anti-angiogenic therapy.

There were some limitations in the study. First, because of the limited sample size, the correlation between IVIM-DWI parameters and histological assessment was only analyzed at the end of follow-up. Second, combined therapy with temozolomide was not used in this study because we aimed to explore the correlation between IVIM-DWI parameters and immunohistochemical indicators, rather than to confirm the mechanism of anti-angiogenic medicines. Thus, future studies should focus on investigating the effects of anti-angiogenic therapy combined with traditional chemotherapy in the same model using IVIM-DWI.

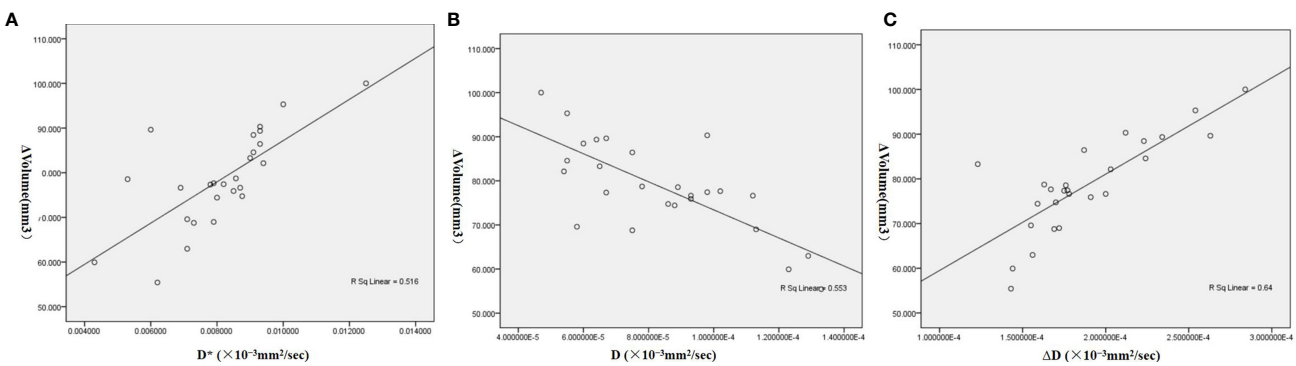


FIGURE 7 | Scatter diagrams of correlations between the relative change of tumor size (ΔVolume) and IVIM-DWI parameters. **(A)** The initial D-value showed moderate negative correlation with ΔVolume. **(B, C)** Initial D*-value and relative change of D-value (ΔD) demonstrated moderate positive correlation with ΔVolume. The Pearson correlation analysis was used.

TABLE 3 | The scores and statistical differences analysis of MVD, PCNA, and Hif1-α between the control and treated groups on day 7 after antiangiogenic therapy.

Groups	MVD	PCNA	Hif1-α
Control	38.890 ± 6.571	10.896 ± 2.326	5.677 ± 1.213
Treated	18.328 ± 4.256	4.875 ± 1.191	2.940 ± 0.046
p	<0.001	<0.001	<0.001

Comparisons of histological assessment between the control and treated groups on day 7 were performed using the independent-samples t-test.

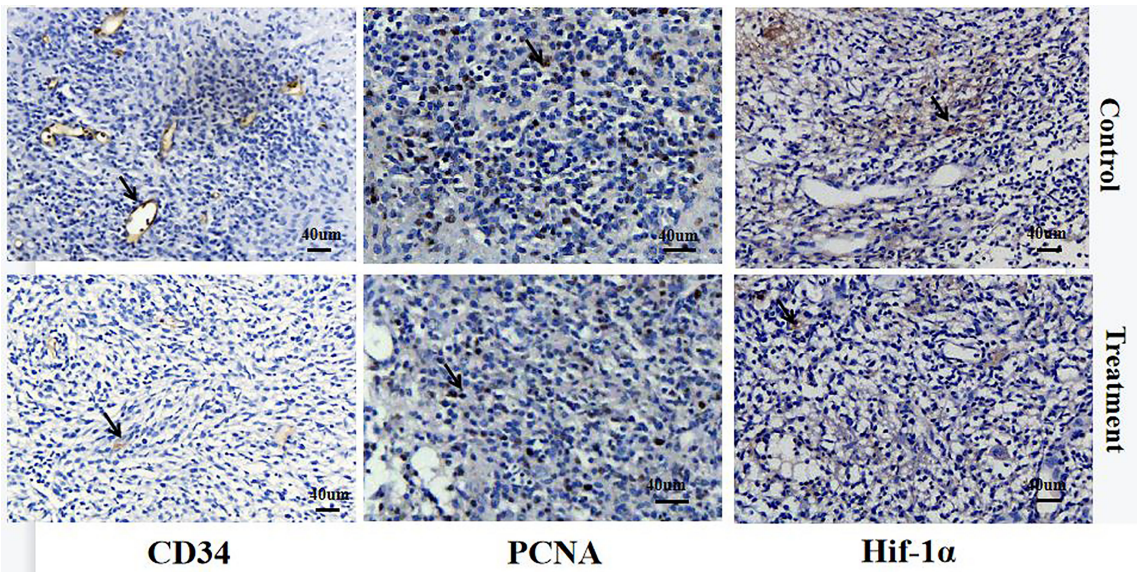


FIGURE 8 | IHC staining of CD34, PCNA, and Hif-1α in the control and treated group on day 7 after therapy. Positive cells are indicated with arrows. The CD34-positive vessels in the control group were dilated and distorted while they were tiny and regular in the treatment group. The PCNA-positive cells in the treatment group was lower than that in the control group. The Hif-1α-positive area in the treatment group was smaller than that of the control group.

TABLE 4 | Correlation coefficients between IVIM-DWI and immunohistochemistry scores of the treated group on day 7 after antiangiogenic therapy.

	MVD		PCNA		Hif-1α	
	Pearson's *r*	*p*	Spearman *r*	*p*	Spearman *r*	*p*
ADC	0.767	0.075	–0.848	0.033*	–0.278	0.594
D	0.527	0.283	–0.928	0.008*	–0.879	0.010*
D*	0.886	0.019*	0.866	0.019*	–0.626	0.052
f	0.802	0.055	–0.314	0.544	–0.463	0.355

*Pearson correlation was used for the comparison between MVD and IVIM-DWI parameters; Spearman's analysis was used for comparisons between PCNA, Hif-1α, and IVIM DWI parameters (*p < 0.05).*

CONCLUSIONS

In conclusion, IVIM-DWI showed decreased perfusion-related parameters and increased diffusion-related parameters were consistent with histological staining in the Bev-treated glioma model. The present study demonstrated that IVIM-DWI was sensitive and accurate in predicting and monitoring the effect of early anti-angiogenic therapy in the C6 glioma rat model.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The animal study was approved by the Ethics Committee for Animal Experimentation of Anhui Medical University of Anhui Province, China (Approval no. SCXK-Wan-2017–001) and was conducted in strict accordance with the Guidelines of the National Institutes of Health for the Care and Use of Laboratory Animals.

AUTHOR CONTRIBUTIONS

WH and YX contributed to conceptualization. YX, HP, and MX contributed to the cell and animal experiments. YQ, WH, and HP contributed to MRI scan and MRI data analysis. YS

contributed to pathological analysis. WH contributed to writing the original draft. XL and YY contributed to writing, reviewing, and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Natural Science Foundation of Anhui Province [No. 2008085QH381].

REFERENCES

1. Long Y, Tao H, Karachi A, Grippin AJ, Jin L, Chang YE, et al. Dysregulation of Glutamate Transport Enhances Treg Function That Promotes VEGF Blockade Resistance in Glioblastoma. *Cancer Res* (2020) 80(3):499–509. doi: 10.1158/0008-5472.CAN-19-1577
2. Seyedmirzaei H, Shobeiri P, Turgut M, Hanaei S, Rezaei N. VEGF Levels in Patients With Glioma: A Systematic Review and Meta-Analysis. *Rev Neurosci* (2020) 32(2):191–202. doi: 10.1515/revneuro-2020-0062
3. Tamura R, Tanaka T, Akasaki Y, Murayama Y, Yoshida K, Sasaki H. The Role of Vascular Endothelial Growth Factor in the Hypoxic and Immunosuppressive Tumor Microenvironment:Perspectives for Therapeutic Implications. *Med Oncol* (2019) 37(1):2. doi: 10.1007/s12032-019-1329-2
4. Kim MM, Umemura Y, Leung D. Bevacizumab and Glioblastoma: Past, Present, and Future Directions. *Cancer J* (2018) 24(4):180–6. doi: 10.1097/PPO.0000000000000326
5. Tamura R, Tanaka T, Miyake K, Yoshida K, Sasaki H. Bevacizumab for Malignant Gliomas: Current Indications, Mechanisms of Action and Resistance, and Markers of Response. *Brain Tumor Pathol* (2017) 34(2):62–77. doi: 10.1007/s10014-017-0284-x
6. Lyon KA, Huang JH. Bevacizumab as an Adjuvant Therapy for Glioblastoma in Elderly Patients: The Facts. *Transl Cancer Res* (2018) 7:S802–5. doi: 10.21037/tcr.2018.08.19
7. Payabvash S. Quantitative Diffusion Magnetic Resonance Imaging in Head and Neck Tumors. *Quant Imaging Med Surg* (2018) 8(10):1052–65. doi: 10.21037/qims.2018.10.14
8. Kong Z, Yan C, Zhu R, Wang J, Wang Y, Wang Y, et al. Imaging Biomarkers Guided Anti-Angiogenic Therapy for Malignant Gliomas. *NeuroImage Clin* (2018) 20:51–60. doi: 10.1016/j.nicl.2018.07.001
9. Qamar S, King AD, Ai QH, So TY, Mo FKF, Chen W, et al. Pre-Treatment Intravoxel Incoherent Motion Diffusion-Weighted Imaging Predicts Treatment Outcome in Nasopharyngeal Carcinoma. *Eur J Radiol* (2020) 129:109127. doi: 10.1016/j.ejrad.2020.109127
10. Federau C, Cerny M, Roux M, Mosimann PJ, Maeder P, Meuli R, et al. IVIM Perfusion Fraction is Prognostic for Survival in Brain Glioma. *Clin Neuroradiol* (2017) 27(4):485–92. doi: 10.1007/s00062-016-0510-7
11. Jabehdar Maralani P, Myrehaug S, Mehrabian H, Chan AKM, Wintermark M, Heyn C, et al. Intravoxel Incoherent Motion (IVIM) Modeling of Diffusion MRI During Chemoradiation Predicts Therapeutic Response in IDH Wildtype Glioblastoma. *Radiother Oncol* (2021) 156:258–65. doi: 10.1016/j.radonc.2020.12.037
12. Hou W, Xue Y, Tang W, Pan H, Xu M, Li X, et al. Evaluation of Tumor Hypoxia in C6 Glioma Rat Model With Dynamic Contrast-Enhanced Magnetic Resonance Imaging. *Acad Radiol* (2019) 26(8):e224–32. doi: 10.1016/j.acra.2018.09.011
13. Hectors SJ, Gordic S, Semaan S, Bane O, Hirten R, Jia X, et al. Diffusion and Perfusion MRI Quantification in Ileal Crohn's Disease. *Eur Radiol* (2019) 29 (2):993–1002. doi: 10.1007/s00330-018-5627-4
14. Liang J, Li Z, Li J, Peng C, Dai W, He H, et al. Application of IVIM-DWI in Detecting the Tumor Vasculogenic Mimicry Under Antiangiogenesis Combined With Oxaliplatin Treatment. *Front Oncol* (2020) 10:1376. doi: 10.3389/fonc.2020.01376
15. Song T, Yao Q, Qu J, Zhang H, Zhao Y, Qin J, et al. The Value of Intravoxel Incoherent Motion Diffusion-Weighted Imaging in Predicting the Pathologic Response to Neoadjuvant Chemotherapy in Locally Advanced Esophageal Squamous Cell Carcinoma. *Eur Radiol* (2021) 31(3):1391–400. doi: 10.1007/s00330-020-07248-z
16. Martens RM, Koopman T, Lavini C, Ali M, Peeters CFW, Noij DP, et al. Multiparametric Functional MRI and ^{18}F-FDG-PET for Survival Prediction in Patients With Head and Neck Squamous Cell Carcinoma Treated With (Chemo) Radiation. *Eur Radiol* (2021) 31(2):616–28. doi: 10.1007/s00330-020-07163-3
17. Hou W, Li X, Pan H, Xu M, Bi S, Shen Y, et al. Dynamic Contrast-Enhanced Magnetic Resonance Imaging for Monitoring the Anti-Angiogenesis Efficacy in a C6 Glioma Rat Model. *Acta Radiol* (2020) 61(7):973–82. doi: 10.1177/0284185119887598
18. De Palma M, Biziato D, Petrova TV. Microenvironmental Regulation of Tumour Angiogenesis. *Nat Rev Cancer* (2017) 17(8):457–74. doi: 10.1038/nrc.2017.51
19. Luo J, Zhou K, Zhang B, Luo N, Bian J. Intravoxel Incoherent Motion Diffusion-Weighted Imaging for Evaluation of the Cell Density and Angiogenesis of Cirrhosis-Related Nodules in an Experimental Rat Model: Comparison and Correlation With Dynamic Contrast-Enhanced MRI. *J Magn Reson Imaging* (2020) 51(3):812–23. doi: 10.1002/jmri.26845
20. Hectors SJ, Wagner M, Besa C, Bane O, Dyvorne HA, Fiel MI, et al. Intravoxel Incoherent Motion Diffusion-Weighted Imaging of Hepatocellular Carcinoma: Is There a Correlation With Flow and Perfusion Metrics Obtained With Dynamic Contrast-Enhanced MRI? *J Magn Reson Imaging* (2016) 44(4):856–64. doi: 10.1002/jmri.25194
21. Wan Q, Bao Y, Xia X, Liu J, Wang P, Peng Y, et al. Intravoxel Incoherent Motion Diffusion-Weighted Imaging for Predicting and Monitoring the Response of Anti-Angiogenic Treatment in the Orthotopic Nude Mouse Model of Lung Adenocarcinoma. *J Magn Reson Imaging* (2021). doi: 10.1002/jmri.27920
22. Wu L, Xu P, Rao S, Yang L, Chen C, Liu H, et al. ADC_{total} Ratio and D Ratio Derived From Intravoxel Incoherent Motion Early After TACE are Independent Predictors for Survival in Hepatocellular Carcinoma. *J Magn Reson Imaging* (2017) 46(3):820–30. doi: 10.1002/jmri.25617
23. Lv Q, Zhang J, Yi Y, Huang Y, Wang Y, Wang Y, et al. Proliferating Cell Nuclear Antigen Has an Association With Prognosis and Risks Factors of Cancer Patients: A Systematic Review. *Mol Neurobiol* (2016) 53(9):6209–17. doi: 10.1007/s12035-015-9525-3
24. Boyd NH, Tran AN, Bernstock JD, Etminan T, Jones AB, Gillespie GY, et al. Glioma Stem Cells and Their Roles Within the Hypoxic Tumor Microenvironment. *Theranostics* (2021) 11(2):665–83. doi: 10.7150/thno.41692
25. Hu M, Zhu Y, Mu D, Fan B, Zhao S, Yang G, et al. Correlation of Hypoxia as Measured by Fluorine-18 Fluoroerythronitroimidazole (^{18}F-FETNIM) PET/CT and Overall Survival in Glioma Patients. *Eur J Nucl Med Mol Imaging* (2020) 47(6):1427–34. doi: 10.1007/s00259-019-04621-z

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Hou, Xue, Qian, Pan, Xu, Shen, Li and Yu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 28 January 2022
doi: 10.3389/fonc.2021.816672

AttR2U-Net: A Fully Automated Model for MRI Nasopharyngeal Carcinoma Segmentation Based on Spatial Attention and Residual Recurrent Convolution

Jiajing Zhang[1], *Lin Gu*[2], *Guanghui Han*[1,3]* *and Xiujian Liu*[1]*

[1] School of Biomedical Engineering, Sun Yat-sen University, Shenzhen, China, [2] RIKEN Center for Advanced Intelligence Project (AIP), Tokyo, Japan, [3] School of Information Engineering, North China University of Water Resources and Electric Power, Zhengzhou, China

OPEN ACCESS

Edited by:
Guang Yang,
Imperial College London,
United Kingdom

Reviewed by:
Xu Chuangye,
Southern University of Science and Technology, China
Jiali Cui,
North China University of Technology, China
Zhenyu Tang,
Beihang University, China

***Correspondence:**
Xiujian Liu
liuxj86@mail.sysu.edu.cn
Guanghui Han
hangh3@mail.sysu.edu.cn

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 17 November 2021
Accepted: 17 December 2021
Published: 28 January 2022

Citation:
Zhang J, Gu L, Han G and Liu X (2022) AttR2U-Net: A Fully Automated Model for MRI Nasopharyngeal Carcinoma Segmentation Based on Spatial Attention and Residual Recurrent Convolution. Front. Oncol. 11:816672. doi: 10.3389/fonc.2021.816672

Radiotherapy is an essential method for treating nasopharyngeal carcinoma (NPC), and the segmentation of NPC is a crucial process affecting the treatment. However, manual segmentation of NPC is inefficient. Besides, the segmentation results of different doctors might vary considerably. To improve the efficiency and the consistency of NPC segmentation, we propose a novel AttR2U-Net model which automatically and accurately segments nasopharyngeal carcinoma from MRI images. This model is based on the classic U-Net and incorporates advanced mechanisms such as spatial attention, residual connection, recurrent convolution, and normalization to improve the segmentation performance. Our model features recurrent convolution and residual connections in each layer to improve its ability to extract details. Moreover, spatial attention is fused into the network by skip connections to pinpoint cancer areas more accurately. Our model achieves a DSC value of 0.816 on the NPC segmentation task and obtains the best performance compared with six other state-of-the-art image segmentation models.

Keywords: nasopharyngeal carcinoma, tumor segmentation, deep learning, spatial attention, recurrent convolution, residual connection

1 INTRODUCTION

Nasopharyngeal cancer is a common malignant tumor occurring in the top and sidewalls of the nasopharyngeal cavity, with 833,019 new cases of nasopharyngeal cancer and 468,745 deaths in China alone during 2015 (1). Nasopharyngeal cancer affects a wide range of areas, from the nasal cavity forward to the conus, down to the oropharynx, and up to the skull. Moreover, it is mainly located in the central part of the head (2), making it difficult to treat with common surgical treatments. As such, the current mainstream treatment is radiotherapy. The lesion segmentation is one of the most critical factors affecting the effectiveness of radiotherapy.

However, traditional manual segmentation (**Figure 1A**) has three main drawbacks. First, the segmentation still currently relies on specialized physicians to manually segment nasopharyngeal

carcinoma. However, manual segmentation is very time-consuming, taking at least 3 hour to complete the segmentation of one single patient (3, 4). Second, as the radiotherapy process progresses, the area of nasopharyngeal cancer keeps changing, making it necessary to re-identify the lesion. This undoubtedly increases the burden of doctors. Third, the effectiveness and quality of manual segmentation may vary significantly among physicians. Studies have shown that the segmentation area of nasopharyngeal cancer might vary from one physician to another for the same case (5, 6). The disagreement among physicians increases the complexity and uncertainty of the treatment. Thus, there is an urgent need for an automated and efficient method for accurate nasopharyngeal carcinoma segmentation.

However, automatic segmentation of nasopharyngeal carcinoma faces challenges from two main aspects, as shown in **Figure 1B**. First, the shape and size of the cancerous area are both very complex (7). Because of its invasive and aggressive nature (8), it can spread to the parapharynx, skull base, and even intracranial region. The size and shape of the cancerous area can vary considerably from patient to patient or from one patient's different stages of the disease. Moreover, nasopharyngeal cancer occurs primarily in and around the nasopharynx, in the center of the head. The MR imaging of nasopharyngeal carcinoma may include a wide range of tissues near the cancerous area, such as the pharyngeal recess, pharyngeal orifice, pharyngeal tonsils, sphenoidal sinus, superior, middle, and inferior nasal concha, and hard and soft palate. Therefore, it is hard to distinguish nasopharyngeal cancer from these peripheral tissues, even for specialized doctors. Second, unlike most computer vision tasks with similar image quality (9, 10), the MR imaging quality may vary significantly for the same nasopharyngeal cancer. The variation of the image can be a result of various factors (1). MR imaging of nasopharyngeal cancer is highly dependent on imaging equipment (2). The imaging quality of the same equipment can be affected by the differences in physician's operation, contrast agent variation, and patient status.

In the last decade, deep learning methods have been increasingly used in computer vision, such as target detection (11, 12), salient object detection (13, 14), video analysis (15, 16), and semantic segmentation (17), especially medical image segmentation (18). For example, Dey et al. propose a hybrid cascaded neural network for liver lesion segmentation (19), Singh et al. use a generative adversarial and convolutional neural network to achieve breast tumor segmentation in mammograms (20), Conze et al. use cascaded convolutional and adversarial deep networks to achieve abdominal multiorgan segmentation (21) and use deep convolutional encoder–decoders to achieve shoulder muscle MRI segmentation (22). Many studies show that deep learning methods can replace manual segmentation in medical image segmentation tasks and achieve satisfactory segmentation results while saving time and costs.

The most widely used structure among image segmentation models is the U-Net (23). It consists of an encoding module and a decoding module. The encoding module uses deep convolutional neural networks to gradually extract the features of the input image from local to global. The decoding module uses upsampling to recover the image resolution. Moreover, the results of each decoding layer are fused with the results obtained from the encoding layer by skip connection. However,

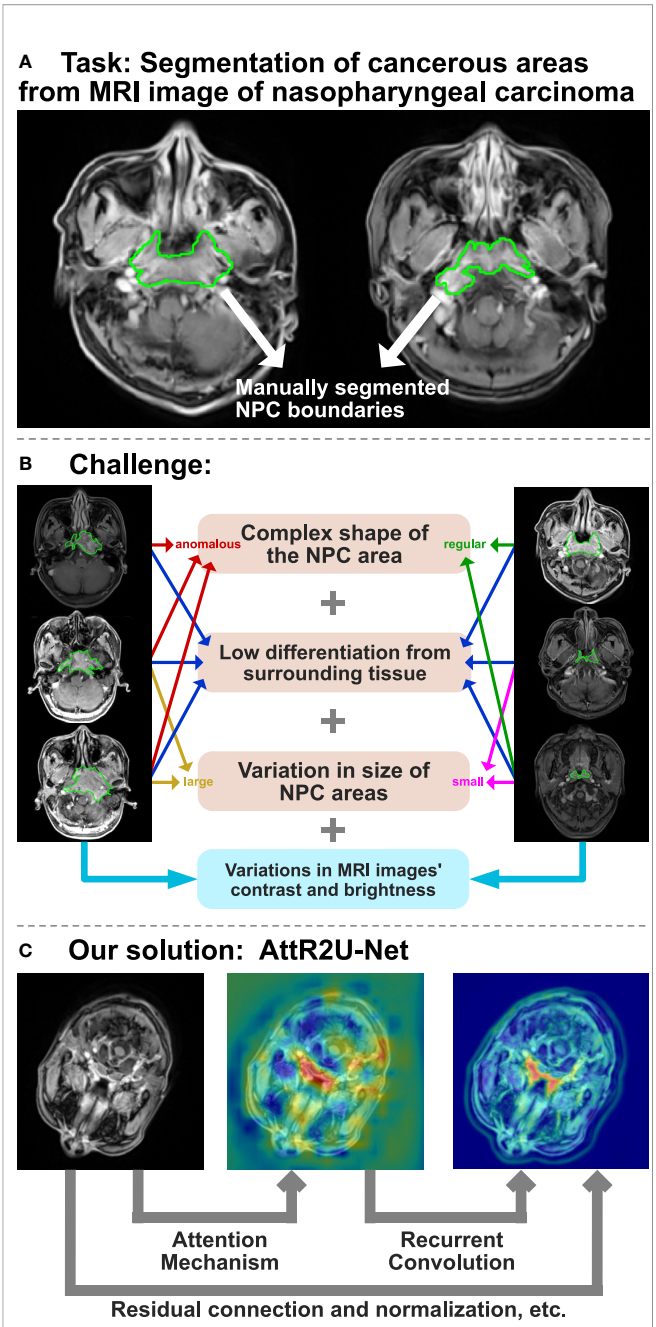


FIGURE 1 | The overview of AttR2U-Net model for automatic segmentation of nasopharyngeal carcinoma. **(A)** The task of the AttR2U-Net model. **(B)** The challenges of segmentation arise from two main aspects, in which orange represents NPC internal causes and blue represents external cause. **(C)** The four main mechanisms of our solution: attention mechanism to localize NPC accurately, recurrent convolution to extract boundary details precisely, residual connection to build a deeper and more efficient model, and normalization to reduce the variations in MRI images.

experiments find that the ordinary U-net structure seems incompetent to segment complex tumors as nasopharyngeal carcinoma. It frequently fails to determine the location of nasopharyngeal cancer, especially when the cancerous area exists only in one side of the patient's head. In addition, the model tends to include the surrounding tissues with similar brightness in the identified cancerous area. As such, the results of previous studies that use the classical U-Net structure to segment nasopharyngeal carcinoma have relatively low precision rates in NPC segmentation tasks. For example, Li et al. (24) only achieve a DSC value of 0.74.

Based on the U-Net structure, we propose AttR2U-Net and integrate several advanced deep learning method (**Figure 1C**), including spatial attention, residual connection, recurrent convolution, and normalization. Compared with general U-Net models, our method is novel with the following four main advantages:

- Our model introduces the attention mechanism so that it can learn the importance offeatures at different spatial locations (25).
- Our model uses the recurrent convolution (26) instead of the standard convolution to enhance the feature extraction capability.
- Our model adds residual connections to the neural network. It allows us to efficiently train deeper networks and solve the network degradation problem (27).
- Our model applies normalization to all the input. It solves the problem of contrast differences between different MRI images.

2 ATTR2U-NET FOR NPC SEGMENTATION

The AttR2U-Net model proposed is shown in **Figure 2**. We incorporate three advanced computer vision methods, spatial attention, recurrent convolution, and residual connection, into the general U-Net through a sophisticated structural design.

The AttR2U-Net model consists of two parts: an encoder and a decoder. It first normalizes the input nasopharyngeal carcinoma image to solve the problem of contrast differences between MRI images, improving the model's efficiency. The encoder with the residual connection and recurrent convolution gradually reduces the image's resolution. The decoder implements upsampling through deconvolution with residual connection and recurrent convolution. The encoder can accurately extract local low-level and global high-level semantic information layer by layer, whereas the decoder recovers the high-resolution nasopharyngeal carcinoma segmentation map.

Additionally, by incorporating spatial attention in the feature fusion process, our model fuses the feature map from each layer of the encoder with the feature map from the decoder at the corresponding scales after passing through the attention block. Thus, our model can preserve both global macro-semantic information and local micro-detail in the upsampling process. Finally, the decoding module upsamples the feature map to the exact resolution as the input image and outputs it as the segmentation map.

2.1 Attention Mechanism

The attention mechanism aims to extract the different importance of each part of the input content, so that the model can focus on the

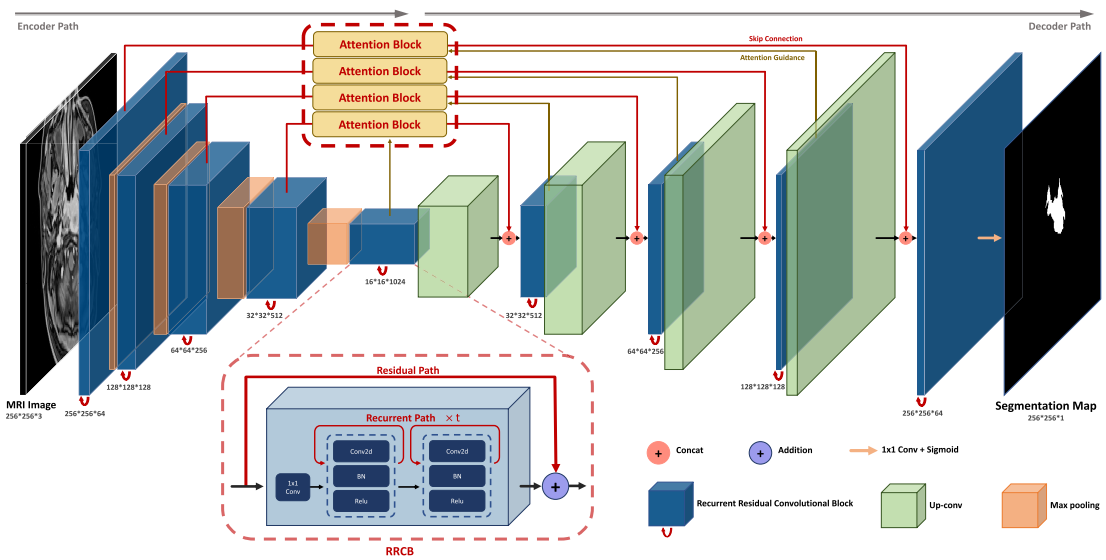


FIGURE 2 | The architecture of AttR2U-Net: The encoder in AttR2U-Net consists of five Recurrent Residual Convolutional Blocks (RRCB). Each RRCB consists of a 1×1 convolution and two Recurrent Convolutional Layers, which loop Conv2d+BN+ReLU t times. Our model adds the input and output of each RRCN to achieve the residual connection. After each RRCB layer, our model sets a 2 × 2 max pooling layer to reduce the scale of the image by half. The decoder is symmetrically composed of four RRCBs, which are respectively connected with the first, second, third, and fourth encoding layer through the attention mechanism. The attention map is fed to the RRCB after concatenating with the upsampling output. After each RRCB, our model places an Up-Conv layer, including an upsampling layer and a Conv2d+BN+ReLU operation. Then output the segmentation result after a 1 × 1 convolution and a sigmoid layer.

needed parts of the input and ignore the less relevant parts. DeepMind firstly proposes attention for image classification (28). Then, it is widely used in various deep learning fields in recent years, including natural language processing (29, 30), image classification (31), video classification (32), and emotion recognition (33). Moreover, there are continuous studies to apply attention in deep convolutional neural networks (34). It results in a number of studies using attention for image segmentation, such as DANet (35) which applies self-attention on the spatial domain and channel domain, OCNet (36) which applies a semantic aggregation strategy, and PANet (37) which applies pyramidal structure attention on layer domain.

Inspired by Attention U-Net (38), our model implements spatial attention by adding attention blocks to the general U-net structure for the high-precision nasopharyngeal carcinoma segmentation.

Our model inserts the attention block in the skip connection connecting the encoder and decoder. The input of each attention block is the feature map $X_{e(i)}$ output from the encoder on the corresponding downsampling scale and the feature map $X_{d(i\text{-}1)}$ output from the upper layer in the decoder. The attention block first linearly transforms them using channel-wise 1×1 convolutions with the parameters of $W_{e(i)}$ and $W_{d(i\text{-}1)}$, where $W_{e(i)}$ represents the parameters of the convolution to the input from the encoder and $W_{d(i\text{-}1)}$ represents the parameters of the convolution to the attention guidance from the decoder. Our model calculates the attention coefficients α_i as

$$\lambda_i = W_{e(i)}^T X_{e(i)} + W_{d(i-1)}^T X_{d(i-1)} + b_{\lambda_i} \quad (1)$$

$$\mu_i = \Psi^T(\mathrm{Re}\,LU(\lambda_i)) + b_{\mu_i} \quad (2)$$

$$\alpha_i = \frac{1}{1 + e^{-\mu_i}} \quad (3)$$

where i refers to the level of the encoder or decoder, Ψ^T corresponds to a 1×1 convolution, λ_i denotes the result of a linear transformation of input $X_{e(i)}$ and attention guidance $X_{d(i\text{-}1)}$, and μ_i is the intermediate variable we define. Our model defines the attention coefficients as the output of a sigmoid layer α_i and $\alpha_i \in [0, 1]$. According to the formula, our model utilizes more accurate semantic information from the higher level to guide the extraction of the attention region from the lower-level feature map. The output of the attention block is the attention coefficients and element-wise product of the input feature map. We visualize attention coefficients to the attention map, as shown in **Figure 3**.

Therefore, the output extracts the vital part of the input feature map for the segmentation while weakening its irrelevant part. In addition, our model uses soft attention. Its main feature is differentiable, making it easier to learn by backpropagation.

2.2 Recurrent Convolutional Block

The recurrent convolution is a commonly used method in text classification (39). Some studies apply it to the field of computer vision, such as Liang et al. (27) who use it in the Object Recognition and Alom et al. (40) who use it in the image segmentation. Inspired by the above studies, our model adds the recurrent convolution block to make it applicable for nasopharyngeal cancer segmentation.

In our model, recurrent convolutional layers replace standard convolutional layers so that our model can keep deepening and integrating contextual information on discrete time steps. It effectively improves the ability to extract detailed features. By setting the total time step parameter to t, the Conv+BN+ReLU operation repeats t times in each recurrent convolutional block. The input for each convolution is defined as follows:

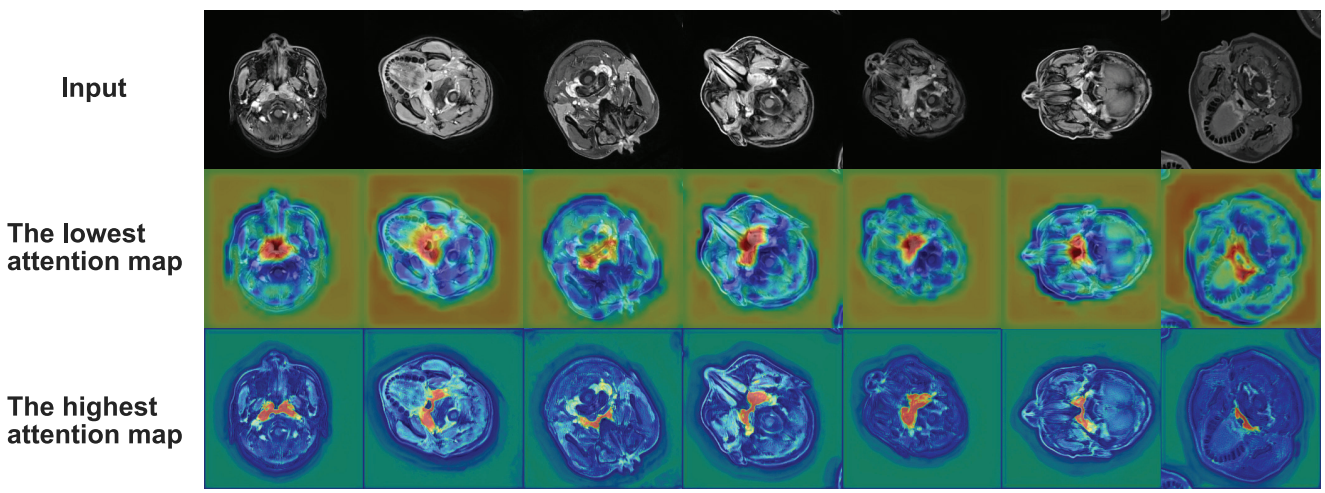


FIGURE 3 | Radar plot showing the comparison of evaluation metrics obtained from each model in the comparison experiment. The five vertices of the radar plot respectively represent the average DSC value, average sensitivity, average specificity, average precision, and average Jaccard similarity coefficient. Note that this radar plot uses percentages to express the above evaluation metrics to make the comparison more distinguishable, as shown in equation 11.

$$I(t) = (W_f)^T X + (W_r)^T O(t-1) + b \quad (4)$$

showing that the input $I(t)$ consists of the summation of the original input X and the output $O(t-1)$ of the previous convolution. W_f and Wr refer to the weights of the feedforward path and recurrent path, and b is the bias. In addition, it is essential to note that X remains constant throughout the iteration.

In our Recurrent Convolutional Block, each loop contains the following operations:

$$y_{1,\tau} = (W_{1,\tau})^T x + b_{1,\tau} \quad (5)$$

$$y_{2,\tau,i} = \gamma \times \frac{y_{1,\tau,i} - \mu_{y_{1,\tau}}}{\sqrt{\sigma_{1,\tau}^2 + e^{-5}}} + \beta \quad (6)$$

where τ refers to the current time step, $W_{1,\tau}$ refers to the weight of an ordinary convolution, $b_{1,\tau}$ refers to the bias, $\mu_{y_{1,\tau}}$ refers to the mini-batch mean of $y_{1,\tau}$, $\sigma_{1,\tau}^2$ refers to its mini-batch variance, and γ and β are the two parameters to be learned. The second equation actually completes a batch normalization operation on $y_{1,\tau}$.

In general, at each time step t, the model does a convolution on the input $I(t)$. Then, the batch normalization (BN) and the activation function (ReLU) were applied to get the output $O(t)$ of that time step and use it as one of the input terms for the next time step.

2.3 Residual Connection

Residual connection (26) is a well-known deep learning method. It represents the final output as a linear superposition of its input and a non-linear transformation of its original output. It combines them by direct summation. Many studies show that the application of residual connection in deep neural networks can effectively solve the network degradation problem (41, 42) and the shattering gradient problem (43) during the backpropagation of training. Besides, it can also make the training more accessible.

Let us assume that the mathematical representation of the residual connection of the ith layer in the τ-th time step is as:

$$x_{i+1,\tau} = F[\alpha_{i,\tau} x_{i,\tau} + f((W_{i,\tau})^T x_{i,\tau} + b_{i,\tau})] \quad (7)$$

where $x_{i,\tau}$ is the input of the ith layer in the τth time step, α refers to the modulating scalar, $W_{i,\tau}$ refers to the weight of the recurrent convolution of the i-th layer in the τ-th time step, and $b_{i,\tau}$ refers to its bias. f refers to the operation after recurrent convolution in each time step, which denotes the batch normalization and the ReLU activation function. F refers to the operation after each time step. Here, F is set to identity. Thus, when introducing the residual connection, our model defines a linear summation of the convolutional output and input after modulating as a new input for the next layer.

As the number of layers increases, the residual connections keep recurring; the residual connections on layer I relative to the previous layer i can be expressed as:

$$x_{I,i,\tau} = (\prod_{k=i}^{I-1} \alpha_k) x_{i,\tau} + \sum_{k=i}^{I-1} \theta_{I,k,\tau} \quad (8)$$

$$\theta_{I,k,\tau} = (\prod_{h=k+1}^{I-1} \alpha_h) f((W_{k,\tau})^T x_{k,\tau} + b_{k,\tau}) \quad (9)$$

where $x_{i,\tau}$ refers to the input of the i-th layer in the τ-th time step, α refers to the modulating scalar, $W_{k,\tau}$ refers to the weight of the recurrent convolution of the kth layer in the τth time step, and $b_{k,\tau}$ refers to its bias. f refers to the operation after recurrent convolution in each time step. Moreover, $\theta_{I,k,\tau}$ is the intermediate variable we define. Finally, $x_{I,i,\tau}$ is the input of layer I with residual connections.

2.4 Normalization Method

Since the MRI data come from multiple batches of different MR scans, and there are differences in contrast agents, physician operations, etc., the brightness and contrast between images vary significantly. For example, the brightness of some images may be very dark, in which nasopharyngeal cancer appears as dark areas, while the brightness of other images may be very bright, with high contrast, in which nasopharyngeal cancer appears as bright areas, as shown in **Figure 1B**. It often requires precise adjustment during manual segmentation. In order to minimize the additional difficulties caused by data variation for training, our model normalizes all the input data.

Several previous studies also point that normalization can make the data more uniformly distributed. It can effectively improve the speed of solving the optimal solution by gradient descent, making it easier for the model to converge and potentially improve the model performance. Therefore, our model applies the Z-score normalization to all images based on all original pixels' mean and standard deviation. It is also known as standardization. Images after processing have a mean of 0 and a variance of 1 conforming to the standard normal distribution.

2.5 Implementation Details

In the training process of AttR2U-Net, our model uses the Adam optimizer (44) to implement the gradient descent method and set the parameters β_1 and β_2 to 0.9 and 0.999, respectively. The learning rate is dynamically adjusted using the learning rate decay strategy. The model also uses L2 regularization (45), dropout (46), and other mechanisms to solve the overfitting problem to some extent. We train the model with the shuffled nasopharyngeal carcinoma images with segmentation labels. Moreover, we promptly test the model on the validation set at the end of each training epoch to adjust the parameters.

3 EXPERIMENT

3.1 Materials

3.1.1 Data Acquisition

Our model uses MRI images from a total of 93 patients diagnosed with nasopharyngeal carcinoma. The patients are scanned by the Siemens Aera MRI system. The resulting MRI images are stored in Digital Imaging and Communications in Medicine (DICOM) file format. Specifically, all MRI data are T1-weighted, and contrast agents are used during the imaging process. The scanning area

includes the entire head and neck region (approximately 100 slices), resulting in T1W+C images with a voxel spacing of $0.71875 \times 0.71875 \times 3mm^3$. We annotate the MRI images of each patient with the nasopharyngeal cancer boundary at the pixel level according to their pathological characteristics.

3.1.2 Data Preprocessing

To utilize the raw data for 2D image segmentation, we extract 956 slices containing the nasopharyngeal cancer region from the DICOM files and transform them into 2D images. However, the raw images contain a large amount of black backgrounds. Our model crops each image to the region of interest (ROI) to reduce unnecessary computing workload. The cropped image includes the nasopharyngeal carcinoma area and the rest of the head region. Then our model resizes them to 256 × 256. Our model applies data augmentation methods, including rotation, mirroring, and affine transformation. In the end, our dataset includes 4,775 nasopharyngeal carcinoma images and their segmentation labels. 80% of the data are used for the training set, 10% are reserved as test set, and 10% are reserved as validation set.

3.2 Evaluation Method

In the training process, our model uses the most commonly used binary cross-entropy loss function, also known as BCELoss, to evaluate the training effect. The loss value L is defined as:

$$L = -\frac{1}{N}\sum(y_n ln(p_n) + (1 - y_n) ln(1 - p_n)) \quad (10)$$

where N refers to the total number of pixels, y_n denotes whether the nth pixel belongs to the NPC region (if yes, then $y_n = 1$), and p_n denotes the probability that the pixel belongs to the NPC region according to our model's prediction.

In the testing process, our model uses the test set of nasopharyngeal carcinoma MRI images (478 images in total) to evaluate the segmentation effect. To quantify the segmentation performance of our model, DSC value (47), Jaccard similarity (48), precision, specificity, sensitivity, and PR curve are used. In addition, we also perform patient-wise 5-fold cross validation experiments. We discuss our model performance in *Results*.

3.3 Comparison Method

In this study, we thoroughly review the field of computer vision and identify six advanced image segmentation models: SEUNet (49), FCDenseNet (50), NestedUNet (51), DeepLabV3 (52), DANet (35), and FCN (53). We use the official model of the corresponding papers and use the optimal parameters suggested by the authors. We then train them with the same training set for nasopharyngeal carcinoma segmentation and keep the same maximum number of training epochs for each model. For each comparison model, we select the best-performing trained model from the training process.

In the testing process, we use the same nasopharyngeal carcinoma segmentation test set to evaluate each model and calculate five evaluation metrics (the same as those used by AttR2U-Net) and plot PR curves. For demonstration, we also select seven typical MRI images of nasopharyngeal carcinoma as inputs and compare the output of each model, which are then compared with the manual segmentation result.

4 RESULTS

4.1 Performance of Our Model

The results show that our model is superior to other models for nasopharyngeal carcinoma segmentation. The test results of our model are as follows: The average DSC value obtained from the test is 0.816, the average Jaccard similarity is 0.692, the average precision is 0.825, the average specificity is 0.998, and the average sensitivity is 0.814. The area under the curve (AUC) in the PR curve (**Figure 4**, purple line) is 0.8945. We further compare the segmentation map output from the model with the ground truth obtained from manual segmentation in **Figure 5** (row 2). Our model achieves excellent segmentation results consistently across most nasopharyngeal cancer cases, where the segmentation target areas are large (**Figures 5C, F, G**) and small (**Figures 5A, B, E**). Our model also handles shape variations in nasopharyngeal carcinoma well. The performance is consistent across cases where nasopharyngeal carcinoma is more regularly shaped (**Figures 5D, F**) or less regularly shaped (**Figures 5B, E**). Overall, the segmentation results of our model overlap well with the manual segmentation ground truth.

We try to adjust the hyperparameter t of our model, where t refers to the number of loops in each recurrent convolution block. We set t to 1, 2, 3, and 4, respectively, and the comparison results are shown in **Table 1**. We also use PR curves to visualize the comparison, as shown in **Figure 4A**. From the above comparison, our model achieves the best performance at $t = 3$.

4.2 Ablation Analysis

Ablation analysis aims to analyze the effectiveness of various components of AttR2U-Net in improving the segmentation accuracy of nasopharyngeal carcinoma. We compare the performance of eight variants of AttR2U-Net on the nasopharyngeal carcinoma segmentation task, including a variant only with spatial attention, a variant only with recurrent convolution, a variant only with residual connection, a variant with spatial attention and recurrent convolution, a variant with spatial attention and residual connection, a variant with recurrent convolution and residual connection, the skeleton, and the model without the normalization method.

The results shown in **Table 2** indicate that our intact model outperforms the eight comparison variants in several vital metrics, especially DSC value. Given the small differences in mean values for some metrics in **Table 2**, we perform a Kruskal–Wallis test, where the p-values obtained from testing each variant against our full model in terms of DSC values are 0.009 (Att +Rec), 2×10^{-4} (Att+Res), 7×10^{-4} (Rec+Res), 7×10^{-6} (Only Att), 1×10^{-14} (Only Rec), 3×10^{-7} (Only Res), 3×10^{-7} (Skeleton), 8×10^{-21} (Without Norm), indicating that each variant is significantly different from our full model on DSC values. We also plot PR curves for each ablation experiment as shown in (**Figure 4B**). Therefore, combining spatial attention,

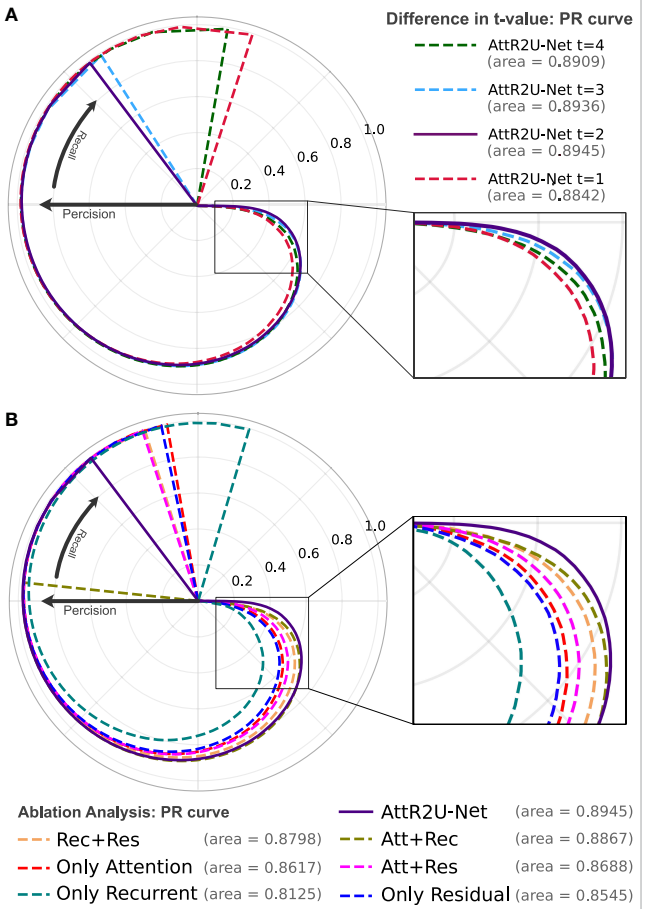


FIGURE 4 | **(A)** The PR curve of AttR2U-Net with t values of 1, 2, 3, and 4. **(B)** The PR curve of the ablation analysis experiment includes 7 models with different parts of AttR2U-Net. The PR curves use a form of the rose diagram to show the comparison results more visually. The area value refers to the area under the PR curve in a plane right-angle coordinate system, also known as AUC. From **(A)**, our model achieves the best performance when t is taken as 2. From **(B)**, our model has the largest AUC value and is at the outermost part of the rose diagram.

residual connection, recurrent convolution, and normalization method can improve the segmentation performance on the nasopharyngeal carcinoma to varying degrees. Specifically, spatial attention improves the accuracy of macroscopic localization, whereas residual connection and recurrent convolution improve segmentation boundary detail. Normalization solves the problem of contrast differences between MRI images.

4.3 Comparison With State-of-the-Art Models

We compare our model with six state-of-the-art models in a comprehensive manner. The comparison models include SEUNet, DANet, FCDenseNet, NestedUNet, DeepLabV3, and FCN16. Examples of NPC segmentation results for each model and their corresponding DSC values are shown in **Figure 5**. For cancerous regions with complex shapes, our model performs better in detail retention and has a higher degree of agreement with manual segmentation. For cases with abnormal cancerous location, our model can find the exact region more accurately than comparison.

We use the same five evaluation metrics as before to quantify the comparison results. Our model outperforms the above six comparison models in the critical metrics for nasopharyngeal carcinoma segmentation. Here we list the test results of the three models with the highest DSC values as follows: For NestedUnet, the average DSC value is 0.790±0.047, 3.19% lower than ours, the average Jaccard similarity is 0.657±0.063, 5.06% lower than ours, the average precision is 0.741±0.058, 10.18% lower than ours, the average specificity is 0.997±0.001, 0.10% lower than ours, and the average sensitivity is 0.856±0.064, 4.91% higher than ours. For DeepLabV3, the average DSC value is 0.787±0.040, 3.55% lower than ours, the average Jaccard similarity is 0.651±0.054, 5.92% lower than ours, the average precision is 0.706±0.057, 14.42% lower than ours, the average specificity is 0.996±0.001, 0.20% lower than ours, and the average sensitivity is 0.894±0.041, 8.95% higher than ours. For SEUNet, the average DSC value is 0.787±0.039, 3.55% lower than ours, the average Jaccard similarity is 0.651±0.053, 5.92% lower than ours, the average precision is 0.702±0.055, 14.91% lower than ours, the average specificity is 0.996±0.001, 0.20% lower than ours, and the average sensitivity is 0.903±0.043, 9.86% higher than ours.

Given the small differences in the mean values of certain metrics in the model comparisons, we perform the Kruskal–Wallis test, in which the p-values obtained from testing each model against our model on DSC values are 0.001 (NestedUnet), 1.4×10^{-5} (DeepLabV3), 1.4×10^{-5} (SEUnet), etc., indicating that each model is significantly different from our model on DSC values. The box plots of DSC values obtained from the tests of each comparison model and the Kruskal–Walis test results are shown in **Figure 6**.

To show the performance differentiation among the models more intuitively, we plot testing results as radar plots shown in **Figure 7**. Additionally, the radar plot transforms the values of each metric into percentage form to make 0.5 value 0% and the maximum value 100%. The conversion formula is as:

$$Y_i = \frac{X_i - 0.5}{Max_i - 0.5} \times 100\,\% \tag{11}$$

where Max_i denotes the maximum value of the ith evaluation metric, X_i denotes the test result of the model on that evaluation metric. For the radar plots, the more peripheral the hexagon enclosed by the model is, the better the model's performance is. The hexagon enclosed by our model is at the outermost part. Moreover, our model ranks first in all evaluation metrics except for a slightly lower SE value. In addition, we plot the specific values of 3 key metrics (DSC value, Jaccard similarity, and precision) for each model, as shown in **Figure 8**. Our model achieves the maximum value on these three most widely used evaluation metrics.

PR curves shown in **Figure 9** compare our model with the state-of-the-art model. Among the seven models, our model has

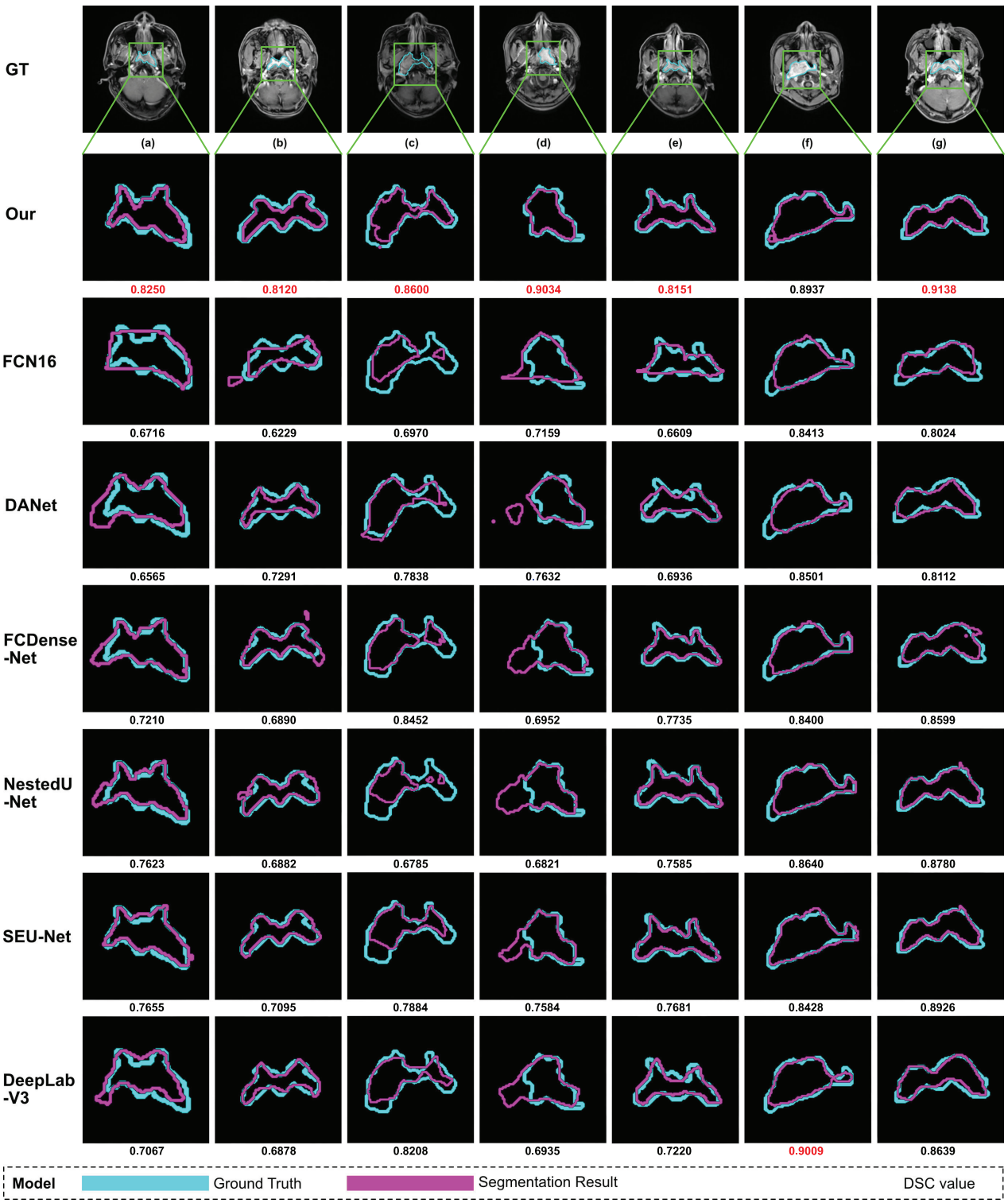


FIGURE 5 | Examples of NPC segmentation results: We select seven typical MRI images of nasopharyngeal carcinoma and present the segmentation results of our model and seven models used for comparison. From top to bottom, the manual segmentation results and model segmentation results of our model, SEUNet, DANet, FCDenseNet, NestedUNet, DeepLabV3, and FCN16, respectively. For each MRI image of nasopharyngeal carcinoma, we draw the manual segmentation boundary with skyblue contours and the segmentation boundary of the corresponding model with purple contours and then calculate the DSC value, which is displayed below the image. We mark the maximum DSC value for each image in red.

TABLE 1 | Experiments indicate that our model achieves the best performance when t = 3.

t	DSC	JS	PC	SP	SE
1	0.792 ± 0.045	0.659 ± 0.061	0.722 ± 0.067	0.996 ± 0.001	0.885 ± 0.046
2	0.815 ± 0.039	0.690 ± 0.055	0.750 ± 0.057	0.997 ± 0.001	0.898 ± 0.045
3	0.816 ± 0.041	0.692 ± 0.058	0.825 ± 0.058	0.998 ± 0.001	0.814 ± 0.057
4	0.803 ± 0.046	0.675 ± 0.063	0.743 ± 0.072	0.997 ± 0.001	0.885 ± 0.046

TABLE 2 | The ablation analysis validates the effectiveness of our model's configuration.

Model	DSC	JS	PC	SP	SE
Att+Rec+Res	0.816 ± 0.041	0.692 ± 0.058	0.825 ± 0.058	0.998 ± 0.001	0.814 ± 0.057
Att+Rec	0.799 ± 0.038	0.668 ± 0.052	0.725 ± 0.050	0.996 ± 0.001	0.896 ± 0.047
Att+Res	0.785 ± 0.051	0.650 ± 0.067	0.780 ± 0.060	0.998 ± 0.001	0.803 ± 0.084
Rec+Res	0.792 ± 0.033	0.658 ± 0.045	0.723 ± 0.047	0.996 ± 0.001	0.879 ± 0.041
Only Att	0.783 ± 0.044	0.646 ± 0.058	0.736 ± 0.065	0.997 ± 0.001	0.846 ± 0.056
Only Rec	0.739 ± 0.058	0.592 ± 0.002	0.679 ± 0.079	0.996 ± 0.001	0.827 ± 0.073
Only Res	0.775 ± 0.048	0.636 ± 0.064	0.753 ± 0.059	0.997 ± 0.001	0.808 ± 0.077
Skeleton	0.769 ± 0.063	0.632 ± 0.080	0.713 ± 0.089	0.996 ± 0.002	0.858 ± 0.076
Without Norm	0.716 ± 0.060	0.562 ± 0.072	0.686 ± 0.071	0.996 ± 0.001	0.764 ± 0.097

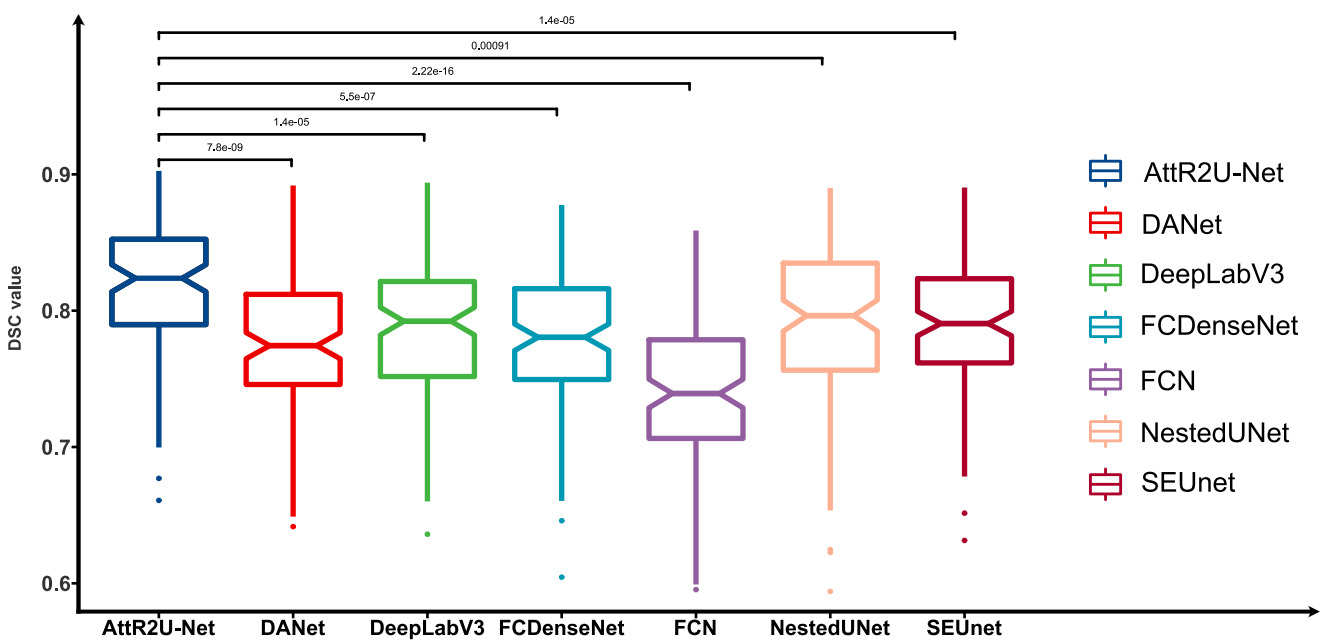


FIGURE 6 | Box plots of the DSC values obtained from the tests on the test set and its Kruskal-Walis test results. The results show that our model achieves the highest DSC values in the test of nasopharyngeal carcinoma segmentation and is significantly different from other models.

the largest area under the curve (AUC). The four models with the best AUC are AttR2U-Net (AUC = 0.8945), SEUnet (AUC = 0.8827, 1.32% lower than ours), DeepLabV3 (AUC = 0.8800, 1.62% lower than ours), and NestedUnet (AUC = 0.8779, 1.86% lower than ours).

To further demonstrate our model's generalization performance, we also perform the patient-wise 5-fold cross validation experiments. **Table 3** presents the model comparison results (DSC value) for the patient-wise 5-fold cross validation. Our model achieves a mean DSC value of 0.7914±0.021 (0.7078~0.7914) in cross-validation. The number of nasopharyngeal carcinoma slices included in each patient case varies widely, ranging from a few to dozens, resulting in different sizes of the training set and test set between different fold numbers. Thus, the results of each experiment in the patient-wise 5-fold cross validation are somewhat different.

The above results suggest that our model is the most reliable for the fully automated end-to-end nasopharyngeal cancer segmentation task. For nasopharyngeal carcinoma, which is relatively complex and difficult to segment accurately, the

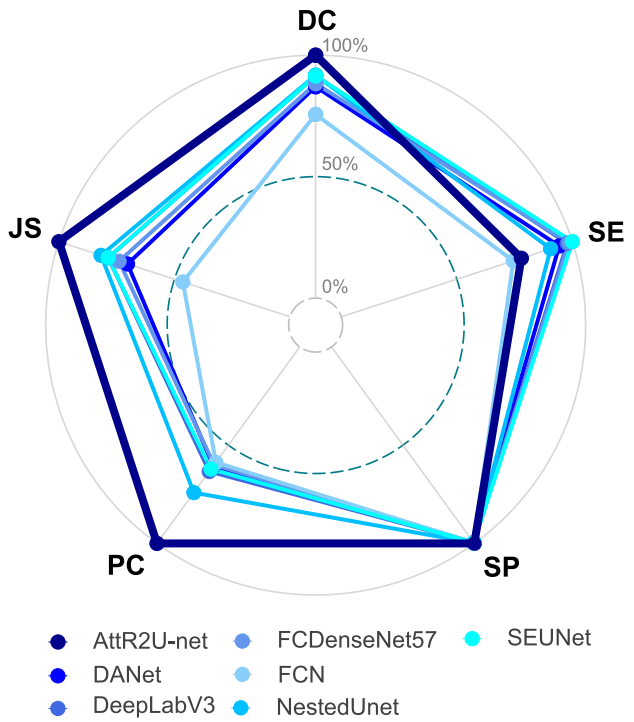


FIGURE 7 | Radar plot showing the comparison of evaluation metrics obtained from each model in the comparison experiment, which includes our model, DANet, DeepLabV3, FCDenseNet, FCN, NestedUNet, and SEUNet. The five vertices of the radar plot respectively represent the average DSC value, average sensitivity, average specificity, average precision, and average Jaccard similarity coefficient. Note that this radar plot uses percentages to express the above evaluation metrics to make the comparison more distinguishable, as shown in equation 11.

AttR2U-Net structure could achieve better performance in terms of accurate localization and detail preservation of the cancerous region.

5 DISCUSSION

This study designed a novel model for automatic nasopharyngeal carcinoma segmentation called AttR2U-Net featuring spatial attention, residual connection, recurrent convolution, and the normalization method. The combination of these features dramatically improves the segmentation performance, allowing our model to take a closer step to the full automation of the nasopharyngeal carcinoma area segmentation. Accurate segmentation of the nasopharyngeal carcinoma area is critical to patients' radiotherapy. Our end-to-end AttR2U-Net model can efficiently segment the nasopharyngeal carcinoma region in a certain degree of accuracy, significantly saving specialized physicians valuable time and circumventing the boundary variances caused by different physicians. We demonstrate that our model is more robust to irregular shapes of nasopharyngeal carcinoma than state-of-the-art models in the literature.

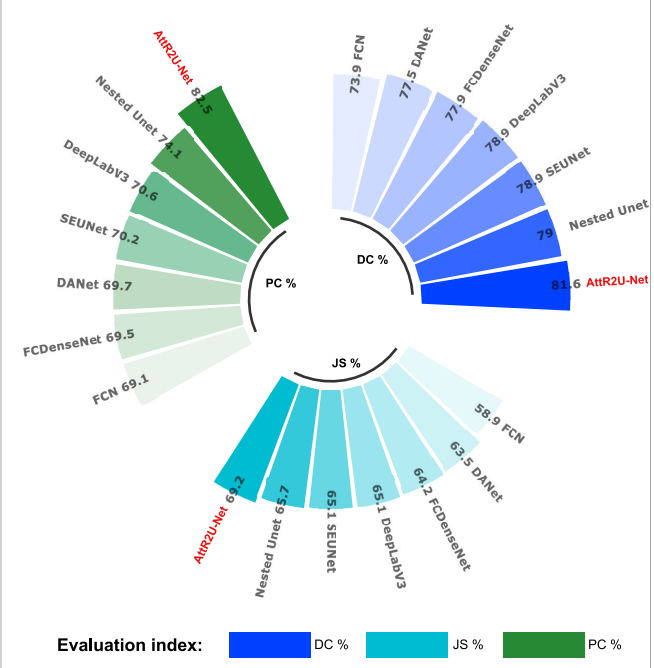


FIGURE 8 | Comparison plots of specific values of 3 key evaluation metrics: include average DSC values, average Jaccard similarity coefficient, and average precision obtained from the test on the nasopharyngeal cancer segmentation test set.

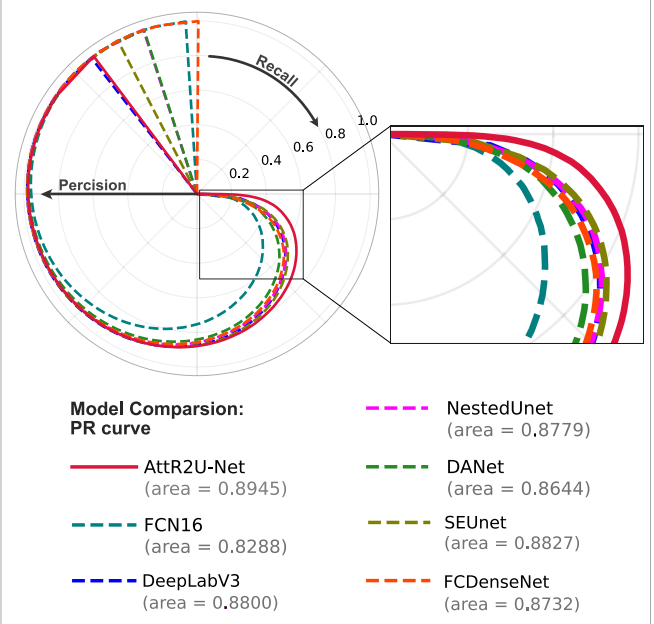


FIGURE 9 | The PR curve of model comparison on the nasopharyngeal carcinoma segmentation test set. The area value indicates the AUC value of each PR curve, and it can be seen that the AttR2U-Net model has the highest AUC value and is located at the outermost part.

TABLE 3 | Results (DSC value) of model comparison experiments using patient-wise 5-fold cross validation.

Model	Fold number					Mean
	1	2	3	4	5	
AttR2U-Net	0.8020	0.8044	0.8030	0.7552	0.7923	0.7914 ± 0.021
Nested UNet	0.8080	0.7870	0.8023	0.7295	0.7627	0.7779 ± 0.032
SEUnet	0.8118	0.7893	0.7926	0.7353	0.7588	0.7776 ± 0.030
FCDenseNet	0.7871	0.7561	0.7932	0.7327	0.7126	0.7563 ± 0.035
DeepLabV3	0.7722	0.7435	0.7708	0.7059	0.7499	0.7485 ± 0.027
DANet	0.7651	0.7269	0.7474	0.6933	0.7146	0.7295 ± 0.028
FCN	0.7461	0.6885	0.7247	0.6727	0.7072	0.7078 ± 0.029

In previous studies of lesion segmentation, deep learning and artificial intelligence techniques are widely used in relatively simple tasks such as segmentation of organs like lung (54), liver (55), and ventricular (56), pancreas (57). Unlike the segmentation in the organs above, the segmentation of nasopharyngeal carcinoma has a higher complexity. It is because its location is in the brain, and the tissue structure in the human brain is the most complex in the human body. The differentiation between nasopharyngeal carcinoma and the surrounding tissues is extremely challenging as its cancerous area is usually mixed with surrounding tissues. Even a professional doctor takes a long time to build one patient's manual segmentation. In addition, compared with other computer vision tasks (58, 59), extremely high precision and detail preservation hold higher stakes for medical image segmentation, as inaccuracies in NPC segmentation may lead to damages of the patient's brain.

Our model achieves a DSC value, a commonly applied measure of the similarity of segmentation results to ground truth, of 0.816 on the nasopharyngeal carcinoma segmentation task. In this study, we compare AttR2U-Net with six other advanced image segmentation models and do ablation analysis experiments. To the best of our knowledge, our model achieves the highest level of performance among fully automated nasopharyngeal carcinoma segmentation models. Specifically, our model design performs more precise localization and higher detail preservation of nasopharyngeal cancer segmentation. As such, our model has significant practical implications for the NPC treatment.

However, our model still has a small number of outlier cases with poor segmentation, mainly those nasopharyngeal carcinomas with a large extent, highly unconventional shape, and extremely remote positions. Our model may be used to assist physicians in segmenting the cancerous region. Nevertheless, given the importance of this segmentation for radiotherapy, it still cannot perform the segmentation task thoroughly in a fully standalone manner. The physician's post-check and calibration remain necessary.

6 CONCLUSION

This study proposed a fully automated end-to-end segmentation model called AttR2U-Net for nasopharyngeal cancer segmentation. Our model creatively combines several advanced computer vision methods, including spatial attention, residual connection, recurrent convolution, and normalization. Compared with other state-of-the-art image segmentation models, our model has the highest segmentation performance. Additionally, our model has excellent efficiency relative to manual segmentation. Thus, it is a promising model to assist physicians in the radiotherapy of nasopharyngeal carcinoma.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the data are not publicly available due to data confidentiality. Requests to access the datasets should be directed to liuxj86@mail.sysu.edu.cn.

ETHICS STATEMENT

Ethical review and approval were not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

JZ: Conceptualization, Methodology, Software, Validation, Writing - Original Draft. LG: Validation, Writing – Review & Editing, Supervision. GH: Conceptualization, Validation, Writing – Review & Editing, Supervision. XL: Conceptualization, Validation, Writing – Review & Editing, Supervision.

FUNDING

This work was supported by the National Natural Science Foundation of China (61901533) and the Shenzhen Fundamental Research Program, China (JCYJ20190807154601663).

REFERENCES

1. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer Statistics in China, 2015. *CA: Cancer J Clin* (2016) 66:115–32. doi: 10.3322/caac.21338
2. Mohammed MA, Abd Ghani MK, Hamed RI, Ibrahim DA. Review on Nasopharyngeal Carcinoma: Concepts, Methods of Analysis, Segmentation, Classification, Prediction and Impact: A Review of the Research Literature. *J Comput Sci* (2017) 21:283–98. doi: 10.1016/j.jocs.2017.03.021
3. Harari PM, Song S, Tomé WA. Emphasizing Conformal Avoidance Versus Target Definition for Imrt Planning in Head-and-Neck Cancer. *Int J Radiat Oncol Biol Phys* (2010) 77:950–8. doi: 10.1016/j.ijrobp.2009.09.062
4. Das IJ, Moskvin V, Johnstone PA. Analysis of Treatment Planning Time Among Systems and Planners for Intensity-Modulated Radiation Therapy. *J Am Coll Radiol* (2009) 6:514–7. doi: 10.1016/j.jacr.2008.12.013
5. Yamazaki H, Shiomi H, Tsubokura T, Kodani N, Nishimura T, Aibe N, et al. Quantitative Assessment of Inter-Observer Variability in Target Volume Delineation on Stereotactic Radiotherapy Treatment for Pituitary Adenoma and Meningioma Near Optic Tract. *Radiat Oncol* (2011) 6:1–6. doi: 10.1186/1748-717X-6-10
6. Feng M, Demiroz C, Vineberg K, Balter J, Eisbruch A. Intra-Observer Variability of Organs at Risk for Head and Neck Cancer: Geometric and Dosimetric Consequences. *Int J Radiat Oncol Biol Phys* (2010) 78:S444–5. doi: 10.1016/j.ijrobp.2010.07.1044
7. Huang KW, Zhao ZY, Gong Q, Zha J, Chen L, Yang R. (2015). Nasopharyngeal Carcinoma Segmentation *via* Hmrf-Em With Maximum Entropy, in: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy: IEEE. pp. 2968–72.
8. Chong V, Fan YF. Detection of Recurrent Nasopharyngeal Carcinoma: Mr Imaging Versus Ct. *Radiology* (1997) 202:463–70. doi: 10.1148/radiology.202.2.9015075
9. Wang C, Dong S, Zhao X, Papanastasiou G, Zhang H, Yang G. Saliencygan: Deep Learning Semisupervised Salient Object Detection in the Fog of Iot. *IEEE Trans Ind Inf* (2019) 16:2667–76. doi: 10.1109/TII.2019.2945362
10. Gao Z, Zhang H, Dong S, Sun S, Wang X, Yang G, et al. Salient Object Detection in the Distributed Cloud-Edge Intelligent Network. *IEEE Netw* (2020) 24:216–24. doi: 10.1109/MNET.001.1900260
11. Dai J, Li Y, He K, Sun J. R-Fcn: Object Detection *via* Region-Based Fully Convolutional Networks. *Adv Neural Inf Process Syst* (2016) 29:379–87.
12. Ghiasi G, Lin TY, Le QV. Nas-Fpn: Learning Scalable Feature Pyramid Architecture for Object Detection. *Proc IEEE/CVF Conf Comput Vision Pattern Recognit* (2019), 7036–45. doi: 10.1109/CVPR.2019.00720
13. Dong S, Gao Z, Sun S, Wang X, Li M, Zhang H, et al. Holistic and Deep Feature Pyramids for Saliency Detection. *BMVC* (2018) 67.
14. Li M, Dong S, Zhang K, Gao Z, Wu X, Zhang H, et al. Deep Learning Intra-Image and Inter-Images Features for Co-Saliency Detection. *BMVC* (2018) 291.
15. Zhang J, Xu C, Gao Z, Rodrigues JJ, de Albuquerque VHC. Industrial Pervasive Edge Computing-Based Intelligence Iot for Surveillance Saliency Detection. *IEEE Trans Ind Inf* (2020) 17:5012–20. doi: 10.1109/TII.2020.3007792
16. Xu C, Gao Z, Zhang H, Li S, de Albuquerque VHC. Video Salient Object Detection Using Dual-Stream Spatiotemporal Attention. *Appl Soft Comput* (2021) 108:107433. doi: 10.1016/j.asoc.2021.107433
17. Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-Decoder With Atrous Separable Convolution for Semantic Image Segmentation. *Proc Eur Conf Comput Vision (ECCV)* (2018) 11211:801–18. doi: 10.1007/978-3-030-01234-2_49
18. Lei T, Wang R, Wan Y, Zhang B, Meng H, Nandi AK. Medical Image Segmentation Using Deep Learning: A Survey. *preprint arXiv* (2020) 2009.13120.
19. Dey R, Hong Y. (2020). Hybrid Cascaded Neural Network for Liver Lesion Segmentation, in: *17th International Symposium on Biomedical Imaging (ISBI)*. Iowa City, USA: IEEE. pp. 1173–7.
20. Singh VK, Rashwan HA, Romani S, Akram F, Pandey N, Sarker MMK, et al. Breast Tumor Segmentation and Shape Classification in Mammograms Using Generative Adversarial and Convolutional Neural Network. *Expert Syst Appl* (2020) 139:112855. doi: 10.1016/j.eswa.2019.112855
21. Conze PH, Kavur AE, Cornec-Le Gall E, Gezer NS, Le Meur Y, Selver MA, et al. Abdominal Multi-Organ Segmentation With Cascaded Convolutional and Adversarial Deep Networks. *Artif Intell Med* (2021) 117:102109. doi: 10.1016/j.artmed.2021.102109
22. Conze PH, Brochard S, Burdin V, Sheehan FT, Pons C. Healthy Versus Pathological Learning Transferability in Shoulder Muscle Mri Segmentation Using Deep Convolutional Encoder-Decoders. *Comput Med Imaging Graphics* (2020) 83:101733. doi: 10.1016/j.compmedimag.2020.101733
23. Ronneberger O, Fischer P, Brox T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation, in: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Munich, Germany: Springer. pp. 234–41.
24. Li S, Xiao J, He L, Peng X, Yuan X. The Tumor Target Segmentation of Nasopharyngeal Cancer in Ct Images Based on Deep Learning Methods. *Technol Cancer Res Treat* (2019) 18:1533033819884561. doi: 10.1177/1533033819884561
25. Jaderberg M, Simonyan K, Zisserman A, Kavukcuoglu K. Spatial Transformer Networks. *Adv Neural Inf Process Syst* (2015) 28:2017–25.
26. Liang M, Hu X. (2015). Recurrent Convolutional Neural Network for Object Recognition, in: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, USA: IEEE. pp. 3367–75.
27. He K, Zhang X, Ren S, Sun J. (2016). Deep Residual Learning for Image Recognition, in: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, USA: IEEE. pp. 770–8.
28. Mnih V, Heess N, Graves A, Kavukcuoglu K. Recurrent Models of Visual Attention. *Adv Neural Inf Process Syst* (2014) 27:2204–12.
29. Bahdanau D, Cho K, Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate. *preprint arXiv* (2014) 1409.0473.
30. Luong MT, Pham H, Manning CD. Effective Approaches to Attention-Based Neural Machine Translation. *preprint arXiv* (2015) 1508.04025. doi: 10.18653/v1/D15-1166
31. Wang F, Jiang M, Qian C, Yang S, Li C, Zhang H, et al. (2017). Residual Attention Network for Image Classification, in: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, USA: IEEE. pp. 3156–64.
32. Wang X, Xiong X, Neumann M, Piergiovanni A, Ryoo MS, Angelova A, et al. (2020). Attentionnas: Spatiotemporal Attention Cell Search for Video Classification, in: *European Conference on Computer Vision (ECCV)*. Glasgow, UK: Springer. pp. 449–65.
33. Pu L, Zou Y, Zhang J, Huang S, Yao L. (2019). Using Dependency Information to Enhance Attention Mechanism for Aspect-Based Sentiment Analysis, in: *Natural Language Processing and Chinese Computing*. Dunhuang, China: Springer. pp. 672–84.
34. Yin W, Schütze H, Xiang B, Zhou B. Abcnn: Attention-Based Convolutional Neural Network for Modeling Sentence Pairs. *Trans Assoc Comput Linguistics* (2016) 4:259–72. doi: 10.1162/tacl_a_00097
35. Fu J, Liu J, Tian H, Li Y, Bao Y, Fang Z, et al. (2019). Dual Attention Network for Scene Segmentation, in: *2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, USA: IEEE. pp. 3146–54.
36. Yuan Y, Huang L, Guo J, Zhang C, Chen X, Wang J. Ocnet: Object Context Network for Scene Parsing. *preprint arXiv* (2018) 1809.00916.
37. Li H, Xiong P, An J, Wang L. Pyramid Attention Network for Semantic Segmentation. *preprint arXiv* (2018) 1805.10180.
38. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, et al. Attention U-Net: Learning Where to Look for the Pancreas. *preprint arXiv* (2018) 1804.03999.
39. Lai S, Xu L, Liu K, Zhao J. (2015). Recurrent Convolutional Neural Networks for Text Classification, in: *Twenty-ninth AAAI conference on artificial intelligence*. Austin, USA: AAAI.
40. Alom MZ, Hasan M, Yakopcic C, Taha TM, Asari VK. Recurrent Residual Convolutional Neural Network Based on U-Net (R2u-Net) for Medical Image Segmentation. *arXiv* (2018) 1802.06955. doi: 10.1109/NAECON.2018.8556686
41. Orhan E, Pitkow X. (2018). Skip Connections Eliminate Singularities, in: *6th International Conference on Learning Representations (ICLR)*. Vancouver, Canada: openreview.net.
42. He K, Zhang X, Ren S, Sun J. (2016). Identity Mappings in Deep Residual Networks, in: *European Conference on Computer Vision (ECCV)*. Amsterdam, The Netherlands: Springer. pp. 630–45.
43. Balduzzi D, Frean M, Leary L, Lewis J, Ma KWD, McWilliams B. (2017). The Shattered Gradients Problem: If Resnets are the Answer, Then What is the

Question?, in: *34th International Conference on Machine Learning*. Sydney, Australia: PMLR. pp. 342–50.
44. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. *preprint arXiv* (2014) 1412.6980.
45. Girosi F, Jones M, Poggio T. Regularization Theory and Neural Networks Architectures. *Neural Comput* (1995) 7:219–69. doi: 10.1162/neco.1995.7.2.219
46. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks From Overfitting. *J Mach Learn Res* (2014) 15:1929–58.
47. Dice LR. Measures of the Amount of Ecologic Association Between Species. *Ecology* (1945) 26:297–302. doi: 10.2307/1932409
48. Bertels J, Eelbode T, Berman M, Vandermeulen D, Maes F, Bisschops R, et al. (2019). Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory and Practice, in: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Shenzhen, China: Springer. pp. 92–100.
49. Xuehua W, Xiangcong X, Yaguang Z, Dingan H. (2021). A New Method With Seu-Net Model for Automatic Segmentation of Retinal Layers in Optical Coherence Tomography Images, in: *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*. Nanchang, China: IEEE. pp. 260–3.
50. Guan S, Khan AA, Sikdar S, Chitnis PV. Fully Dense Unet for 2-D Sparse Photoacoustic Tomography Artifact Removal. *IEEE J Biomed Health Inf* (2019) 24:568–76. doi: 10.1109/JBHI.2019.2912935
51. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. Unet++: A Nested U-Net Architecture for Medical Image Segmentation. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Granada, Spain: Springer (2018). p. 3–11.
52. Chen LC, Papandreou G, Schroff F, Adam H. Rethinking Atrous Convolution for Semantic Image Segmentation. *preprint arXiv* (2017) 1706.05587.
53. Long J, Shelhamer E, Darrell T. (2015). Fully Convolutional Networks for Semantic Segmentation, in: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, USA: IEEE. pp. 3431–40.
54. Skourt BA, El Hassani A, Majda A. Lung Ct Image Segmentation Using Deep Neural Networks. *Proc Comput Sci* (2018) 127:109–13. doi: 10.1016/j.procs.2018.01.104
55. Gotra A, Sivakumaran L, Chartrand G, Vu KN, Vandenbroucke-Menu F, Kauffmann C, et al. Liver Segmentation: Indications, Techniques and Future Directions. *Insights Into Imaging* (2017) 8:377–92. doi: 10.1007/s13244-017-0558-1
56. Pednekar A, Kurkure U, Muthupillai R, Flamm S, Kakadiaris IA. Automated Left Ventricular Segmentation in Cardiac Mri. *IEEE Trans Biomed Eng* (2006) 53:1425–8. doi: 10.1109/TBME.2006.873684
57. Roth HR, Lu L, Farag A, Shin HC, Liu J, Turkbey EB, et al. (2015). Deeporgan: Multi-Level Deep Convolutional Networks for Automated Pancreas Segmentation, in: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Munich, Germany: Springer. pp. 556–64.
58. Gao Z, Xu C, Zhang H, Li S, de Albuquerque VHC. Trustful Internet of Surveillance Things Based on Deeply Represented Visual Co-Saliency Detection. *IEEE Internet Things J* (2020) 7:4092–100. doi: 10.1109/JIOT.2019.2963701
59. Dong S, Gao Z, Pirbhulal S, Bian GB, Zhang H, Wu W, et al. Iot-Based 3d Convolution for Video Salient Object Detection. *Neural Comput Appl* (2020) 32:735–46. doi: 10.1007/s00521-018-03971-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhang, Gu, Han and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 02 February 2022
doi: 10.3389/fonc.2021.811355

WEENet: An Intelligent System for Diagnosing COVID-19 and Lung Cancer in IoMT Environments

Khan Muhammad[1], Hayat Ullah[2], Zulfiqar Ahmad Khan[2], Abdul Khader Jilani Saudagar[3]*, Abdullah AlTameem[3], Mohammed AlKhathami[3], Muhammad Badruddin Khan[3], Mozaherul Hoque Abul Hasanat[3], Khalid Mahmood Malik[4], Mohammad Hijji[5] and Muhammad Sajjad[6,7]*

[1] Visual Analytics for Knowledge Laboratory (VIS2KNOW Lab), School of Convergence, College of Computing and Informatics, Sungkyunkwan University, Seoul, South Korea, [2] Department of Software, Sejong University, Seoul, South Korea, [3] Information Systems Department, College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia, [4] Department of Computer Science & Engineering, Oakland University, Rochester, MI, United States, [5] Faculty of Computers & Information Technology, Computer Science Department, University of Tabuk, Tabuk, Saudi Arabia, [6] Digital Image Processing Laboratory, Islamia College Peshawar, Peshawar, Pakistan, [7] Color and Visual Computing Lab, Department of Computer Science, Norwegian University of Science and Technology (NTNU), Gjøvik, Norway

The coronavirus disease 2019 (COVID-19) pandemic has caused a major outbreak around the world with severe impact on health, human lives, and economy globally. One of the crucial steps in fighting COVID-19 is the ability to detect infected patients at early stages and put them under special care. Detecting COVID-19 from radiography images using computational medical imaging method is one of the fastest ways to diagnose the patients. However, early detection with significant results is a major challenge, given the limited available medical imaging data and conflicting performance metrics. Therefore, this work aims to develop a novel deep learning-based computationally efficient medical imaging framework for effective modeling and early diagnosis of COVID-19 from chest x-ray and computed tomography images. The proposed work presents "WEENet" by exploiting efficient convolutional neural network to extract high-level features, followed by classification mechanisms for COVID-19 diagnosis in medical image data. The performance of our method is evaluated on three benchmark medical chest x-ray and computed tomography image datasets using eight evaluation metrics including a novel strategy of cross-corpse evaluation as well as robustness evaluation, and the results are surpassing state-of-the-art methods. The outcome of this work can assist the epidemiologists and healthcare authorities in analyzing the infected medical chest x-ray and computed tomography images, management of the COVID-19 pandemic, bridging the early diagnosis, and treatment gap for Internet of Medical Things environments.

Keywords: medical imaging, COVID-19 diagnosis, machine learning, Internet of Medical Things, deep learning, x-ray imaging, cancer categorization

OPEN ACCESS

Edited by:
Guang Yang,
Imperial College London,
United Kingdom

Reviewed by:
Kun Wang,
UCLA School of Law, United States
Kemal Polat,
Abant Izzet Baysal University, Turkey
Zhenghua Chen,
Nanyang Technological University,
Singapore

****Correspondence:***
Khan Muhammad
khan.muhammad@ieee.org
Abdul Khader Jilani Saudagar
aksaudagar@imamu.edu.sa

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *08 November 2021*
Accepted: *01 December 2021*
Published: *02 February 2022*

Citation:
Muhammad K, Ullah H, Khan ZA, Saudagar AKJ, AlTameem A, AlKhathami M, Khan MB, Abul Hasanat MH, Mahmood Malik K, Hijji M and Sajjad M (2022) WEENet: An Intelligent System for Diagnosing COVID-19 and Lung Cancer in IoMT Environments. Front. Oncol. 11:811355. doi: 10.3389/fonc.2021.811355

1 INTRODUCTION

In the beginning of December 2019, a novel infectious acute disease called coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has emerged and caused severe impact on health, human lives, and global economy. This COVID-19 disease originated in Wuhan city of China and then spread in several other countries and become a global pandemic (1). This virus is easily transmitted between two persons through petite drops caused by coughing, sneezing, and talking during close contact. The infected person usually has certain symptoms after 7 days that include high fever, continuous cough, shortness of breath, and taste loss. According to the statistical report of the World Health Organization (WHO) (2), COVID-19 affected around 192 countries with 199 million confirmed active cases and 4.2 million confirmed deaths till August 4, 2021. Considering the fast spread and its high contiguous nature, it is essential to diagnose COVID-19 at an early stage to greatly prevent the outbreak by isolating the infected persons, thereby minimizing the possibilities of infection to healthy people. Till date, the most common and convenient technique for diagnosing COVID-19 is the reverse transcription polymerase chain reaction (RT-PCR). However, this technique has very low precision, high delay, and low sensitivity, making it less effective in preventing the spread of COVID-19 (3).

Besides the RT-PCR testing system, there are several other medical imaging-based COVID-19 diagnosing methods such as computed tomography (CT) (4–6) and chest radiography (x-ray) (7, 8). Diagnosis of COVID-19 is typically associated with both the symptoms of pneumonia and medical chest x-ray tests (9, 10). Chest x-ray is the first medical imaging-based technique that plays an important role in the diagnosis and detection of COVID-19 disease. Some attempts have been made in the literature to detect COVID-19 from medical chest x-ray images using machine learning and deep learning approaches (11, 12). For instance, Narin et al. (13) evaluated the performance of five pretrained Convolutional Neural Network (CNN)-based models for the detection of coronavirus pneumonia-infected patients using medical chest x-ray images. Ismail et al. (14) utilized deep feature extraction and fine-tuning of the pretrained CNNs to classify COVID-19 and normal (healthy) chest x-ray images. Tang et al. (15) used chest x-ray images with effective screening for the detection of COVID-19 cases. Furthermore, Jain et al. (16) used transfer learning for the COVID-19 detection using medical chest x-ray images, and they compared the performance of medical imaging-based COVID-19 detection methods.

More recently, several other deep learning-based approaches (17) are presented to overcome the limitations of previous imaging-based COVID-19 detection methods. For instance, Minaee et al. (18) proposed a transfer learning strategy to improve the COVID-19 recognition rate in medical chest x-ray images. They investigated different pretrained CNN architectures on their newly prepared COVID-19 x-ray image datasets and claimed reasonable results. However, their newly created dataset is not balanced and has a smaller number of COVID-19p images compared with non-COVID-19 images. Aniello et al. (19) presented ADECO-CNN to classify infected and noninfected patients *via* medical CT images. They compared their CNN architecture with pretrained CNNs including VGG19, GoogleNet, and ResNet50. Yujin et al. (20) suggested a patch-based CNN approach for efficient classification and segmentation of COVID-19 chest x-ray images. They first preprocessed medical chest x-ray images and then fed them into their proposed network for infected lung area segmentation and classification in medical images. However, their attained performance is relatively low due to the small number of images in their used dataset. Similarly, Yu-Huan et al. (21) presented a joint classification and segmentation framework called JCS for COVID-19 medical chest CT diagnosis. They trained their JCS system on their newly created COVID-19 classification and segmentation dataset. They claimed real-time and explainable diagnosis of COVID-19 in chest CT images with high efficiency in both classification and segmentation. Afshar et al. (22) proposed a deep uncertainty-aware transfer learning framework for COVID-19 detection in medical x-ray and CT images. They first extracted CNN features from images of chest x-ray and CT scan dataset and then evaluated by different machine learning classifiers to classify the input image as COVID or non-COVID.

The current COVID-19 pandemic situation greatly overwhelms the health monitoring systems of even developed countries, leading to an upward trend in the number of deaths on a daily basis. Also, the inaccessibility of healthcare system and required medication to the rural areas caused increase in the loss of human lives. Therefore, an intelligent AI-driven healthcare system is necessarily needed for combating with COVID-19 pandemic and rescuing hospitals and other medical staff. Thanks to the Internet of Things (23–26) and Internet of Medical Things (IoMT) (27, 28) for offering powerful features (i.e., online monitoring, high-speed communication, and remote checkups) that can greatly assist healthcare system of a country against COVID-19 pandemic (29). Also, Healthcare 5.0 with 5G-enabled IoMT environment can effectively improve the accessibility of doctors and nurses to their patients in remote areas, enabling COVID-19 patients to control their health based on daily recommendations from doctors.

Undoubtedly, the deployment of 5G-enabled IoMT protocols can greatly enhance the performance of smart healthcare system by connecting hospitals and patients, transmitting their health-related data between both parties. However, such a smart IoMT healthcare environment demands computationally efficient yet accurate AI algorithms (including both machine learning and deep learning algorithms) (30). Most of the existing deep learning approaches use computationally complex CNN architectures that require high network bandwidth and computational requirements and cannot be employed on resource constrained devices. Thus, the architecture of AI Algorithm (i.e., CNN architecture) to be deployed must meet the requirements of the executional environment (the device used in IoMT environment).

To alleviate the shortcomings of pervious approaches and design energy-efficient model for IoMT-enabled environment,

we propose a computationally efficient, yet accurate CNN architecture called WEENet. The proposed architecture is designed to efficiently detect COVID-19 in medical chest x-ray images, requiring limited computational resources. More precisely, the key contributions of this study are summarized as follows:

1. Deep learning-based models require huge amount of medical imaging data to train effectively, but COVID-19 benchmarks have relatively limited number of samples especially for COVID class. To increase the number of images for effective training of the proposed WEENet framework, we applied offline data augmentation techniques on available medical chest x-ray images such as rotation, flipping, zooming, etc. that brought improvements in the performance as evident from the results.
2. WEENet is developed to detect COVID-19 in medical chest x-ray images and support the management of IoMT environments. WEENet uses EfficientNet (31) model as a backbone for feature extraction from chest x-ray images, followed by stacked autoencoding layers to represent the features in more abstract form before the final classification decision.
3. The performance of several deep learning-based models is evaluated using benchmark medical chest x-ray images datasets and eight evaluation metrics including a novel strategy of cross-corpse and robustness evaluation for COVID-19 detection in chest x-ray images. Furthermore, we also compared the performance of our WEENet with other state-of-the-art (SOTA) methods, where it surpassed in terms of several evaluation metrics.

The remainder of the article is organized as follows: ***Section 2*** covers the proposed IoMT-based WEENet framework with a discussion on datasets. In ***Section 3***, we discuss the experimental setups, the experimental results, and their analysis. Finally, ***Section 4*** concludes this paper and suggests future research directions.

2 PROPOSED IOMT-BASED WEENET FRAMEWORK

This section discusses the overall workflow of our WEENet framework in IoMT environment for efficient and timely detection of COVID-19 in x-ray images over edge computing platforms. For better understanding, the proposed WEENET framework is divided into three phases including Data Acquisition, Preprocessing, and WEENet. The first phase presents the detail of data collection from different sources, followed by the second phase which performs extensive data augmentation on the data collected in the first phase to prevent underfitting/overfitting problems. The third phase contains the WEENet architecture which is responsible for COVID-19 detection in x-ray images. The overall graphical overview of our proposed framework with all phases is given in **Figure 1** and explained in the following subsections.

2.1 Data Acquisition and Preprocessing

During the pandemic, hospitals around the world produced image data related to COVID-19 (such as medical x-ray and CT images), and some of them are publicly available for research purposes in medical imaging. However, the available COVID-19 image datasets are either not well organized or have lack of balance between positive and negative class samples, which often lead network to model overfitting during the training process. Therefore, the research community is working to organize the available COVID-19 image data and make it usable before utilizing it for early diagnosis of COVID-19. To achieve data diversity and balance between positive and negative class samples, we actively used data augmentation approaches, which not only increase the volume of data but can also significantly improve the classification performance of deep learning models as evident from our experiments.

In this research, we have used three different COVID-19 image datasets namely chest x-ray images (CXI) (18), x-ray dataset COVID-19 (XDC) (32), and COVID-19 radiography database (CRD) (33), where each dataset contains medical chest x-ray images of positive and negative patients. To alleviate the chances of model overfitting and class biasness, we performed extensive data augmentation by equalizing the number of positive and negative class samples in each of the abovementioned datasets. Considering the number of images per class in the dataset, we performed data augmentation with different augmentation ratio for each dataset so that we can obtain balance training data. Following this strategy, we augmented the COVID-19 images of CXI (18) dataset with augmentation ratio of 1:15 such that each image is reproduced in 15 different variants. Similarly, for XDC (32), we used the data augmentation ratio of 1:10 for both positive and negative classes. For CRD (33) dataset, we only augmented the COVID-9 class with the augmentation ratio of 1:3, where each image is reconstructed with its 3 different variants. The proposed data augmentation strategy analyzed different augmentation approaches and then selected the most suitable eight distinct operations on each image of the dataset that include Rotation, Zoom, Width shift, Height shift, Shear, Fill mode, Flip, and Brightness operations before forwarding to our proposed WEENet for training. The details of augmentation operations used in our method are listed in **Table 1**.

It can be noticed that images in original XDC dataset are insufficient for training a deep learning algorithm. Also, the number of positive samples in the CXI dataset is comparatively lesser than negative samples prior to data augmentation process. Similarly, the CRD dataset has also a huge difference between the number of positive and negative samples. On the other hand, the augmented version of the listed datasets can be found well balanced and rich in terms of data diversity, thus more suitable for deep learning-based methods.

2.2 EfficientNet: The Backbone Architecture

Several CNN architectures have been explored before choosing the appropriate model that are extensively used in different

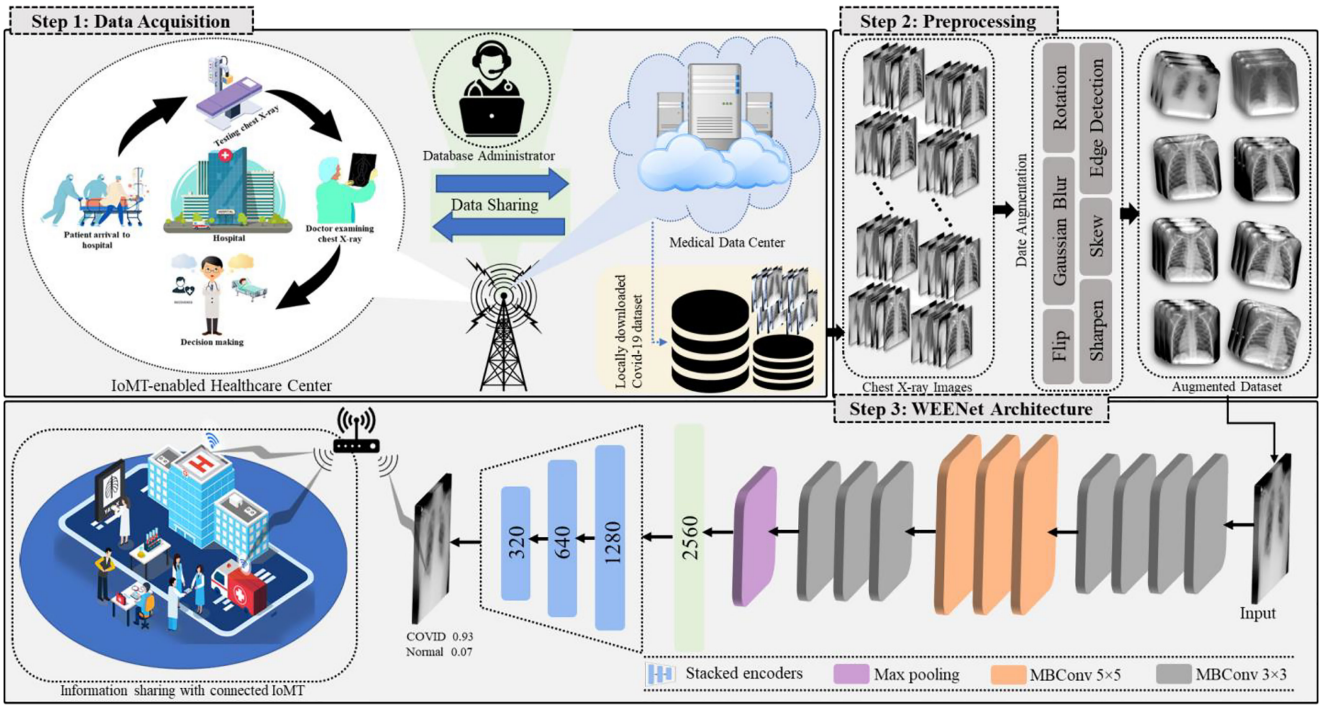


FIGURE 1 | Overview of the proposed WEENet-assisted framework for COVID-19 diagnosis using chest x-ray images with the support of 5G technology and efficient management for IoMT environments.

domain studies such as time series prediction, classification, object detection, and crowed estimation. These architectures include VGG16 (34), VGG19 (34), ResNet18 (35), ResNet50 (35), and ResNet101 (35) that are used by researchers for COVID-19 detection in chest x-ray images, but each CNN model has its own pros and cons. However, researchers investigate these architectures to boost their accuracy by using different scaling strategies to adjust the network depth, width, or resolution. Most of the networks are based on single scaling, that scales only a single dimension from depth, width, and size. Though, scaling two or three dimensions will yield efficiency and suboptimal accuracy. To this end, we investigate EfficientNet that scales all the dimensions through compound scaling technique. This network is developed through leveraging multiobjective architecture search, which optimizes both floating point operations (FLOPs) and accuracy. EfficientNet uses the search space of (36) and ACC (m) × [FLOPS (m)/T]^w as an optimization tool. The ACC (m) and FLOPS (m) represent the accuracy and FLOPs of model m while T and w are the FLOPs target and hyperparameters, respectively. These terms control the tradeoff between the accuracy and FLOPs. This network comprises several convolutional layers where different-sized kernels are equipped in each layer. The input frame having three channels (R, G, B) corresponds to size such as 224 × 224 × 3. The subsequent layers are scaled down in a resolution that reduces the size of feature maps while the width is scaled up to increase the accuracy. This tool ensures the collection of important features from the input frame. For example, the second layer consists of Width = 112 kernels, and the number of kernels by next convolution is Width = 64. The total maximum kernels used are Depth = 2,560 in the last layer, where the resolution is 7 × 7 which represents the most discriminative features. At the end, we added max pooling layer that is followed by encoding layers and a SoftMax layer for the final classification.

TABLE 1 | The operational details of our proposed data augmentation strategy.

No.	Technique	Parameter range
1	Rotation	−25~25
2	Zoom	0.10
3	Width shift	0.01
4	Height shift	0.01
5	Shear	0.1
6	Fill mode	Nearest
7	Flip	Right and left
8	Brightness	0.50, 1.50

2.3 The Proposed WEENet

The proposed WEENet is based on EfficientNet model followed by encoding layers. EfficientNet is used to extract important features from the input data and then the output is feed forward to stacked encoding layers. The stacked encoding layers are based on autoencoder (37) used to compress the data from high dimension into low dimension, while preserving the salient information from the input data. Autoencoders are a type of deep neural networks that

map the data to itself through a process of (nonlinear) dimensionality reduction, followed by dimensionality expansion. The autoencoder models include three layers: input, hidden, and output layer as shown in **Figure 2**. The encoder part is used to map the input data into lower dimension followed by decoding layers to reconstruct it. Let us suppose the input data $(X_n)_{n=1}^{N}$, where X_n belongs to $R^{m\,x\,1}$, h_n is the low-dimension map (hidden state) which is calculated from X_n, and "O_n" is the output decoder. The mathematical representation of encoding and decoding layer is shown in Eq. (1).

$$\boldsymbol{k}_n = F(\boldsymbol{W}_1 X_n +)\boldsymbol{k}_1) \tag{1}$$

here, F represents the encoding function, $\boldsymbol{W}_1$ is the weight metrics, and $\boldsymbol{k}_1$ is the bias term. The mathematical representation of the decoding layer is shown in Eq. (2).

$$O_n = G(\boldsymbol{W}_2 X_n + \boldsymbol{k}_2) \tag{2}$$

In this equation, G is the decoding function, $\boldsymbol{W}_2$ is the weight metrics, and $\boldsymbol{k}_2$ is the bias term of decoding layer. In our WEENet, we used the encoding part of the autoencoder to represent the features in more abstract form. In these layers, the high-dimension EfficientNet features is encoded to low-dimension features. In the proposed model, two encoding layers are incorporated with EfficientNet architecture. The output of EfficientNet is 2,560 dimension feature vector which is encoded to 1,280 dimension feature vector. Furthermore, 1,280 dimension feature vector is then encoded to 640 dimension feature vector is then 320. The proposed model is trained for 50 epochs, using SGD optimizer with 0.0001 learning rate, and its performance is tested against SOTA as given in ***Section 3***.

3 EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we evaluated our WEENet on three publicly available COVID-19 datasets and compared the classification performance with other methods. For this, we first provide the details of experimental settings of this research study, followed by information about datasets and metrics for performance evaluation. Subsequently, we compare the proposed WEENet with other SOTA CNN architectures used for COVID-19 classification. Finally, we close this section by emphasizing on the feasibility of our proposed WEENet framework for COVID-19 diagnosis in 5G-enabled IoMT environments.

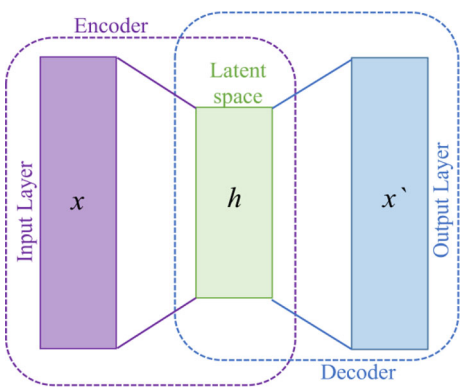


FIGURE 2 | General overview of autoencoder architecture.

3.1 Implementation Details

This section provides the detail of experimental settings and the execution environment used for implementing our proposed WEENet framework. The proposed method is purely implemented in Python (version 3.5) language using Visual Studio Code (VSCode)-integrated development environment. The WEENet concepts are implemented by utilizing a very prominent deep learning framework called Keras on Intel Core i7 CPU equipped with a GPU of Nvidia GTX having 6 GB onboard memory. The proposed WEENet architecture is trained on three different datasets including CXI, XDC, and CRD with the same configuration of hyperparameters, i.e., number of epochs, batch size, learning rate, weight decay, etc. The training and validation performances of our WEENet on CXI, XDC, and CRD datasets are visually depicted in **Figures 3**–**5**, respectively.

3.2 Details of the Datasets

For experimental evaluation, we have used three publicly available datasets (18, 32, 33) to validate the performance of our proposed method compared with other SOTA CNN architectures. These datasets contain chest x-ray images of positive and negative COVID-19 patients assigned with corresponding labels, i.e., COVID-19 and normal. The statistical details of the abovementioned datasets are listed in **Table 2**. Besides these datasets, there are several other publicly available datasets commonly used for COVID-19 classification. However, most of them are either imbalance or have weak diversity leading to poor performance. Therefore, we selected CXI, XDC, and CRD datasets from the publicly available listed datasets in **Table 3**. The detail of each dataset is given in **Table 3** including publishing year, name of the dataset, number of samples in COVID class and non-COVID class, methods of evaluation, and experimental outcomes in terms of sensitivity, specificity, and accuracy.

3.2.1 CXI Dataset

It is one of the most used datasets for COVID-19 diagnosis in medical image analysis community. This dataset contains a total of 184 COVID-19 infected and more than 5,000 normal chest x-ray images. Clearly, the original CXI (18) dataset has imbalance class samples that significantly affect a model's performance during training. Considering the chances of overfitting during training, we augmented the dataset and balanced the number of images for both COVID-19 and normal class. The number of per class images for both original and augmented CXI dataset is listed in **Table 2**.

3.2.2 XDC Dataset

This dataset is created by collecting a small number of chest x-ray images of positive and negative COVID-19 patients. Overall, this

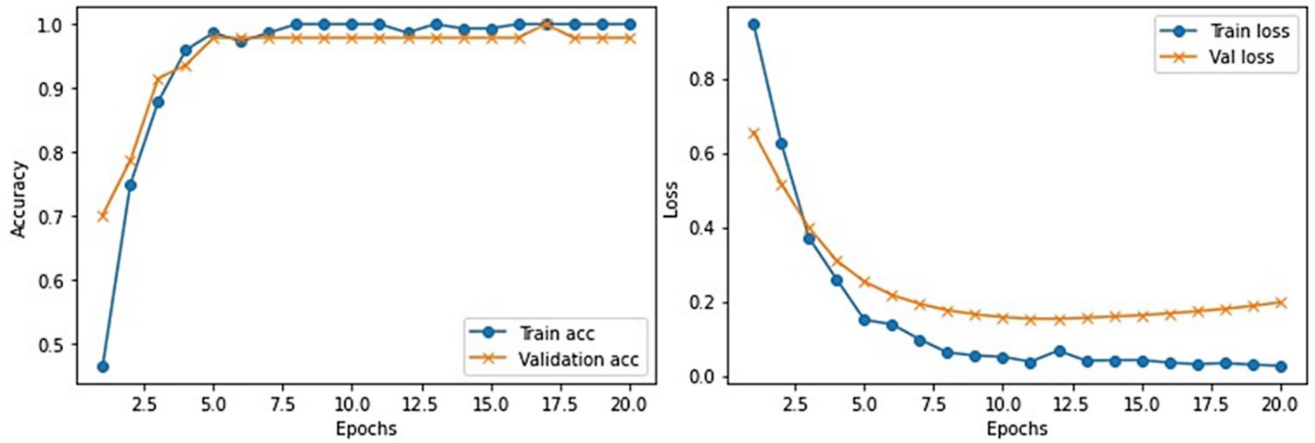


FIGURE 3 | Training and validation performance of our proposed WEENet over medical CXI image dataset.

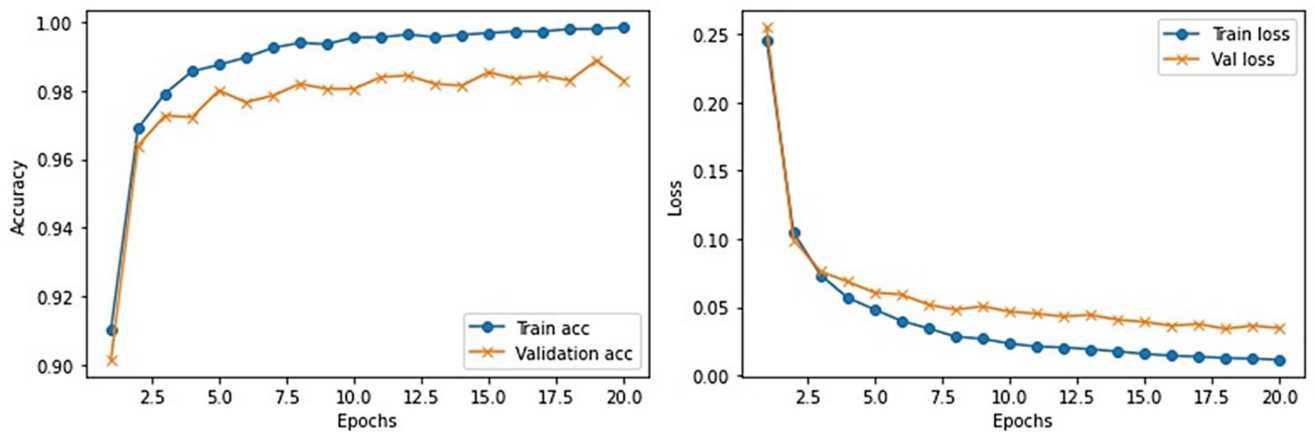


FIGURE 4 | Training and validation performance of our proposed WEENet over medical XDC image dataset.

dataset is very small and cannot be used for training large CNN networks. Also, such a lesser amount of image data often leads to model underfitting where model struggles to learn from the data under observation. To avoid such kind of hurdles during training, we augmented the XDC (32) dataset and increased the number of images from 94 to 940 for both COVID-19 and normal class as given in **Table 2**.

TABLE 2 | Number of samples per class in original and augmented datasets.

Dataset	Original dataset		Augmented dataset	
	COVID-19	Normal	COVID-19	Normal
CXI (18)	200	5,000	3,000	3,000
XDC (32)	94	94	940	940
CRD (33)	3,616	10,192	10,848	10,848

3.2.3 CRD Dataset

The COVID-19 radiography dataset is the large-scale chest x-ray image dataset released in different versions. In the first release, they publicly share 219 COVID-19 infected and 1,341 normal chest x-ray images. In the second release, they increased the number of COVID-19 infected chest x-ray images to 1,200. Following this, in the third release, the number of COVID-19 infected chest x-ray images is increased to 3,616 and normal chest x-ray images to 10,192. In this paper, we used the final release of the CRD (33) dataset, whose statistical details are presented in **Table 2**.

3.3 Evaluation Metrics

In image classification problem, the performance of trained CNN model is mostly evaluated by conducting quantitative assessment *via* commonly used classification performance metrics. These

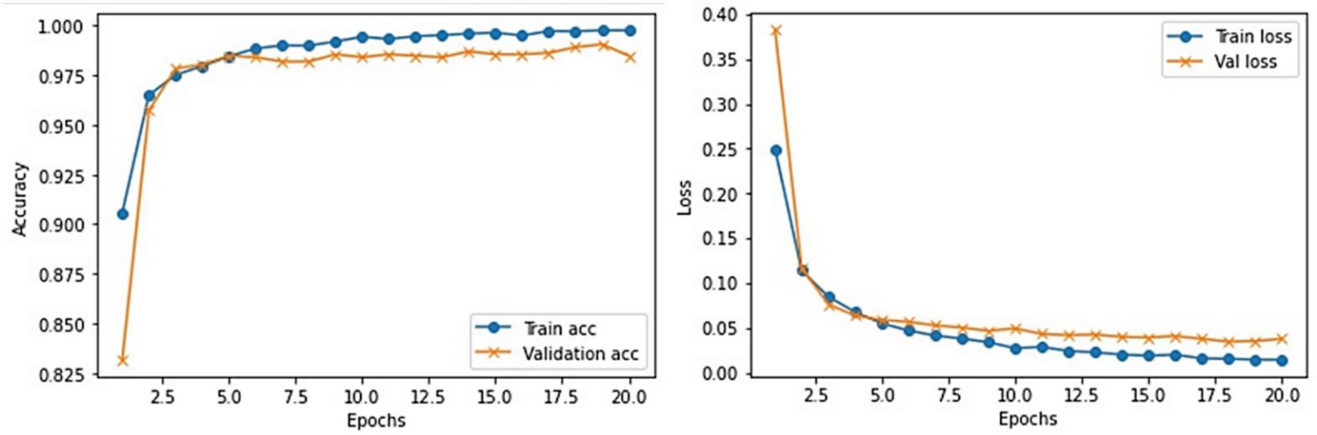


FIGURE 5 | Training and validation performance of our proposed WEENet over medical CRD image dataset.

metrics can be easily computed with the help of confusion matrix by forwarding the actual class labels and predicted labels. Following this strategy, in this paper, we used eight commonly used performance evaluation metrics that include true positive (TP), false positive (FP), false negative (FN), true negative (TN), sensitivity, specificity, accuracy, and receiver operating characteristics (ROC) for validating the classification performance of our WEENet. The values of TP, FP, FN, and TN are retrieved from the confusion matrix and the sensitivity, specificity, accuracy, and ROC metrics are computed accordingly as Eq. (3) to Eq. (6).

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{4}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{5}$$

$$\text{ROC} = \text{Sensitivity} \sim \text{R} \sim \text{Specificity} \tag{6}$$

TABLE 3 | Detailed information of the collected SOTA including techniques and their other important remarks.

Ref.	Year	Dataset	COVID	Non-COVID	Technique	Inclusion	Performance (%)
(18)	2020	CXI	200	5,000	ResNet18, ResNet50, SqueezeNet, and DenseNet121	✓	Sensitivity = 98 (± 3) Specificity = 90
(19)	2021	SARS-COV-2 CT scan dataset	1,252	1,230	ADECO-CNN	✗	Accuracy = 98.99
(21)	2021	COVID-CS	68,626	75,541	Novel joint classification and segmentation along fine-grained pixel-level labels of opacifications	✗	Sensitivity = 95 Specificity = 93
(22)	2021	CT scan dataset	349	397	Four models for feature extraction and machine learning classifiers for classification	✗	Accuracy = 87.9
(32)	2021	XDC	94	94	Centralized-VGG16 + data augmentation. Centralized-ResNet50 + data augmentation.	✓	Sensitivity = 95.1 Specificity = 93.0 *Sensitivity = 96.8* Specificity = 96.2
(33)	2021	CRD	3,616	10,192	U-Net Modified U-Net	✓	Accuracy = 98.21 Accuracy = 98.63
Ours	2021	CXI	200	5,000	WEENet	✓	Sensitivity = 85.0 **Specificity = 100** **Accuracy = 99.5**
		XDC	94	94		✓	**Sensitivity = 100** Specificity = 95.7 Accuracy = 97.8
		CRD	3,616	10,192		✓	Sensitivity = 98.6 *Specificity = 99.6* *Accuracy = 99.3*

Here, sensitivity indicates the number of correctly classified positive samples over the total number of positive samples. Similarly, specificity represents the number of correctly classified negative samples over the total number of negative samples. Accuracy is a generic classification metric that indicates the total number of correct classifications over the total number of samples. Finally, ROC metric represents the relationship (indicated by symbol ~R~) between specificity and sensitivity.

3.4 Cross-Corpse Evaluation and Robustness Analysis

The generalization of a system plays an important role especially when dealing with uncertain computational environment, where data under the observation is semantically different from the data used for training the algorithm. Bearing this in mind, we proposed a new evaluation strategy called cross-corpse evaluation for validating the generalization and robustness of our proposed system in uncertain environment. In this new evaluation strategy, first, we evaluated the performance of our method against other SOTA on test sets of the same datasets used for training. While in the second round of experiments, we assessed the performance of the proposed approach compared with the underlined investigated CNNs on test sets of the datasets other than training datasets, which is termed as cross-corpse evaluation. The obtained quantitative results for both the same dataset and cross-corpse evaluation strategy are presented in **Tables 4**, **5**. It can be easily perceived that the obtained accuracy score for cross-corpse evaluation is comparatively lower than that of the original dataset, yet the accuracy scores indicate the better generalization performance. Furthermore, the reported quantitative results in **Tables 4**, **5** verify the overwhelming performance of our method by obtaining the highest accuracy across each dataset in both the same dataset and cross-corpse evaluation. We also evaluated the qualitative performance of our method against SOTA by doing classification on randomly collected images from the test sets of each experimented dataset. The prediction results for randomly selected images from each experiment dataset are shown in **Figure 6**, where it can be noticed that our method provides the best prediction results compared with other SOTA methods for COVID-19 classification.

3.5 Comparison With Other CNN Models for COVID-19 Classification

This section presents the comparative analysis of our proposed WEENet with other SOTA methods for COVID-19 classification on test sets of CXI, XDC, and CRD datasets. For comparative analysis, we evaluated the performance of our method and compared it with SOTA including MobileNet (38), NASNet-Mobile (39), VGG16 (34), ResNet101 (35), RestNet50 (35), VGG19 (34), and EfficientNet (31). To investigate the performance of our method and validate the effectiveness of the proposed data augmentation strategy, we conducted experiments on both original datasets and augmented datasets and compared the results with the SOTA methods. The obtained results for the original dataset are given in **Table 4**, where it can be perceived that our proposed WEENet outperforms all comparative CNNs on original datasets across each evaluation metric except NASNet-Mobile (39) that performs comparatively better than our method in terms of TP and FP on the CXI dataset. On the other hand, the obtained results on augmented datasets are given in **Table 5**, where it can be noticed that our proposed WEENet achieved the best results by overwhelming the

TABLE 4 | Performance comparison of several deep learning-based models over benchmark datasets.

Dataset name	Model	Original dataset								Cross-corpse evaluation
CXI (18)		TP↑	FP↓	FN↓	TN↑	Sensitivity↑	Specificity↑	Accuracy↑	ROC↑	Accuracy↑
	MobileNet (38)	12	24	485	515	0.024	0.955	0.508	0.424	0.435
	NASNet-Mobile (39)	**26**	**10**	254	746	0.092	0.986	0.745	0.734	0.714
	VGG16 (34)	17	19	380	620	0.042	0.970	0.614	0.546	0.597
	ResNet101 (35)	21	15	402	598	0.049	0.975	0.597	0.591	0.557
	ResNet50 (35)	19	17	341	659	0.052	0.974	0.654	0.593	0.604
	VGG19 (34)	16	20	258	742	0.058	0.973	0.731	0.593	0.710
	EfficientNet (31)	24	12	179	811	0.118	0.985	0.813	0.743	0.847
	WEENet	25	11	**97**	**903**	**0.200**	**0.988**	**0.895**	**0.799**	**0.887**
XDC (32)	MobileNet (38)	11	9	8	12	0.578	0.571	0.585	0.575	0.553
	NASNet-Mobile (39)	14	6	4	16	0.777	0.727	0.750	0.750	0.782
	VGG16 (34)	13	7	9	11	0.590	0.611	0.600	0.600	0.574
	ResNet101 (35)	16	4	6	14	0.723	0.777	0.750	0.750	0.745
	ResNet50 (35)	15	5	7	13	0.681	0.722	0.700	0.700	0.617
	VGG19 (34)	12	8	6	14	0.666	0.636	0.650	0.650	0.592
	EfficientNet (31)	18	2	4	16	0.818	0.888	0.850	0.850	0.817
	WEENet	**19**	**1**	**2**	**18**	**0.904**	**0.947**	**0.925**	**0.925**	**0.894**
CRD (33)	MobileNet (38)	231	132	466	554	0.331	0.807	0.567	0.590	0.517
	NASNet-Mobile (39)	309	54	274	746	0.530	0.930	0.762	0.791	0.698
	VGG16 (34)	244	119	436	584	0.358	0.830	0.598	0.622	0.578
	ResNet101 (35)	251	112	464	556	0.451	0.833	0.583	0.618	0.514
	ResNet50 (35)	223	140	499	521	0.308	0.788	0.538	0.563	0.482
	VGG19 (34)	187	176	238	782	0.440	0.816	0.700	0.641	0.697
	EfficientNet (31)	**329**	**34**	138	882	0.704	0.962	0.875	0.886	0.837
	WEENet	**329**	**34**	**84**	**936**	**0.796**	**0.964**	**0.914**	**0.912**	**0.894**

TABLE 5 | Performance comparison of several deep learning-based models over augmented datasets.

Dataset name	Model	Original dataset								Cross-corpse evaluation
CXI (18)		TP↑	FP↓	FN↓	TN↑	Sensitivity↑	Specificity↑	Accuracy↑	ROC↑	Accuracy↑
	MobileNet (38)	219	149	134	266	0.620	0.641	0.631	0.630	0.583
	NASNet-Mobile (39)	307	61	64	336	0.827	0.846	0.837	0.837	0.784
	VGG16 (34)	280	88	126	388	0.689	0.815	0.757	0.758	0.693
	ResNet101 (35)	249	119	101	299	0.711	0.715	0.713	0.712	0.647
	ResNet50 (35)	307	61	98	302	0.758	0.832	0.793	0.795	0.691
	VGG19 (34)	291	77	79	321	0.786	0.806	0.796	0.797	0.738
	EfficientNet (31)	339	29	37	363	0.901	0.926	0.914	0.914	0.883
	WEENet	**364**	**4**	**3**	**397**	**0.991**	**0.990**	**0.991**	**0.991**	**0.947**
XDC (32)	MobileNet (38)	61	33	39	55	0.610	0.625	0.617	0.617	0.587
	NASNet-Mobile (39)	83	11	19	75	0.813	0.873	0.840	0.840	0.816
	VGG16 (34)	71	23	30	64	0.703	0.735	0.718	0.718	0.687
	ResNet101 (35)	77	17	20	74	0.793	0.813	0.803	0.803	0.774
	ResNet50 (35)	72	22	19	75	0.791	0.773	0.781	0.782	0.698
	VGG19 (34)	68	26	20	74	0.772	0.740	0.755	0.755	0.714
	EfficientNet (31)	88	6	12	82	0.880	0.931	0.904	0.904	0.817
	WEENet	**93**	**1**	**2**	**92**	**0.979**	**0.989**	**0.984**	**0.984**	**0.974**
CRD (33)	MobileNet (38)	759	261	402	618	0.653	0.703	0.675	0.675	0.597
	NASNet-Mobile (39)	822	198	174	846	0.825	0.810	0.817	0.818	0.798
	VGG16 (34)	784	236	433	587	0.644	0.713	0.672	0.672	0.586
	ResNet101 (35)	833	187	116	904	0.877	0.828	0.851	0.851	0.880
	ResNet50 (35)	751	269	478	542	0.611	0.668	0.633	0.634	0.595
	VGG19 (34)	798	222	198	822	0.801	0.787	0.794	0.794	0.742
	EfficientNet (31)	947	73	103	917	0.901	0.926	0.913	0.914	0.847
	WEENet	**1004**	**16**	**19**	**1001**	**0.981**	**0.984**	**0.982**	**0.983**	**0.914**

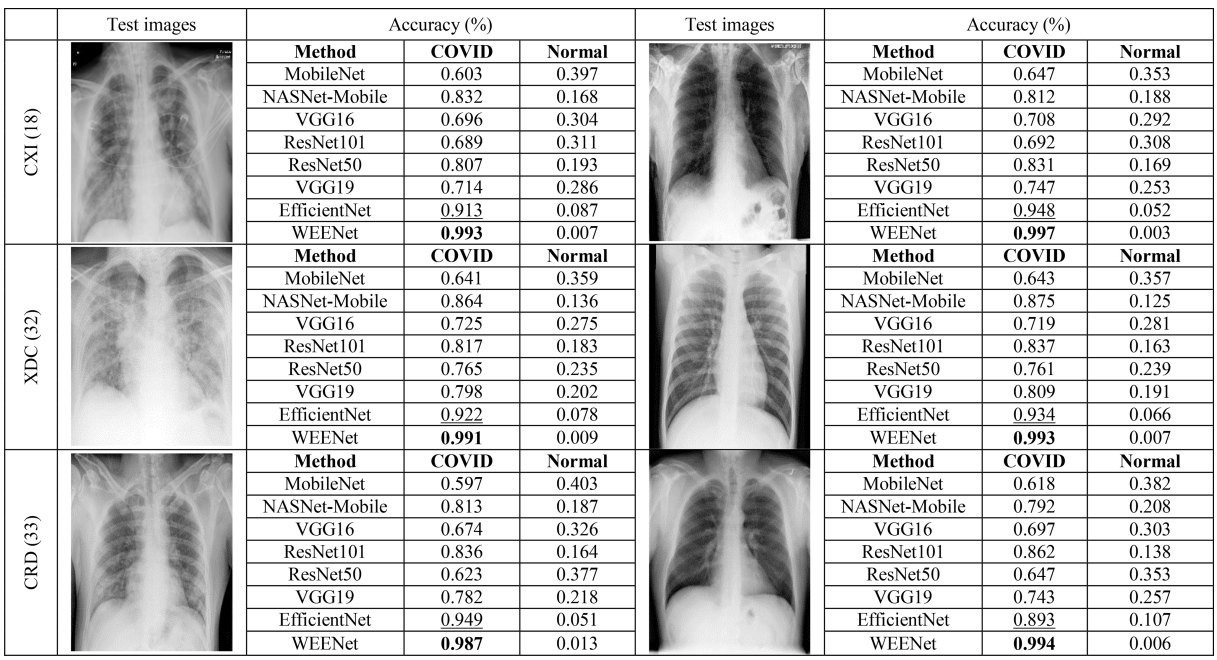


FIGURE 6 | Robustness analysis of our proposed method against other SOTA CNNs on randomly selected test images from each dataset and the corresponding predictions made by each method.

SOTA CNNs across each evaluation metric, thus showing its superiority and efficiency for COVID-19 classification in medical chest x-ray images. We also compared our WEENet architecture with other SOTA CNN-based COVID-19 classification approaches and reported the results in **Table 6**. The reported results reflect the dominancy of our WEENet on CRD dataset across each evaluation metric Although our method obtained comparatively lower values for sensitivity and specificity on the CXI and XDC datasets, still our method attained best results on the same datasets across the other two evaluation metrics. The

TABLE 6 | Performance comparison of the proposed WEENet with other baseline models.

Dataset	Model	Original dataset		
CXI (18)		Sensitivity↑	Specificity↑	Accuracy↑
	SqueezeNet (40)	**0.980**	0.920	–
	WEENet	0.850	**1.0**	**0.995**
XDC (32)	ResNet50 (35)	0.981	**0.958**	0.970
	WEENet	**1.0**	0.957	**0.978**
CRD (33)	ChestNet (41)	0.962	0.972	0.962
	WEENet	**0.986**	**0.996**	**0.993**

best reported results presented in **Tables 3–6** are highlighted in the bold text while the runner up scores are indicated in the underlined text. Furthermore, some visual results of the proposed WEENet over test set of each dataset are given in **Figure 7**.

3.6 Feasibility Analysis for 5G-Enabled IoMT Environment

Considering the requirements of 5G-enabled IoMT environment for rapid and accurate smart healthcare systems (42–44), it is essential to analyze the feasibility of a system before deploying in the real world. The feasibility assessment protocols involved different steps to investigate the suitability of a given system for the problem under observation in various aspects such as the robustness of decision-making system, automation, real-time response, and employability on edge-computing platforms. Having this in mind, we conducted feasibility analysis experiments and investigated our proposed WEENet in the abovementioned aspects. Based on the obtained quantitative results in the previous section, we estimated the robustness of our WEENet by averaging the attained accuracy score across all datasets and achieved an average of 90% accuracy. Next, the proposed method meets automation requirements thereby providing fully end-to-end deep learning system. Although, our method takes relatively more time for diagnosing COVID-19 in chest x-ray image, it has limited memory storage requirements for deployment on edge devices, making it a suitable approach for early COVID-19 detection in 5G-enabled IoMT environments. The conducted feasibility assessment findings are depicted in **Figure 8**.

3.7 WEENet for Lung Cancer Detection

In this section, we discuss the effectiveness and reusability of our proposed WEENet framework for early detection of lung cancer in chest CT scan images of infected patients. The deep learning-based early detection of lung cancer (45) can greatly facilitate the doctors and other medical-related individuals to eliminate the cancer cell at first place by providing proper care and treatment to the infected patients. Considering the relevancy in the image data (chest CT scan images) used for COVID-19 detection and lung cancer CT scan images (46), the proposed WEENet framework can be used for lung cancer detection by fine-tunning the architecture on lung cancer image data using transfer learning strategy (47). For efficient retraining of the WEENet architecture, the trained weights (already learned knowledge during training on COVID-19 image data) can be

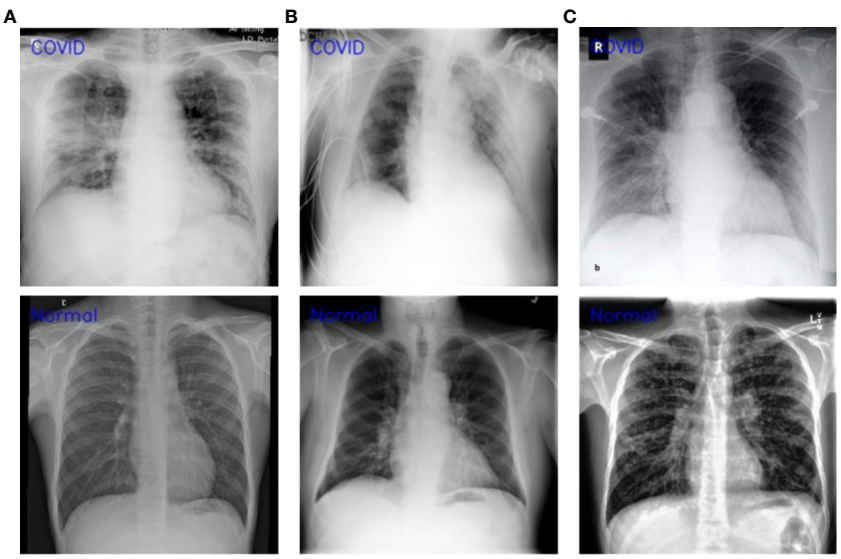


FIGURE 7 | Visual results of WEENet over each dataset **(A)** CXI, **(B)** XDC, and **(C)** CRD datasets.

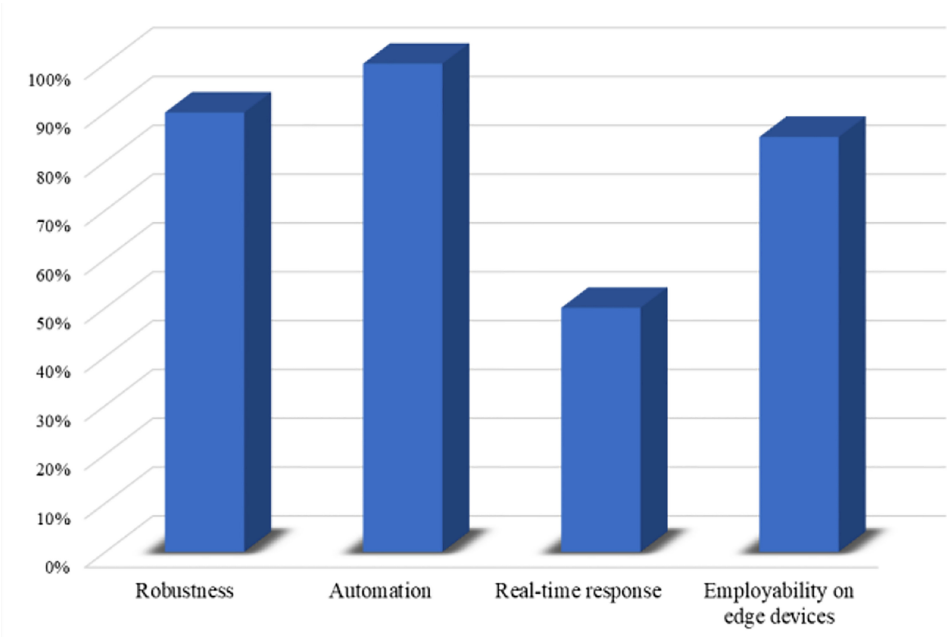


FIGURE 8 | Feasibility assessment of our proposed WEENet for 5G-enabled IoMT environment.

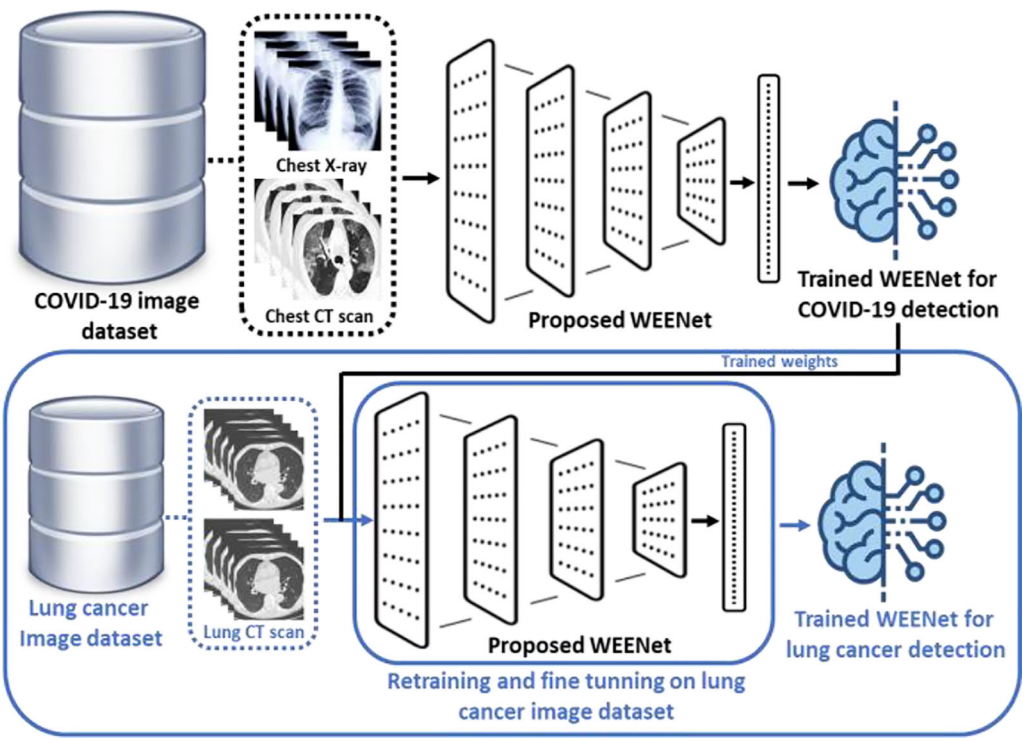


FIGURE 9 | The graphical overview of the reusability process of our proposed WEENet for lung cancer detection task.

used while training the proposed WEENet on lung cancer image data. The utilization of trained weights will not only reduce the training efforts (in terms of training time) but can also improve the performance of retrained architecture for lung cancer detection. The reusability workflow of our proposed WEENet for lung cancer detection task is depicted in **Figure 9**.

4 CONCLUSION AND FUTURE WORK

The COVID-19 pandemic started in 2019 and has severely affected human life and the world economy for which different actions are initiated to stop its spread and efficiently handle the pandemic. Such actions include the concept of smart lockdown, development of new devices for temperature checking, early detection of COVID-19 using medical imaging techniques, and treatment plans for patients with different risk levels. This work supports the necessary action of early COVID-19 detection using medical chest x-ray images in 5G-enabled IoMT environment, contributing to the management of COVID-19 pandemic. Considering the limited available medical imaging data and different conflicting performance metrics for early COVID-19 detection, in this work, we investigated deep learning-based frameworks for effective modeling and early diagnosis of COVID-19 from medical chest x-ray images in IoMT-enabled environment. We proposed "WEENet" for COVID-19 diagnosis using efficient CNN architecture and evaluated its performance on three benchmark medical chest x-ray and CT image datasets using eight different evaluation metrics such as accuracy, ROC, robustness, specificity, and sensitivity etc. We also tested the performance of our method using cross-corpse evaluation strategy. Our results are encouraging against SOTA methods and will support healthcare authorities in analyzing medical chest x-ray images of infected patients and will assist the management of the COVID-19 pandemic in IoMT environments.

The reported results are better than SOTA methods, but model size is not the best among all methods under consideration (though better than majority of the models). This is due to some of the architectural layers, tuned to balance the performance metrics towards optimization. More investigation is needed to further reduce the model size without affecting the performance, which is one of our plans. We also plan to extend this work to a multiclass problem including mild, moderate, and severe as discussed in COVIDGR dataset (48) from the University of Granada, Spain.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

KM, HU, and ZK contributed to the idea conceptualization, data acquisition, implementation, experimental assessment, manuscript writing, and revision. AS contributed to the data acquisition and manuscript revision. AA, MA, MK, MHAH contributed to acquisition and interpretation of data. KMM, MH, and MS contributed to the design of study and revision of manuscript.

FUNDING

This work was supported by the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number 959.

ACKNOWLEDGMENTS

The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number 959.

REFERENCES

1. Song F, Shi N, Shan F, Zhang Z, Shen J, Lu H, et al. Emerging 2019 Novel Coronavirus (2019-Ncov) Pneumonia. *Radiology* (2020) 295(1):210–7. doi: 10.1148/radiol.2020200274
2. WHO. *WHO Coronavirus (COVID-19) Dashboard* (2021). Available at: https://covid19.who.int/.
3. Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu JJR. Chest CT for Typical Coronavirus Disease 2019 (COVID-19) Pneumonia: Relationship to Negative RT-PCR Testing. *Radiology* (2020) 296(2):E41–5. doi: 10.1148/radiol.2020200343
4. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical Features of Patients Infected With 2019 Novel Coronavirus in Wuhan, China. *Lancet* (2020) 395(10223):497–506. doi: 10.1016/S0140-6736(20)30183-5
5. Liang X, Zhang Y, Wang J, Ye Q, Liu Y, Tong JJF I. M. Diagnosis of COVID-19 Pneumonia Based on Graph Convolutional Network. *Front Med* (2021) 7:1071. doi: 10.3389/fmed.2020.612962
6. Heidarian S, Afshar P, Enshaei N, Naderkhani F, Rafiee MJ, Fard FB, et al. Covid-Fact: A Fully-Automated Capsule Network-Based Framework for Identification of Covid-19 Cases From Chest Ct Scans. *Front Artif Intell* (2021) 4:598932. doi: 10.3389/frai.2021.598932
7. Bogoch II, Watts A, Thomas-Bachli A, Huber C, Kraemer MU, Khan K. Pneumonia of Unknown Aetiology in Wuhan, China: Potential for International Spread via Commercial Air Travel. *J Travel Med* (2020) 27(2): taaa008. doi: 10.1093/jtm/taaa008
8. Yoo SH, Geng H, Chiu TL, Yu SK, Cho DC, Heo J, et al. Deep Learning-Based Decision-Tree Classifier for COVID-19 Diagnosis From Chest X-Ray Imaging. *Front Med* (2020) 7:427. doi: 10.3389/fmed.2020.00427
9. Shi H, Han X, Jiang N, Cao Y, Alwalid O, Gu J, et al. Radiological Findings From 81 Patients With COVID-19 Pneumonia in Wuhan, China: A Descriptive Study. *Lancet Infect Dis* (2020) 20(4):425–34. doi: 10.1016/S1473-3099(20)30086-4
10. Mohammad-Rahimi H, Nadimi M, Ghalyanchi-Langeroudi A, Taheri M, Ghafouri-Fard S. Application of Machine Learning in Diagnosis of COVID-19 Through X-Ray and CT Images: A Scoping Review. *Front Cardiovasc Med* (2021) 8:185. doi: 10.3389/fcvm.2021.638011
11. Hu S, Gao Y, Niu Z, Jiang Y, Li L, Xiao X, et al. Weakly Supervised Deep Learning for Covid-19 Infection Detection and Classification From Ct Images. *IEEE Access* (2020) 8:118869–83. doi: 10.1109/ACCESS.2020.3005510
12. Ye Q, Xia J, Yang G. Explainable AI For COVID-19 CT Classifiers: An Initial Comparison Study. *arXiv Preprint arXiv* (2021). doi: 10.1109/CBMS52027.2021.00103

13. Narin A, Kaya C, Pamuk ZJPAApplications. Automatic Detection of Coronavirus Disease (Covid-19) Using X-Ray Images and Deep Convolutional Neural Networks. *Pattern Anal Appl* (2021) 24:1207–20. doi: 10.1007/s10044-021-00984-y
14. Ismael AM, Şengür A. Deep Learning Approaches for COVID-19 Detection Based on Chest X-Ray Images. *Expert Syst Appl* (2021) 164:114054. doi: 10.1016/j.eswa.2020.114054
15. Tang S, Wang C, Nie J, Kumar N, Zhang Y, Xiong Z, et al. EDL-COVID: Ensemble Deep Learning for COVID-19 Cases Detection From Chest X-Ray Images. *IEEE Trans Ind Inf* (2021) 17(9):6539–49. doi: 10.1109/TII.2021.3057683
16. Jain R, Gupta M, Taneja S, Hemanth D. J. J. A. I. Deep Learning Based Detection and Analysis of COVID-19 on Chest X-Ray Images. *Appl Intell* (2021) 51(3):1690–700. doi: 10.1007/s10489-020-01902-1
17. Wang L, Zhang Y, Wang D, Tong X, Liu T, Zhang S, et al. Artificial Intelligence for COVID-19: A Systematic Review. *Front Med* (2021) 8:1457. doi: 10.3389/fmed.2021.704256
18. Minaee S, Kafieh R, Sonka M, Yazdani S, Soufi GJ. Deep-Covid: Predicting Covid-19 From Chest X-Ray Images Using Deep Transfer Learning. *Med Image Anal* (2020) 65:101794. doi: 10.1016/j.media.2020.101794
19. Castiglione A, Vijayakumar P, Nappi M, Sadiq S, Umer M. COVID-19: Automatic Detection of the Novel Coronavirus Disease From CT Images Using an Optimized Convolutional Neural Network. *IEEE Trans Ind Inf* (2021) 17(9):6480–8. doi: 10.1109/TII.2021.3057524
20. Oh Y, Park S, Ye JC. Deep Learning Covid-19 Features on Cxr Using Limited Training Data Sets. *IEEE Trans Med Imaging* (2020) 39(8):2688–700. doi: 10.1109/TMI.2020.2993291
21. Wu Y-H, Gao S-H, Mei J, Xu J, Fan D-P, Zhang R-G, et al. Jcs: An Explainable Covid-19 Diagnosis System by Joint Classification and Segmentation. *IEEE Trans Image Process* (2021) 30:3113–26. doi: 10.1109/TIP.2021.3058783
22. Shamsi A, Asgharnezhad H, Jokandan SS, Khosravi A, Kebria RM, Nahavandi D, et al. An Uncertainty-Aware Transfer Learning-Based Framework for COVID-19 Diagnosis. *IEEE Trans Netural Networks* (2021) 32(4):1408–17. doi: 10.1109/TNNLS.2021.3054306
23. Zhang W, Lu Q, Yu Q, Li Z, Liu Y, Lo SK, et al. Blockchain-Based Federated Learning for Device Failure Detection in Industrial IoT. *IEEE Internet Things J* (2020) 8(7):5926–37. doi: 10.1109/JIOT.2020.3032544
24. Lu H, He X, Du M, Ruan X, Sun Y, Wang K. Edge QoE: Computation Offloading With Deep Reinforcement Learning for Internet of Things. *IEEE Internet Things J* (2020) 7(10):9255–65. doi: 10.1109/JIOT.2020.2981557
25. Du M, Wang K, Chen Y, Wang X, Sun Y. J. I. C. M. Big Data Privacy Preserving in Multi-Access Edge Computing for Heterogeneous Internet of Things. *IEEE Commun Mag* (2018) 56(8):62–7. doi: 10.1109/MCOM.2018.1701148
26. Xu C, Wang K, Sun Y, Guo S, Zomaya NSEngineering. Redundancy Avoidance for Big Data in Data Centers: A Conventional Neural Network Approach. *IEEE Trans Network Sci Eng* (2018) 7(1):104–14. doi: 10.1109/TNSE.2018.2843326
27. Parah SA, Kaw JA, Bellavista P, Loan NA, Bhat GM, Muhammad K, et al. Efficient Security and Authentication for Edge-Based Internet of Medical Things. *IEEE Internet Things J* (2020) 8(21):15652–62. doi: 10.1109/JIOT.2020.3038009
28. Dourado CM, Da Silva SPP, Da Nóbrega RVM, Rebouças Filho PP, Muhammad K, De Albuquerque VHC. An Open IoHT-Based Deep Learning Framework for Online Medical Image Recognition. *IEEE J Selected Areas Commun* (2020) 39(2):541–8. doi: 10.1109/JSAC.2020.3020598
29. Bellavista P, Torello M, Corradi A, Foschini L. Smart Management of Healthcare Professionals Involved in COVID-19 Contrast With SWAPS. *Front Sustain Cities* (2021) 03:73. doi: 10.3389/frsc.2021.638743
30. Zhang W, Zhou T, Lu Q, Wang X, Zhu C, Sun H, et al. Dynamic Fusion-Based Federated Learning for COVID-19 Detection. *IEEE Internet Things J* (2021) 8 (21):15884–91. doi: 10.1109/JIOT.2021.3056185
31. Tan M, Le Q. Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks. In: *International Conference on Machine Learning*. California: PMLR (2019). p. 6105–14.
32. Feki I, Ammar S, Kessentini Y, Muhammad K. J. A. S. C. Federated Learning for COVID-19 Screening From Chest X-Ray Images. *Appl Soft Comput* (2021) 106:107330. doi: 10.1016/j.asoc.2021.107330
33. Rahman T, Khandakar A, Qiblawey Y, Tahir A, Kiranyaz S, Kashem SBA, et al. Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection Using Chest X-Ray Images. *Comput Biol* (2021) 132:104319. doi: 10.1016/j.compbiomed.2021.104319
34. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv* (2014).
35. He K, Zhang X, Ren S, Sun J. (2016). Deep Residual Learning for Image Recognition, in: *29th IEEE Conference on Computer Vision and Pattern Recognition*. (Las Vegas: IEEE) pp. 770–8.
36. Tan M, Chen B, Pang R, Vasudevan V, Sandler M, Howard A, et al. Mnasnet: Platform-Aware Neural Architecture Search for Mobile. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. California: IEEE (2019). p. 2820–8.
37. Kramer MA. Nonlinear Principal Component Analysis Using Autoassociative Neural Networks. *AIChE J* (1991) 37(2):233–43. doi: 10.1002/aic.690370209
38. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv preprint arXiv* (2017).
39. Zoph B, Vasudevan V, Shlens J, Le QV. Learning Transferable Architectures for Scalable Image Recognition In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City: IEEE (2018). pp. 8697–710.
40. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K. SqueezeNet: AlexNet-Level Accuracy With 50x Fewer Parameters and< 0.5 MB Model Size. *arXiv preprint arXiv* (2016).
41. Wang H, Xia Y. Chestnet: A Deep Neural Network for Classification of Thoracic Diseases on Chest Radiography. *arXiv preprint arXiv* (2018).
42. Rachakonda L, Bapatla AK, Mohanty SP, Kougianos E. SaYoPillow: Blockchain-Integrated Privacy-Assured IoMT Framework for Stress Management Considering Sleeping Habits. *IEEE Trans Consumer Electron* (2020) 67(1):20–9. doi: 10.1109/TCE.2020.3043683
43. Salem O, Alsubhi K, Shaafi A, Gheryani M, Mehaoua A, Boutaba R. Man in the Middle Attack Mitigation in Internet of Medical Things. *IEEE Trans Ind Inf* (2021) 18:2053–62. doi: 10.1109/TII.2021.3089462
44. Sun Y, Liu J, Yu K, Alazab M, Lin K. PMRSS: Privacy-Preserving Medical Record Searching Scheme for Intelligent Diagnosis in IoT Healthcare. *IEEE Trans Ind Inf* (2021) 18:1981–90. doi: 10.1109/TII.2021.3070544
45. Ibrahim DM, Elshennawy NM, Sarhan AM. Deep-Chest: Multi-Classification Deep Learning Model for Diagnosing COVID-19, Pneumonia, and Lung Cancer Chest Diseases. *Comput Biol Med* (2021) 132:104348. doi: 10.1016/j.compbiomed.2021.104348
46. Quarato CMI, Mirijello A, Maggi MM, Borelli C, Russo R, Lacedonia D, et al. Lung Ultrasound in the Diagnosis of COVID-19 Pneumonia: Not Always and Not Only What Is COVID-19 "Glitters". *Front Med* (2021) 8. doi: 10.3389/fmed.2021.707602
47. Shrey SB, Hakim L, Kavitha M, Kim HW, Kurita T. Transfer Learning by Cascaded Network to Identify and Classify Lung Nodules for Cancer Detection. In: *International Workshop on Frontiers of Computer Vision*. Ibusuki: Springer (2020). p. 262–73.
48. D. A. R. I. i. D. S. a. C. Intelligence. *COVIDGR*. Available at: https://dasci.es/es/transferencia/open-data/covidgr/.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Muhammad, Ullah, Khan, Saudagar, AlTameem, AlKhathami, Khan, Abul Hasanat, Mahmood Malik, Hijji and Sajjad. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 08 February 2022
doi: 10.3389/fonc.2022.822756

Baseline Amide Proton Transfer Imaging at 3T Fails to Predict Early Response to Induction Chemotherapy in Nasopharyngeal Carcinoma

Zhou Liu[1,2†], Liyan Zou[1†], Qian Yang[1], Long Qian[3], Tianran Li[1], Honghong Luo[1], Canwen Che[1], Yuanyuan Lei[4], Peng Chen[5], Chunyan Qiu[5], Xin Liu[2], Yin Wu[2] and Dehong Luo[1,6*]*

[1] Department of Radiology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen, China, [2] Paul C. Lauterbur Research Center for Biomedical Imaging, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, [3] MR Research, GE Healthcare, Beijing, China, [4] Department of Pathology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen, China, [5] Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen, China, [6] Department of Radiology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

OPEN ACCESS

Edited by:
Kyung Hyun Sung, UCLA Health System, United States

Reviewed by:
Nguyen Minh Duc, Pham Ngoc Thach University of Medicine, Vietnam
Zheng Han, Johns Hopkins University, United States

***Correspondence:**
Dehong Luo
13926236152@163.com
Yin Wu
yin.wu@siat.ac.cn

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *26 November 2021*
Accepted: *14 January 2022*
Published: *08 February 2022*

Citation:
Liu Z, Zou L, Yang Q, Qian L, Li T, Luo H, Che C, Lei Y, Chen P, Qiu C, Liu X, Wu Y and Luo D (2022) Baseline Amide Proton Transfer Imaging at 3T Fails to Predict Early Response to Induction Chemotherapy in Nasopharyngeal Carcinoma. Front. Oncol. 12:822756. doi: 10.3389/fonc.2022.822756

Background: Early identification of nasopharyngeal carcinoma (NPC) patients with high risk of failure to induction chemotherapy (IC) would facilitate prompt individualized treatment decisions and thus reduce toxicity and improve overall survival rate. This study aims to investigate the value of amide proton transfer (APT) imaging in predicting short-term response of NPC to IC and its potential correlation with well-established prognosis-related clinical characteristics.

Methods and Materials: A total of 80 pathologically confirmed NPC patients receiving pre-treatment APT imaging at 3T were retrospectively enrolled. Using asymmetry analysis, APT maps were calculated with mean (APT_{mean}), 90th percentile (APT_{90}) of APT signals in manually segmented NPC measured. APT values were compared among groups with different histopathological subtypes, clinical stages (namely, T, M, N, and overall stages), EBV-related indices (EBV-DNA), or responses to induction chemotherapy, using Mann–Whitney U test or Kruskal–Wallis H test.

Results: NPC showed significantly higher APT_{mean} than normal nasopharyngeal tissues (1.81 ± 0.62% *vs.* 1.32 ± 0.56%, P <0.001). APT signals showed no significant difference between undifferentiated and differentiated NPC subtypes groups, different EBV-DNA groups, or among T, N, M stages and overall clinical stages of II, III, IVA and IVB (all P >0.05). Similarly, baseline APT-related parameters did not differ significantly among

different treatment response groups after IC, no matter if evaluated with RECIST criteria or sum volumetric regression ratio (SVRR) (all P >0.05).

Conclusion: NPC showed significantly stronger APT effect than normal nasopharyngeal tissue, facilitating NPC lesion detection and early identification. However, stationary baseline APT values exhibited no significant correlation with histologic subtypes, clinical stages and EBV-related indices, and showed limited value to predict short-term treatment response to IC.

Keywords: amide proton transfer (APT), nasopharyngeal carcinoma (NPC), Epstein–Barr virus (EBV), clinical stage, induction chemotherapy

INTRODUCTION

The number of the newly diagnosed nasopharyngeal carcinoma (NPC) patients has been reported to be over 129,000 worldwide in 2018, among which 70% is aggregated in East Asia and Southeast Asia (1). In particular, the incidence is as high as 9.69/100,000 in Southeastern China (2). Currently, induction chemotherapy (IC) followed by concurrent chemoradiotherapy (CCRT) has become the standard treatment for locoregionally NPC for its better prognosis than CCRT alone (3, 4). However, around 10% patients still do not respond well to IC. Identification of those non-responders in advance would facilitate more prompt personalized treatment decisions, avoiding chemotherapy toxicity and unnecessary costs. In addition, tumor response to IC has been identified as an independent prognostic factor for long-term survival (5, 6). Hence, pretreatment prediction of short-term response to IC is essential for better individualized management of NPC patients.

Due to substantial biological heterogeneity, NPCs with the same clinical stage might have completely different responses to IC. Identification of biomarkers that could reflect tumor microenvironment characteristics may benefit the prediction of treatment response. With the advancement of morphological and function imaging techniques, MRI has been widely used for NPC evaluation by providing information at both macroscopic and microscopic levels (7–10). Recently, APT-weighted (APT_w) imaging has emerged as a powerful tool to probe the chemical exchange saturation transfer (CEST) effect between endogenous proteins/peptide amide protons and bulk water (11). Because of overexpressed proteins in tumors, APT_w imaging has been increasingly used in tumor detection, grading, and treatment response evaluation of various types of cancer (12–26). Although APT_w imaging has shown promising efficacy in predicting middle-term outcome after complete standard treatment (induction chemotherapy followed by concurrent chemoradiotherapy) (26) or 2-year long-term outcome prediction in NPC patients (27), little is known about whether APT signal could predict short-term treatment response to IC only.

Besides, NPC can be classified into distinct subtypes by well-established prognosis-related characteristics, e.g. Epstein–Barr virus (EBV) infection status (28), different histopathological subtypes (29) and clinical stages, etc. It remains unknown whether these distinct NPC subtypes with different prognosis have different protein metabolism levels, which potentially could be quantified by APT_w imaging. Therefore, if there is any association between baseline APT measurements and well-established prognosis-related clinical characteristics of NPC, such as histopathological subtypes, clinical stages, and status of Epstein–Barr (EB) virus infection, the value of APT in predicting prognosis would be further strengthened.

In this study, we aimed to investigate the potential value of APT_w imaging in predicting short-term response to IC and its potential association with prognosis-related histopathological subtypes, clinical stages, and EBV-related indices.

MATERIALS AND METHODS

Patient Enrollment

This retrospective study was approved by the local ethics committee of the Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences, and informed consent was waived. From December 2019 to August 2021, 114 patients with suspected NPC who had undergone baseline APT-weighed imaging were enrolled for further selection. The inclusion and exclusion criteria were as follows: 1) biopsy confirmed NPC; 2) the maximal diameter of the primary nasopharyngeal lesion was larger than 1 cm; 3) conventional MRI and APT-weighted imaging were performed before treatment; 4) each patient had received at least one examination for clinical staging work-up, such as chest X-ray, abdominal ultrasonography, chest and abdomen CT, bone SPECT scan, or whole-body PET-CT scan. Clinical staging for each NPC patient was performed based on the 8th edition of the American Joint Committee on Cancer (AJCC) (30); APT-weighted images were reviewed by a board-certified radiologist (ZL with more than 10 years of experience in head and neck imaging) and those with degraded image quality,

Abbreviations: NPC, nasopharyngeal carcinoma; APT, amide proton transfer; CEST, chemical exchange saturation transfer; EBV, Epstein–Barr virus; AJCC, American Joint Committee on Cancer; WASSR, water saturation shift referencing; APT_{mean}, Mean value of APT asymmetry; APT_{90}, 90th percentile of APT asymmetry; RECIST, Response Evaluation Criteria in Solid Tumors; SVRR, sum volumetric regression ratio; EBER, EBV-encoded small RNA; ISH, *in situ* hybridization; EBV-DNA, Epstein–Barr deoxyribonucleic acid; EA-IgA, early antigen antibody immunoglobulin A; VCA-IgA, viral capsid antibody immunoglobulin A; MTR, magnetization transfer ratio.

namely, overt signal lost, motion artifacts, geometric distortion, were excluded from analysis.

Treatment Strategy

All NPC patients received standard treatment. For NPC patients with clinical stages of I and II, concurrent chemoradiotherapy or radiotherapy alone was given, whereas IC followed by concurrent chemoradiotherapy were prescribed to patients with clinical stages of III and IV. The brief treatment procedure was: 1) IC (GP regimen): Gemcitabine (1 g/m^2, Days 1 and 8) + cisplatin (75 mg/m^2, Day 1); or Paclitaxel (135–175 mg/m^2, Day 1) + cisplatin (75 mg/m^2, Day 1), 3 weeks per cycle × 2–3 cycles; 2) Concurrent chemoradiotherapy: cisplatin (30 mg/m^2, Day 1, once a week for 7 times, iv) and concurrent radiation therapy (in a total of 6.6–7.6 weeks, accumulated dose of 66–78 Gy for primary lesion, while 60–70 Gy for regional lymph nodes). With intensity-modulated radiation therapy being used, conventional 6MV-X fractionated radiotherapy (200 cGy each time × 1 time per day × 5 days per week) was performed, which covers the primary NPC lesion and bilateral cervical lymphatics at risk.

Follow-Up and Study Endpoint

In this study, the endpoint was set between after IC completion and before next-step concurrent chemoradiotherapy. Standard clinical MRI follow-up was performed at the endpoint for each subject.

MRI Study

MRI study was performed on a 3T MR scanner (Discovery MR 750w, General Electric) using a 32-channel head and neck joint coil. Local shimming was placed to cover the whole nasopharyngeal carcinoma to reduce static magnetic field inhomogeneity. Referencing to T2-weighted images, a single-slice axial plane covering the largest primary nasopharyngeal carcinoma was selected for APT imaging with chemical shift-selective fat suppression and single-shot fast spin echo (FSE) readout (4 RF saturation pulses with duration of 500 ms for each pulse and without interpulse delay, RF saturation power B_1 of 2 μT, TR = 6.5 s with minimum TE, slice thickness = 4 mm, field of view = 22 × 22 cm^2, matrix = 128 × 128). In additional to an unsaturated scan, 27 CEST-weighted images with frequency offsets of 0, ± 0.5, ± 1.0, ± 1.5, ± 2.0, ± 2.5, ± 3.0, ± 3.5, ± 4.5, ± 5.0, ± 5.5, ± 6.0, ± 6.5, and ±7.0 ppm were acquired. The total acquisition time is approximately 4.5 min. A water saturation shift referencing (WASSR) map was collected with B_1 of 0.5 μT (frequency offsets between ±240 Hz with intervals of 48 Hz) to correct B_0 inhomogeneity.

Data Analysis

APT data was analyzed in MATLAB (MathWorks). Z-spectra (S) was normalized by the unsaturated scan (S_0), interpolated by smoothing splines, and corrected the field inhomogeneity with WASSR (31, 32). Pixel-wise APT-weighted value was calculated using asymmetry analysis of APT_w (APT-weighted) = S(−3.5 ppm)/S_0 − S(3.5 ppm)/S_0.

Region of interest (ROI) of primary NPC was manually delineated on the unsaturated image by a radiologist (LZ with 6 years of experience in head and neck imaging) using open-source ITK-SNAP software (http://itksnap.org) with referencing conventional anatomical images, namely, T1-weighted, T2-weighted, and contrast enhanced T1-weighted images. Overt hemorrhage, air space, necrosis and vessels were excluded. For patients with tumor confined in one side of nasopharynx without involving the midline of posterior wall of nasopharynx, the normal nasopharyngeal tissue was delineated as well. ROI delineation was then reviewed and approved by a senior radiologist (ZL with more than 10 years of experience in head and neck imaging). APT_w signals, namely, the average APT (APT_{mean}) and 90th percentile (APT_{90}), were obtained from each ROI.

Treatment Response Evaluation

The largest diameter of NPC was manually measured on the slice with the largest area of lesion on axial T_1-weighted contrast enhanced images by a board-certified radiologist (LZ with 6 years of experience in head and neck imaging) on the vendor workstation. The measurement was reviewed by a senior radiologist (ZL), with any discrepancies resolved through consensus discussion. Then, the treatment response of primary tumor was determined by calculating the relative change percentage of tumor maximal diameter between pre- and post-treatment, as $\Delta D = (D_{pre\text{-}treatment} - D_{post\text{-}treatment})/D_{pre\text{-}treatment}$, and classified as complete response (CR, ΔD = 100%), partial response (PR, ΔD ≥30%), stable disease (SD, −20%< ΔD <30%), and progression disease (PD, ΔD ≤−20%) according to the RECIST criteria version 1.1 (33).

Tumor volume was measured by manually delineating and summing NPC lesion from consecutive 2D slices on contrast-enhanced T_1-weighted image by LZ and then reviewed by a senior radiologist (ZL). The sum volumetric regression ratio (SVRR) was defined as the tumor volume reduction percentage before and after treatment, and 50% reduction in SVRR was arbitrarily chosen as the cut-off value.

Histopathology and Epstein–Barr (EB) Virus-Related Indices

The following pathological information were collected: 1) histology of nasopharyngeal carcinoma based on WHO classification, namely, nonkeratinizing squamous cell carcinoma (differentiated subtype and undifferentiated subtype), keratinizing squamous cell carcinoma and basaloid squamous cell carcinoma (34); 2) EBV-encoded small RNA (EBER) status using *in situ* hybridization (ISH). Serum test reports before treatment were also examined to obtain the following information: 1) copy numbers of viral Epstein–Barr deoxyribonucleic acid (EBV-DNA) in plasma samples before treatment, whereas copy numbers of 400 as threshold; 2) serum level of early antigen antibody (EA-IgA); and 3) serum level of viral capsid antibody (VCA-IgA).

Statistical Analysis

All the statistical analyses were performed using SPSS software v. 26.0 (IBM, Armonk, NY). Paired Student's t-test was performed to evaluate the difference in APT_w signals between normal and

cancerous nasopharyngeal tissues. Continuous variables were displayed as median (25th–75th percentile). Mann–Whitney U test was used between the two histopathological subtypes. Kruskal–Wallis H test was conducted to assess any significant differences among groups with different clinical stages or treatment responses. Receiver operating characteristic (ROC) analysis was performed and the area under the curve (AUC) was calculated. The optimal cutoff values were determined by maximizing the sum of sensitivity and specificity using YOUDEN index. Pearson's correlation analysis was used to evaluate the relationship between APT_w measurements and tumor regression rate.

RESULTS

A total of 80 patients were included in the analysis (**Figure 1**) with clinical characteristics summarized in **Table 1**.

Tissue Characterization

A total of 17 patients showed NPC confined in one side and not involving the midline of nasopharynx. **Figures 2A, B** illustrate APT_w maps and histograms in NPC and normal nasopharyngeal tissue regions in a 66-year-old male patient. The APT_{mean} of the NPC were significantly higher than that of corresponding normal nasopharyngeal tissue (1.81 ± 0.62% *vs.*1.32 ± 0.56%, P <0.001) (**Figure 2C**). However, APT_{90} showed no significant difference between NPC and corresponding normal nasopharyngeal tissue (3.68 ± 0.84% *vs.*3.25 ± 1.07%, P = 0.115). The optimal cutoff value of APT_{mean} in discriminating NPC with contralateral nasopharyngeal tissue was 1.66% with a sensitivity of 0.765, a specificity of 0.765, and an AUC of 0.727 (**Figure 2D**).

Among the enrolled 80 patients, 66 had undifferentiated carcinoma, 14 had differentiated carcinoma, and no patients had nonkeratinizing squamous cell carcinoma. All APT_w signals did not differ significantly between undifferentiated and differentiated NPC groups (all P >0.05) (**Figure 3**), as summarized in **Table 2**.

TNM Stage

No significant difference was found in APT_w signals among primary tumor stages of T1 to T4, among lymph node stages of N1 to N3 (**Figure 4**), between M0 and M1 stages, among overall clinical stages of II, III, IVA, and IVB (all P >0.05) (**Table 2**).

EBV-Related Indices

The 67 patients with available EBER status all showed positive EBER on the specimen using *in situ* hybridization in all cases. For the 76 cases with available serum EBV-DNA quantification, APT_w signals was found to be comparable between EBV-DNA positive group and EBV-DNA negative group (**Figure 5**). Similarly, there was no significant difference in APT_w signals between EA-IgA positive group (N = 18) and EA-IgA negative group (N = 16), and between VCA-IgA positive group (N = 27) and VCA-IgA negative group (N = 7) (all P >0.05) for the 34 NPC patients with available serum expression of EBV antibody (**Table 3**).

Treatment Response

Among 66 patients with available treatment response evaluation after overall treatment, 53 patients had received induction chemotherapy (**Table 4**). For patients receiving overall treatment, no significant correlation was found neither between APT_{mean} and SVRR (R^2 = −0.014, P = 0.738) (**Figure 6A**) nor between APT_{90} and SVRR (R^2 = −0.001, P = 0.339) (**Figure 6B**). Similarly, for patients who received induction chemotherapy, SVRR did not significantly correlate with neither APT_{mean} (R^2 = −0.019, P = 0.883) (**Figure 6C**), nor APT_{90} (R^2 = 0.008, P = 0.238) (**Figure 6D**). Baseline APT-related parameters did not differ significantly among different treatment

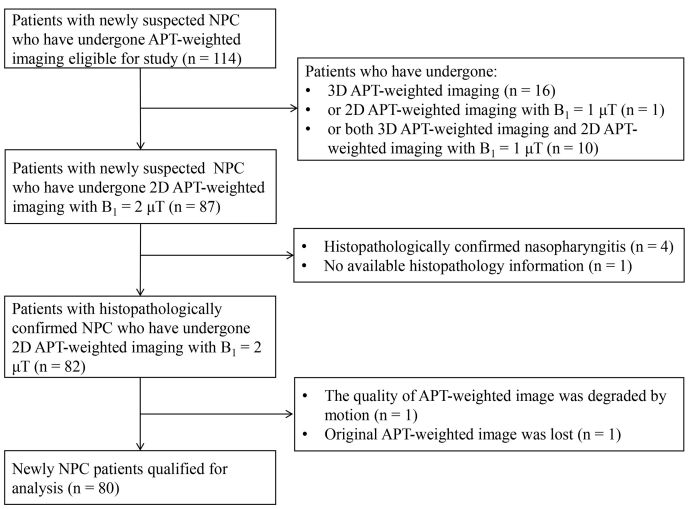


FIGURE 1 | Flowchart of patient enrollment.

TABLE 1 | Clinical characteristics of the enrolled 80 patients.

Clinical characteristics	Values/No. of patients
Age (years)	41.5 (34-51.25)
Gender	
Male	55
Female	25
Histology	
Nonkeratinizing squamous cell carcinoma	0
Undifferentiated carcinoma	66
Differentiated carcinoma	14
Primary tumor stage	
T1	16
T2	8
T3	41
T4	15
Nodal stage	
N0	1
N1	25
N2	25
N3	29
Metastatic stage	
M0	66
M1	11
Mx	3
Clinical stage	
I	0
II	5
III	33
IVA	27
IVB	12
Not available	3
Treatment strategy	
Radiotherapy only	1
Concurrent chemoradiotherapy	15
Induction chemotherapy and concurrent chemoradiotherapy	58
Treatment strategy not available	6
MRI scan	
Baseline MR scan	80
Follow-up MR scan after induction chemotherapy	53
Follow-up MR scan after radiotherapy only	1
Follow-up MR scan after concurrent chemoradiotherapy	12
No follow-up MR scan	14

response groups after overall treatment or after induction chemotherapy, no matter evaluated with the RECIST criteria or SVRR (all P >0.05). Besides, no significant difference in tumor regression rate was found between EBV-DNA (+) (n = 16) and EBV-DNA (−) (n = 37) group who received induction chemotherapy, as evaluated with the RECIST criteria (52.0 (38.3–57.2)% *vs.*58.0 (35.0–100.0)%, P = 0.407) and SVRR (67.0 (52.5–88)% *vs.* 79.0 (56.0–100.0)%, P = 0.441), respectively.

DISCUSSION

In this study, we investigated the value of pre-treatment APT_w signal in predicting short-term response after IC in NPC patients and its potential association with prognosis-related clinical characteristics, namely, histopathological subtypes, clinical stages, and EBV-related indices. Unfortunately, based on current result, the stationary baseline APT-weighted imaging showed limited value to predict response to IC and no significant correlation with different histopathological subtypes, EBV infection statuses and clinical stages.

The nasopharynx APT MRI is challenging due to susceptibility artifacts, and swallowing and jaw motion (22, 35). Our image quality was visually comparable to those shown in previous APT studies (22, 35) without apparent signal loss, geometric distortion, or any artifacts. In this study, significantly higher mean APT value in NPC tissue was observed than that in normal nasopharyngeal tissue, which could facilitate NPC lesion detection and early identification. The stronger APT effect may reflect rapidly proliferating tumor cells and abnormally active proteo-synthesis in cancerous tissues, as postulated previously (36).

Although APT-weighted imaging has been proven to be useful in differentiating head and neck malignant and benign tumors (22, 23), it was reported to be difficult in distinguishing malignant tumor types in head and neck regions (23). Our study showed that APT_w signals failed to discriminate between undifferentiated and differentiated NPC histology subtypes, which may imply similar protein synthesis status of these two histology subtypes before treatment. As for clinical stages, no significant difference of APT_w signals was shown among T, N, and M stages, and overall clinical stages.

EB-virus infection was recognized as the most common etiology of NPC, especially for nasopharyngeal non-keratinizing squamous cell carcinoma in endemic area (37). Consistently, all cases with available EBER *in situ* hybridization were tested EBER positive in our study. The serum EBV-DNA copy number has already been widely used to screen population with the high risk of developing NPC in endemic areas (38) and may serve as a biomarker to monitor and predict treatment response (39). However, pretreatment EBV-DNA status does not show apparent relationship with tumor regression rate after IC in this study. In contrast, a multi-center phase II randomized controlled study showed that an early decrease of EBV-DNA copy number indicates better treatment response (40). This might imply that monitoring the dynamic change of EBV-infection status throughout the treatment course using EBV-DNA copy number might be more effective than merely detecting a stationary pretreatment status in predicting treatment response. However, EBV-DNA copy number quantification data were not available after treatment in this retrospective study. Besides, although it seems no significant correlation of APT signals with EBV-related indices was presented, the correlation between APT_w signals change and EBV-DNA copy number alteration should be investigated in the future.

This study focused on the investigation of the value of APT imaging in predicting short-term survival of NPC patients. Although Qamar et al. (26, 27) demonstrated that the mean APT value change between pretreatment and intra-treatment 2 weeks later indeed could predict the middle-term outcome six months after the completion of standard treatment, both pretreatment and intra-treatment mean APT value did not

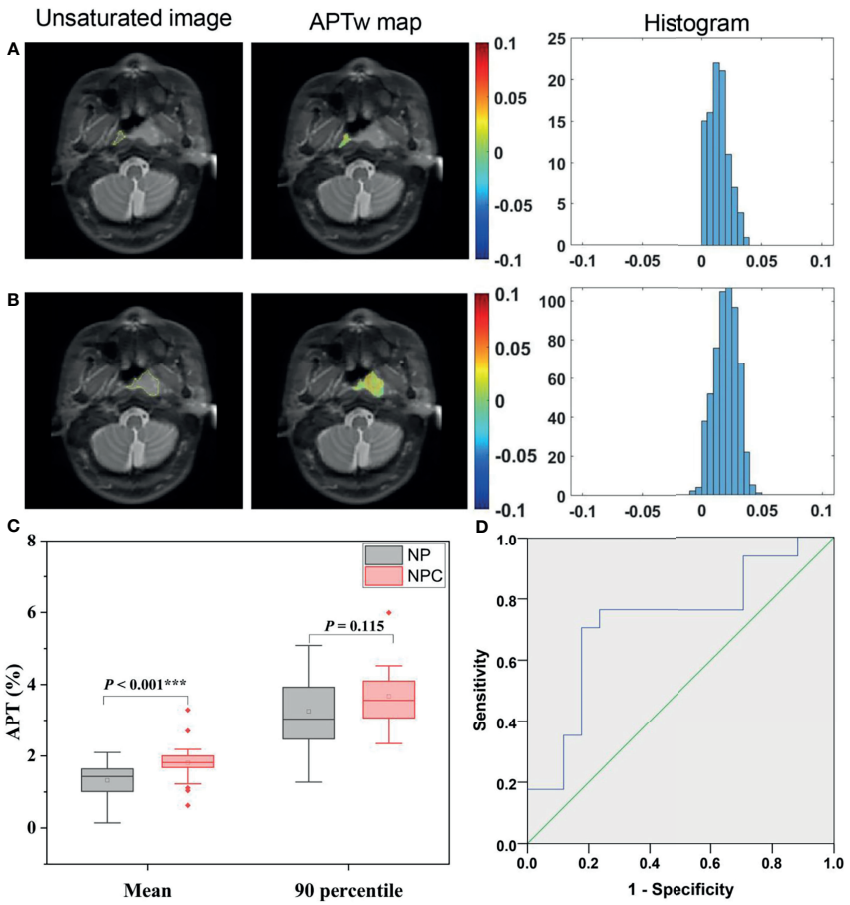


FIGURE 2 | Representative ROI delineation on unsaturated S0 images, APT_w maps, and histograms in NPC **(A)** and normal nasopharyngeal tissue **(B)** regions in a 66-year-old male patient. Comparison of APT_w signals, including APT_{mean}, APT_{90} **(C)** between NPC and corresponding normal nasopharynx tissues; ROC analyses of APT_{mean} in discriminating NPC and normal nasopharynx tissue **(D)**.

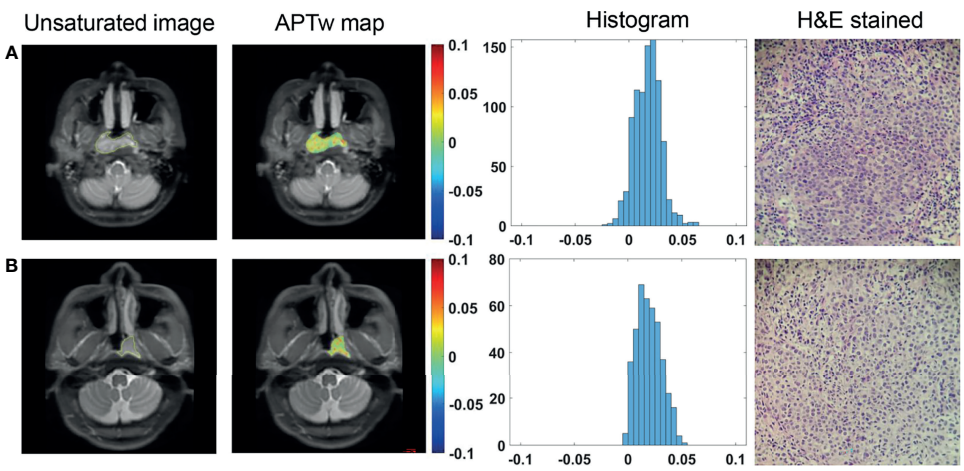


FIGURE 3 | ROI delineation on unsaturated S0 images with ROIs overlaid, APT_w maps and histograms, and H&E stained slides (40 × 10) of a 34-year-old female patient with differentiated NPC **(A)** and a 36-year-old male patient with undifferentiated NPC **(B)**.

TABLE 2 | APT-related parameters in different histology and TNM stage groups.

	No.	APT_{mean} (%)	APT_{90} (%)
Nonkeratinizing squamous cell carcinoma	0	–	–
Undifferentiated carcinoma	66	1.72 (1.36–2.10)	3.54 (3.05–4.04)
Differentiated carcinoma	14	1.77 (1.45–1.99)	3.20 (3.14–3.77)
***P*-value**	–	0.810	0.695
Primary tumor stage			
T1	16	1.81 (1.60–2.07)	3.70 (3.46–4.11)
T2	8	2.01 (1.72–2.42)	3.67 (3.23–4.52)
T3	41	1.64 (1.32–1.96)	3.39 (2.86–3.74)
T4	15	1.80 (1.31–2.26)	3.22 (3.03–4.16)
***P*-value**	–	0.330	0.103
Nodal stage			
N0	1	–	–
N1	25	1.76 (1.48–2.00)	3.44 (3.14–4.09)
N2	25	1.86 (1.48–2.14)	3.59 (3.06–4.28)
N3	29	1.72 (1.32–2.00)	3.45 (3.07–3.82)
***P*-value**	–	0.644	0.867
Metastatic stage			
M0	66	1.74 (1.48–2.01)	3.42 (3.05–4.03)
M1	11	1.53 (1.02–2.21)	3.74 (3.64–4.05)
Mx[a]	3	–	–
***P*-value**	–	0.432	0.048*
Overall tumor stage			
I	0	–	–
II	5	1.76 (1.62–1.80)	3.25 (2.89–4.04)
III	33	1.86 (1.58–2.14)	3.56 (3.14–4.09)
IVA	27	1.59 (1.36–1.91)	3.16 (2.95–3.58)
IVB	12	1.62 (1.08–2.18)	3.75 (3.67–4.10)
Not available	3	–	–
***P*-value**	–	0.343	0.068

$P < 0.05$ indicates significant difference. Mx[a], metastatic status undetermined.

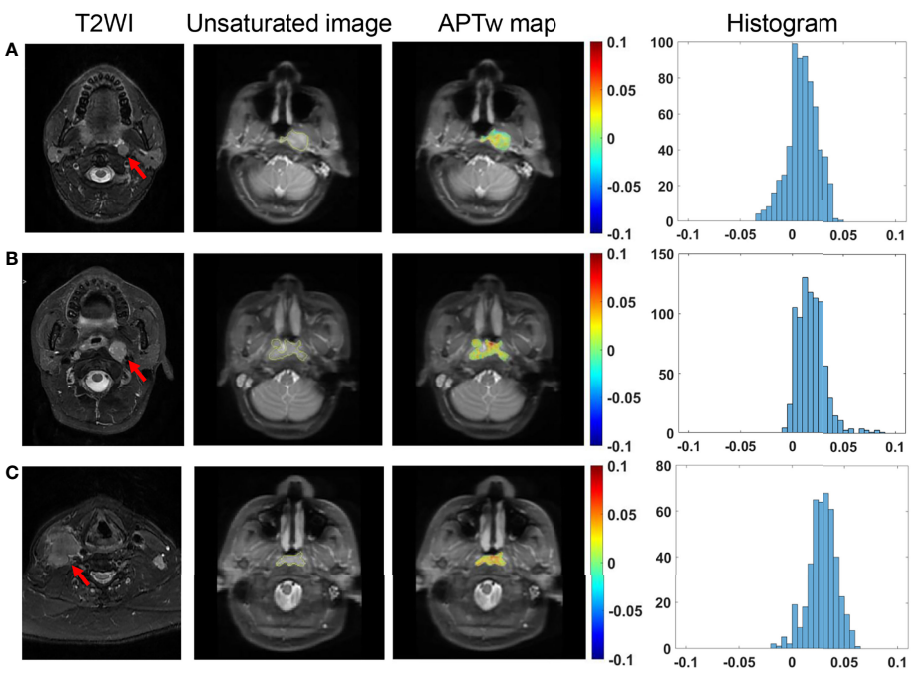


FIGURE 4 | Comparison of T2-weighted images of metastatic lymph nodes (red arrows), and unsaturated S0 images with ROI overlaid, APT_w maps and histograms in NPC tissues among different N stages. A 39-year-old male patient with N1 stage NPC **(A)**, a 33-year-old female patient with N2 stage NPC **(B)**, and a 65-year-old male patient with N3 stage NPC **(C)**.

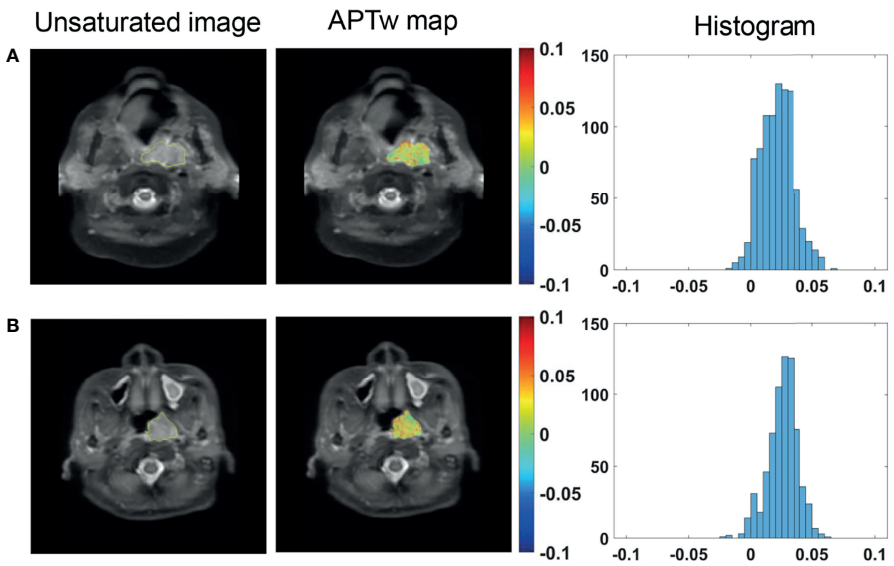


FIGURE 5 | Unsaturated S0 image with ROI overlaid, APT_w maps and histograms of a 65-year-old male NPC patient with serum EBV-DNA (+) status **(A)**, and a 47-year-old female NPC patient with serum EBV-DNA (–) status **(B)**.

TABLE 3 | APT-related parameters in groups with different EBV-related indices.

	No.	APT_{mean} (%)	APT_{90} (%)
EBV-DNA			
+	23	1.72 (1.48–2.04)	3.44 (3.12–4.04)
–	53	1.69 (1.28–2.06)	3.45 (3.05–3.78)
Not available	4	–	–
***P*-value**	–	0.995	0.730
EA-IgA			
+	18	1.70 (1.27–1.78)	3.14 (2.94–3.60)
–	16	1.84 (1.49–2.20)	3.54 (3.16–4.15)
Not available	46	–	–
***P*-value**	–	0.164	0.126
EB-VCA-IgA			
+	27	1.72 (1.59–1.74)	3.05 (2.92–3.38)
–	7	1.80 (1.59–2.00)	3.45 (2.92–4.01)
Not available	46	–	–
***P*-value**	–	0.531	0.206
EBER			
+	67	1.73 (1.40–2.08)	3.45 (3.05–4.03)
–	0	–	–
Not available	13	–	–

show significant difference between responders and non-responders. Such observations were consistent with our results that no significant difference in pretreatment APT_w signals was presented among different treatment response groups. This might also imply that monitoring the dynamic protein change throughout the treatment course using APT_w imaging may be more effective than detecting a stationary baseline pretreatment status in predicting treatment response.

Aside from the relatively limited sample size and imbalanced distribution in some groups, several limitations should be noted. First, conventional MTR asymmetry analysis was used to quantify APT effect, which would be contaminated with magnetization transfer and nuclear overhauser enhancement effects. However, the method has been histologically demonstrated to be applicable in tumor characterization (41). In addition, the APT_w signal in tumor has proved mostly attributed to the APT contribution under the currently utilized RF power of 2 μT (42). Therefore, the findings regarding APT are valid. Second, this study employed single-slice APT imaging to shorten the scan time, and the limited spatial coverage may be inadequate to characterize intra-tumoral heterogeneity. 3D APT scanning integrated with fast data sampling strategies would be desired for whole-tumor characterization in the future.

TABLE 4 | Baseline APT-related parameters for different treatment response group.

	No.	APT_{mean} (%)	APT_{90} (%)
Overall treatment response			
PD	0	–	–
SD	8	2.11 (1.59–2.36)	3.52 (2.76–4.16)
PR	41	1.77 (1.50–2.04)	3.39 (3.13–3.74)
CR	17	1.72 (1.32–2.18)	3.65 (2.95–4.10)
***P*-value**	–	0.562	0.880
Overall treatment response based on SVRR			
<50%	13	1.62 (1.38–2.11)	3.21 (3.05–3.56)
>50%	53	1.80 (1.48–2.16)	3.45 (3.13–4.04)
***P*-value**	–	0.514	0.306
Treatment response after induction chemotherapy			
PD	0	–	–
SD	6	2.11 (1.33–2.46)	3.52 (2.65–3.92)
PR	35	1.77 (1.45–2.01)	3.25 (3.10–3.78)
CR	12	1.72 (1.53–2.22)	3.89 (3.15–4.20)
***P*-value**	–	0.838	0.590
Treatment response after induction chemotherapy based on SVRR			
<50%	11	1.53 (1.33–2.01)	3.21 (3.01–3.57)
>50%	42	1.80 (1.49–2.17)	3.49 (3.13–4.04)
***P*-value**	–	0.380	0.369

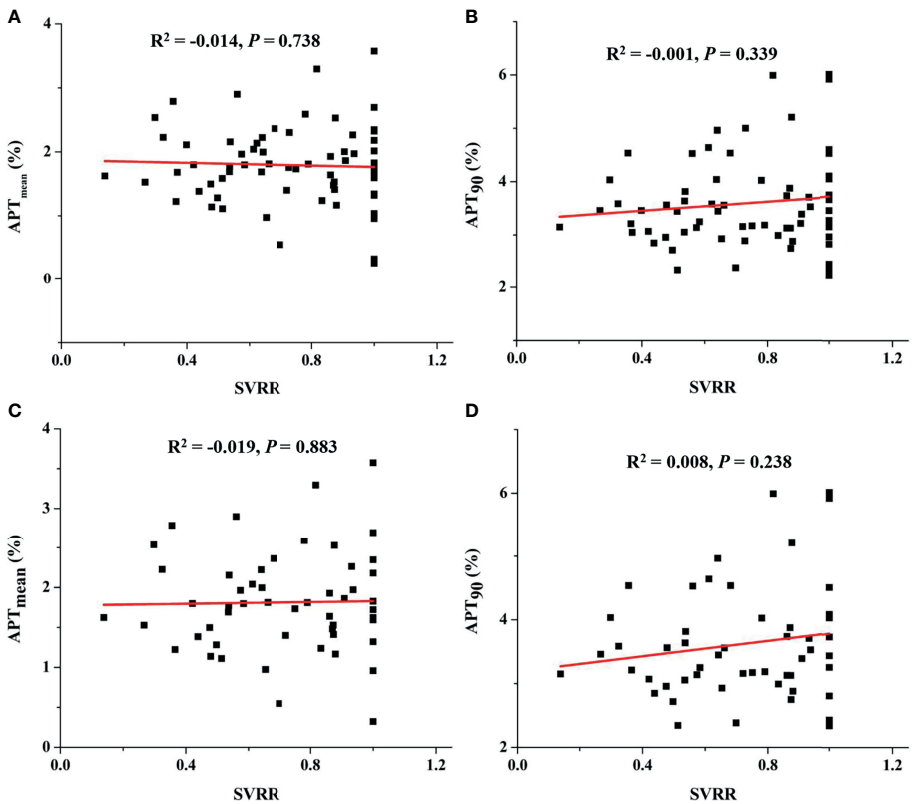


FIGURE 6 | For patients receiving overall treatment, correlation between APT_{mean} and SVRR **(A)**, between APT_{90} and SVRR **(B)**, and for patients who received induction chemotherapy, correlation between APT_{mean} and SVRR **(C)**, between APT_{90} and SVRR **(D)**.

Conclusion

In conclusion, NPC showed significantly stronger APT effect than normal nasopharyngeal tissue, facilitating NPC lesion detection and early identification. However, baseline APT values showed limited value in predicting short-term response to IC and no significant correlation with different histopathological subtypes, EBV infection status, and clinical stages.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

DL and YW contributed to project idea and supervision and to manuscript revision, and maintained integrity of this manuscript. ZL and LZ implemented the whole study, analyzed the data, and drafted the manuscript. QY and LQ contributed to statistical analysis and reviewed the manuscript. TL, HL, and CC collected the raw MRI data. YL collected the pathological data. PC and CQ collected clinical information. XL provided technical support. All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

The work described in this paper was supported by grants from the Key Areas Research and Development Program of Guangdong (2019B020235001) and the Shenzhen High-level Hospital Construction Fund.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2018) 68 (6):394–424. doi: 10.3322/caac.21492
2. Wei KR, Zheng RS, Zhang SW, Liang ZH, Li ZM, Chen WQ. Nasopharyngeal Carcinoma Incidence and Mortality in China, 2013. *Chin J Cancer* (2017) 36 (1):90. doi: 10.1186/s40880-017-0257-9
3. Blanchard P, Lee A, Marguet S, Leclercq J, Ng WT, Ma J, et al. Chemotherapy and Radiotherapy in Nasopharyngeal Carcinoma: An Update of the MAC-NPC Meta-Analysis. *Lancet Oncol* (2015) 16(6):645–55. doi: 10.1016/s1470-2045(15)70126-9
4. Cao SM, Yang Q, Guo L, Mai HQ, Mo HY, Cao KJ, et al. Neoadjuvant Chemotherapy Followed by Concurrent Chemoradiotherapy Versus Concurrent Chemoradiotherapy Alone in Locoregionally Advanced Nasopharyngeal Carcinoma: A Phase III Multicentre Randomised Controlled Trial. *Eur J Cancer* (2017) 75:14–23. doi: 10.1016/j.ejca.2016.12.039
5. Peng H, Chen L, Li WF, Guo R, Mao YP, Zhang Y, et al. Tumor Response to Neoadjuvant Chemotherapy Predicts Long-Term Survival Outcomes in Patients With Locoregionally Advanced Nasopharyngeal Carcinoma: A Secondary Analysis of a Randomized Phase 3 Clinical Trial. *Cancer* (2017) 123(9):1643–52. doi: 10.1002/cncr.30520
6. Peng H, Chen L, Zhang Y, Li WF, Mao YP, Liu X, et al. The Tumour Response to Induction Chemotherapy has Prognostic Value for Long-Term Survival Outcomes After Intensity-Modulated Radiation Therapy in Nasopharyngeal Carcinoma. *Sci Rep* (2016) 6:24835. doi: 10.1038/srep24835
7. Xu J, Zhang K, Wang H, Cui Z, Wu Y, Wang W, et al. Clinical Application of Diffusion-Weighted Magnetic Resonance Imaging in Radiotherapy for Nasopharyngeal Carcinoma. *J Xray Sci Technol* (2019) 27(2):297–306. doi: 10.3233/xst-180441
8. Xiao Y, Chen Y, Chen Y, He Z, Yao Y, Pan J. Longitudinal Assessment of Intravoxel Incoherent Motion Diffusion Weighted Imaging in Evaluating the Radio-Sensitivity of Nasopharyngeal Carcinoma Treated With Intensity-Modulated Radiation Therapy. *Cancer Res Treat* (2019) 51(1):345–56. doi: 10.4143/crt.2018.089
9. Huang WY, Li MM, Lin SM, Chen F, Yang K, Zhu XL, et al. *In Vivo* Imaging Markers for Prediction of Radiotherapy Response in Patients With Nasopharyngeal Carcinoma: RESOLVE DWI Versus DKI. *Sci Rep* (2018) 8 (1):15861. doi: 10.1038/s41598-018-34072-9
10. Qin Y, Yu X, Hou J, Hu Y, Li F, Wen L, et al. Prognostic Value of the Pretreatment Primary Lesion Quantitative Dynamic Contrast-Enhanced Magnetic Resonance Imaging for Nasopharyngeal Carcinoma. *Acad Radiol* (2019) 26(11):1473–82. doi: 10.1016/j.acra.2019.01.021
11. Zhou J, Heo HY, Knutsson L, van Zijl PCM, Jiang S. APT-Weighted MRI: Techniques, Current Neuro Applications, and Challenging Issues. *J Magn Reson Imaging* (2019) 50(2):347–64. doi: 10.1002/jmri.26645
12. Meissner JE, Korzowski A, Regnery S, Goerke S, Breitling J, Floca RO, et al. Early Response Assessment of Glioma Patients to Definitive Chemoradiotherapy Using Chemical Exchange Saturation Transfer Imaging at 7 T. *J Magn Reson Imaging* (2019) 50(4):1268–77. doi: 10.1002/jmri.26702
13. Paech D, Dreher C, Regnery S, Meissner JE, Goerke S, Windschuh J, et al. Relaxation-Compensated Amide Proton Transfer (APT) MRI Signal Intensity is Associated With Survival and Progression in High-Grade Glioma Patients. *Eur Radiol* (2019) 29(9):4957–67. doi: 10.1007/s00330-019-06066-2
14. Park JE, Kim HS, Park KJ, Choi CG, Kim SJ. Histogram Analysis of Amide Proton Transfer Imaging to Identify Contrast-Enhancing Low-Grade Brain Tumor That Mimics High-Grade Tumor: Increased Accuracy of MR Perfusion. *Radiology* (2015) 277(1):151–61. doi: 10.1148/radiol.2015142347
15. Togao O, Hiwatashi A, Yamashita K, Kikuchi K, Keupp J, Yoshimoto K, et al. Grading Diffuse Gliomas Without Intense Contrast Enhancement by Amide Proton Transfer MR Imaging: Comparisons With Diffusion- and Perfusion-Weighted Imaging. *Eur Radiol* (2017) 27(2):578–88. doi: 10.1007/s00330-016-4328-0
16. Choi YS, Ahn SS, Lee SK, Chang JH, Kang SG, Kim SH, et al. Amide Proton Transfer Imaging to Discriminate Between Low- and High-Grade Gliomas: Added Value to Apparent Diffusion Coefficient and Relative Cerebral Blood Volume. *Eur Radiol* (2017) 27(8):3181–9. doi: 10.1007/s00330-017-4732-0
17. Ma B, Blakeley JO, Hong X, Zhang H, Jiang S, Blair L, et al. Applying Amide Proton Transfer-Weighted MRI to Distinguish Pseudoprogression From True Progression in Malignant Gliomas. *J Magn Reson Imaging* (2016) 44(2):456–62. doi: 10.1002/jmri.25159

18. Park KJ, Kim HS, Park JE, Shim WH, Kim SJ, Smith SA. Added Value of Amide Proton Transfer Imaging to Conventional and Perfusion MR Imaging for Evaluating the Treatment Response of Newly Diagnosed Glioblastoma. *Eur Radiol* (2016) 26(12):4390–403. doi: 10.1007/s00330-016-4261-2
19. Donahue MJ, Donahue PC, Rane S, Thompson CR, Strother MK, Scott AO, et al. Assessment of Lymphatic Impairment and Interstitial Protein Accumulation in Patients With Breast Cancer Treatment-Related Lymphedema Using CEST MRI. *Magn Reson Med* (2016) 75(1):345–55. doi: 10.1002/mrm.25649
20. Takayama Y, Nishie A, Sugimoto M, Togao O, Asayama Y, Ishigami K, et al. Amide Proton Transfer (APT) Magnetic Resonance Imaging of Prostate Cancer: Comparison With Gleason Scores. *Magma* (2016) 29(4):671–9. doi: 10.1007/s10334-016-0537-4
21. Ohno Y, Yui M, Koyama H, Yoshikawa T, Seki S, Ueno Y, et al. Chemical Exchange Saturation Transfer MR Imaging: Preliminary Results for Differentiation of Malignant and Benign Thoracic Lesions. *Radiology* (2016) 279(2):578–89. doi: 10.1148/radiol.2015151161
22. Law BKH, King AD, Ai QY, Poon DMC, Chen W, Bhatia KS, et al. Head and Neck Tumors: Amide Proton Transfer MRI. *Radiology* (2018) 288(3):782–90. doi: 10.1148/radiol.2018171528
23. Yu L, Li C, Luo X, Zhou J, Zhang C, Zhang Y, et al. Differentiation of Malignant and Benign Head and Neck Tumors With Amide Proton Transfer-Weighted MR Imaging. *Mol Imaging Biol* (2019) 21(2):348–55. doi: 10.1007/s11307-018-1248-1
24. Liu R, Jiang G, Gao P, Li G, Nie L, Yan J, et al. Non-Invasive Amide Proton Transfer Imaging and ZOOM Diffusion-Weighted Imaging in Differentiating Benign and Malignant Thyroid Micronodules. *Front Endocrinol (Lausanne)* (2018) 9:747. doi: 10.3389/fendo.2018.00747
25. Bae YJ, Choi BS, Jeong WJ, Jung YH, Park JH, Sunwoo L, et al. Amide Proton Transfer-Weighted MRI in the Diagnosis of Major Salivary Gland Tumors. *Sci Rep* (2019) 9(1):8349. doi: 10.1038/s41598-019-44820-0
26. Qamar S, King AD, Ai QY, Law BKH, Chan JSM, Poon DMC, et al. Amide Proton Transfer MRI Detects Early Changes in Nasopharyngeal Carcinoma: Providing a Potential Imaging Marker for Treatment Response. *Eur Arch Otorhinolaryngol* (2019) 276(2):505–12. doi: 10.1007/s00405-018-5231-x
27. Qamar S, King AD, Ai QH, Mo FKF, Chen W, Poon DMC, et al. Pre-Treatment Amide Proton Transfer Imaging Predicts Treatment Outcome in Nasopharyngeal Carcinoma. *Eur Radiol* (2020) 30(11):6339–47. doi: 10.1007/s00330-020-06985-5
28. Lertbutsayanukul C, Kannarunimit D, Prayongrat A, Chakkabat C, Kitpanit S, Hansasuta P. Prognostic Value of Plasma EBV DNA for Nasopharyngeal Cancer Patients During Treatment With Intensity-Modulated Radiation Therapy and Concurrent Chemotherapy. *Radiol Oncol* (2018) 52(2):195–203. doi: 10.2478/raon-2018-0016
29. Wu SG, Lian CL, Wang J, Zhang WW, Sun JY, Lin Q, et al. The Effect of Histological Subtypes on Survival Outcome in Nasopharyngeal Carcinoma After Extensive Follow Up. *Ann Transl Med* (2019) 7(23):768. doi: 10.21037/atm.2019.11.75
30. Huang SH, O'Sullivan B. Overview of the 8th Edition TNM Classification for Head and Neck Cancer. *Curr Treat Options Oncol* (2017) 18(7):40. doi: 10.1007/s11864-017-0484-y
31. Stancanello J, Terreno E, Castelli DD, Cabella C, Uggeri F, Aime S. Development and Validation of a Smoothing-Splines-Based Correction Method for Improving the Analysis of CEST-MR Images. *Contrast Media Mol Imaging* (2008) 3(4):136–49. doi: 10.1002/cmmi.240
32. Kim M, Gillen J, Landman BA, Zhou J, van Zijl PC. Water Saturation Shift Referencing (WASSR) for Chemical Exchange Saturation Transfer (CEST) Experiments. *Magn Reson Med* (2009) 61(6):1441–50. doi: 10.1002/mrm.21873
33. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1). *Eur J Cancer* (2009) 45(2):228–47. doi: 10.1016/j.ejca.2008.10.026
34. Stelow EB, Wenig BM. Update From The 4th Edition of the World Health Organization Classification of Head and Neck Tumours: Nasopharynx. *Head Neck Pathol* (2017) 11(1):16–22. doi: 10.1007/s12105-017-0787-0
35. Yuan J, Chen S, King AD, Zhou J, Bhatia KS, Zhang Q, et al. Amide Proton Transfer-Weighted Imaging of the Head and Neck at 3 T: A Feasibility Study on Healthy Human Subjects and Patients With Head and Neck Cancer. *NMR BioMed* (2014) 27(10):1239–47. doi: 10.1002/nbm.3184
36. Chen Y, Li X, Song Y, Zhu X, Zhao J, Yan X, et al. The Diagnostic Efficacy of Amide Proton Transfer Imaging in Grading Gliomas and Predicting Tumor Proliferation. *Neuroreport* (2019) 30(2):139–44. doi: 10.1097/wnr.0000000000001174
37. Wang HY, Chang YL, To KF, Hwang JS, Mai HQ, Feng YF, et al. A New Prognostic Histopathologic Classification of Nasopharyngeal Carcinoma. *Chin J Cancer* (2016) 35:41. doi: 10.1186/s40880-016-0103-5
38. Ng RH, Ngan R, Wei WI, Gullane PJ, Phillips J. Trans-Oral Brush Biopsies and Quantitative PCR for EBV DNA Detection and Screening of Nasopharyngeal Carcinoma. *Otolaryngol Head Neck Surg* (2014) 150(4):602–9. doi: 10.1177/0194599813520136
39. Zhang Y, Tang LL, Li YQ, Liu X, Liu Q, Ma J. Spontaneous Remission of Residual Post-Therapy Plasma Epstein-Barr Virus DNA and Its Prognostic Implication in Nasopharyngeal Carcinoma: A Large-Scale, Big-Data Intelligence Platform-Based Analysis. *Int J Cancer* (2019) 144(9):2313–9. doi: 10.1002/ijc.32021
40. Wang FH, Wei XL, Feng J, Li Q, Xu N, Hu XC, et al. Efficacy, Safety, and Correlative Biomarkers of Toripalimab in Previously Treated Recurrent or Metastatic Nasopharyngeal Carcinoma: A Phase II Clinical Trial (POLARIS-02). *J Clin Oncol* (2021) 39(7):704–12. doi: 10.1200/jco.20.02712
41. Togao O, Yoshiura T, Keupp J, Hiwatashi A, Yamashita K, Kikuchi K, et al. Amide Proton Transfer Imaging of Adult Diffuse Gliomas: Correlation With Histopathological Grades. *Neuro Oncol* (2014) 16(3):441–8. doi: 10.1093/neuonc/not158
42. Zhou J, Hong X, Zhao X, Gao JH, Yuan J. APT-Weighted and NOE-Weighted Image Contrasts in Glioma With Different RF Saturation Powers Based on Magnetization Transfer Ratio Asymmetry Analyses. *Magn Reson Med* (2013) 70(2):320–7. doi: 10.1002/mrm.24784

Conflict of Interest: Author LQ was employed by GE Healthcare.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Liu, Zou, Yang, Qian, Li, Luo, Che, Lei, Chen, Qiu, Liu, Wu and Luo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 17 February 2022
doi: 10.3389/fonc.2022.794101

Establishment of a Novel Mouse Hepatocellular Carcinoma Model for Dynamic Monitoring of Tumor Development by Bioluminescence Imaging

Xiangyi Cao[1,2], Yulong Zhang[1], Qianqian Zhou[1], Sujing Sun[1], Minwei He[1], Xiaohui Wang[1], Ping Ma[1], Xiaoang Yang[2], Liping Lv[1] and Linsheng Zhan[1]**

[1] Department of Transfusion Medicine, Institute of Health Service and Transfusion Medicine, Beijing, China, [2] Zhengzhou University, BGI College and Henan Institute of Medical and Pharmaceutical Sciences, Academy of Medical Science, Zhengzhou University, Zhengzhou, China

OPEN ACCESS

Edited by:
Guang Yang, Imperial College London, United Kingdom

Reviewed by:
John A. Ronald, Western University, Canada
Yang Deng, Shandong First Medical University, China

***Correspondence:**
Linsheng Zhan
lszhan91@yahoo.com
Liping Lv
lvlp2020@163.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *13 October 2021*
Accepted: *21 January 2022*
Published: *17 February 2022*

Citation:
Cao X, Zhang Y, Zhou Q, Sun S, He M, Wang X, Ma P, Yang X, Lv L and Zhan L (2022) Establishment of a Novel Mouse Hepatocellular Carcinoma Model for Dynamic Monitoring of Tumor Development by Bioluminescence Imaging. Front. Oncol. 12:794101. doi: 10.3389/fonc.2022.794101

In this study, a novel mouse model of hepatocellular carcinoma (HCC) was established by simultaneously knocking out Pten and p53 suppressor genes and overexpressing c-Met and △90-β-catenin proto-oncogenes in the livers of mice *via* hydrodynamic injection (HDI). The mutations were introduced using the CRISPR/Cas9 and Sleeping Beauty transposon systems. In this way, a primary liver cancer model was established within six weeks. In addition, macrophages expressing arginase-1(Arg1) promoter coupled with firefly luciferase were engineered for bioluminescence imaging (BLI) of the tumor microenvironment. This novel, rapidly-generated model of primary hepatocellular carcinoma can be monitored noninvasively, which can facilitate not only applications of the model, but also the development of new drugs and treatment strategies of HCC.

Keywords: HCC, mouse hepatocellular carcinoma model, hydrodynamic injection, CRISPR/Cas9, noninvasive monitoring

INTRODUCTION

Hepatocellular carcinoma (HCC), a highly prevalent malignancy with over 800,000 cases diagnosed annually worldwide, is the second leading cause of cancer-related deaths (1, 2). Early diagnosis and therapy of HCC are important for improvement of the patients' prognosis. Currently, HCC is diagnosed by imaging (3) and measuring serum alpha-fetoprotein (AFP) levels (4). The traditional imaging techniques used in HCC early diagnosis are computed tomography (CT), positron emission computed tomography (PET) and magnetic resonance imaging (MRI). However, diagnostic has it's limitation in detecting earlier mini-size cancer (5, 6). Though serum tumor markers, such as AFP and tumor-associated antigens (TAAs), are engaging alternatives for monitoring and early diagnosis of hepatocellular carcinoma, their sensitivities and specificities remain disappointing. As a result, 40% of affected patients are already in the advanced stage of HCC at the time of diagnosis, making surgical treatments impossible. In addition, the first-line treatment for advanced HCC is sorafenib, which provides finite clinical benefit (7). And other first or second-

line therapeutic options, such as immune checkpoint inhibitors, have also shown efficacy against HCC recently (8). Thus, the study of early diagnosis and new treatment strategies of HCC remains an attractive issue. Corresponding small animal models play an important role in studies on hepatocellular carcinoma early diagnosis and therapy. Building animal models that are similar to human liver cancer occurrence and development can help us better study the mechanism of liver cancer, while realize early diagnosis and drug development.

Genetically engineered mouse (GEM) models of HCC have been developed based on oncogene knock-in or the knockout of tumor suppressor genes. However, the establishment of such models is an expensive and time-consuming process, and it takes almost three months to establish an HCC mouse model with this technique. Hydrodynamic injection (HDI) is a novel gene transfection technology that can be used to establish a mouse model of liver cancer in a fraction of the time taken by a GEM model. Studies have shown that directly overexpressing c-Met and △90-β-catenin or simultaneously knocking out Pten and p53 in the mouse liver *via* HDI initiates tumorigenesis (9, 10). Since the activation of oncogenes and inactivation of tumor suppressors often synergize during cancer development (11), a murine HCC model was constructed in our study by simultaneously expressing oncogenes and knocking out tumor suppressor genes in the liver using the HDI method.

In vivo bioluminescence imaging (BLI) is routinely used to detect gene enhancer/promoter function in tissues or cells through reporter gene constructs (12), and monitor tumor growth, bacterial and viral infection, gene expression and therapeutic responses in real time by fluorescent monitoring devices. Usually, the reporter gene constructs used for BLI is formed by cloning the enhancer/promoter sequence upstream of the luciferase cDNA. Luciferase activities are utilized to reflect the expression of the protein of interest in-time. In addition, during tumor development, immune cells infiltrate in and around tumor tissues through chemotaxis, of which the tumor-associated macrophages (TAMs) are in the majority. TAMs are immune regulatory cells formed by peripheral monocytes under the influence of the tumor microenvironment (13), which can secrete a variety of cytokines and play a key role in promoting cancer cell proliferation, invasion and metastasis (14). The TAMs are classified into the pro-inflammatory M1 macrophages and immunosuppressive M2 macrophages, and M2 polarization is seen during tumor development and characterized by the upregulation of arginase-1 (Arg1) (15). Therefore, combining luciferase expression with activation of the Arg1 promoter is an ideal tool to dynamic monitoring of tumor development.

In this study, a mouse liver cancer model was established quickly by simultaneously overexpressing proto-oncogenes and knocking out tumor suppressor genes through hydrodynamic gene transfection, and sensor macrophages were engineered that could express the luciferase reporter gene under the control of the Arg1 promoter. With the development of HCC, the transgenic macrophages injected in the mouse underwent M1/M2 polarization, followed by Arg1 expression upregulation in macrophages. Thus, the luciferase reporter gene mediated by Arg1 promoter herein was constructed to track M2 polarization activated and thereby the progression of HCC. Using this model, liver cancer development can be monitored based on serological changes and BLI of tumor-infiltrating M2 macrophages.

METHODS

Experimental Animals

SPF C57BL/6J mice were purchased from Beijing Vital River Laboratory Animal Technology Co., Ltd., weighing 18 ~ 20 g and male. The experimental Animal Center of the Academy of Military Medical Sciences provided the SPF level breeding environment, and animal experiments were approved by the Experimental Animal Ethics Committee of the Academy of Military Medical Sciences 2020-680.

Plasmids

LentiCRISPRv2 expression plasmid was stored in our lab. PCMV/SB expression plasmid, PT3-EF1a-c-Met expression plasmid, PT3-△90-β-catenin and pGL3-mArg1 promoter/enhancer -31/-3810 expression plasmid were purchased from Addgene Company in USA.

Construction of Primary Liver Cancer Model in C57BL/6J Mice

The injection volume of normal saline containing plasmids was about 10% of the body weight of each mouse (about 2 mL). The plasmids content was generally about 10 μg/mouse, and the injection time was generally about 3-5 s/time. The amount of plasmids was calculated according to the concentration of extracted plasmids. Plasmids pCMV/SB, PT3-EF1a-c-Met, PT3-△90-β-catenin, LentiCRISPR-sgPten and LentiCRISPR-sgp53 were prepared in the ratio of 1:2:2:2:2, which was dissolved in a proper amount of normal saline, mixed well, and used in an environment free of high temperatures and ultraviolet radiation.

Western Blot

Western blot was used to detect whether Pten (anti-Pten Cell Signaling, 9559, 1:1000) or p53(anti- p53 Cell Signal,2545, 1:200) was successfully knocked down in recombinant cells and whether c-Met (anti-c-Met Cell Signaling, 8198, 1:1000)or △90-β-catenin (anti-△90-β-catenin Abcam, ab196204, 1:1000)was overexpressed in liver tissue. Lysis buffer RIPA containing protease inhibitor was used for 20 min at 0°C and centrifugated at 12 000 RPM at 4°C for 20 min. Samples were heated at 100°C for 5 min before operation. 25 mg protein was electrophoretized in 10% SDS-PAGE gel and transferred to PVDF membrane. 5% skim milk was incubated overnight with the primary antibody at 4°C After being rinsed with TBS-Tween 20 (0.1%), they were incubated at room temperature for 1 hour with enzymed-labeled goat anti-rabbit or goat anti-rat antibody (1:2000). Finally, chemiluminescence HRP substrate was used to detect the binding of the antibody to the protein.

Immunohistochemistry (IHC)

Liver tissue was fixed overnight in 4% formalin, embedded in paraffin, sected at 4 μm, and stained with hematoxylin and eosin (H&E) for pathological examination. Oil red staining was used to observe lipids. Immunohistochemical main antibodies included anti-Ki67 (Abcam, ab16667, 1:200), anti-CK19 (Abcam, ab133496, 1:100), anti-p53 (Cell Signaling, 2545, 1:160), Anti-Pten (Cell Signaling, 2545, 1:100), Anti-Met (Cell Signaling, 8198, 1:100) and Anti-△90-β-catenin (anti-△90-β-catenin Abcam, ab196204, 1:1000).

Detection of TAAs in Liver Tissue

The mRNA expression levels of GNAS, PTCH, NPM1 and PAX5 in recombinant cells were detected by qRT-PCR. Total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the instructions. cDNA was then obtained by reverse transcription using Rever Tra AceR qPCR RT Master Mix. PCR amplification was performed using synthetic primers and SYBR-PCR kit (SYBR qPCR Mix).The reaction was carried out in eight consecutive tubes, which were incubated at 95°C for 1 min, 95°C for 15 s, 58°C for 30 s, and 72°C for 30 s. A total of 40 cycles were incubated at 60°C for 5 s and 95°C for 5 s. PCR was performed on a Bio-Rad CFX amplification device. GAPDH mice were used as internal control. Primer sequences of GNAS, PTCH, NPM1 and PAX5 are as follows:

NPM1	Forward	GTATCTGGACCTGTACCTGATG
NPM1	Reverse	GTGGAGTCGTAGCATATAGTCC
PAX5	Forward	CTGACATCTTCACCACCACGGAAC
PAX5	Reverse	GGTTGTGCTCGCCAAGTCTCG
GNAS	Forward	TGAATCTGAGTCTGATCACGAG
GNAS	Reverse	ATTTCGGTCTCGGATTCGATAT
PTCH1	Forward	TCCATCAGCAATGTCACCGC
PTCH1	Reverse	TCAGTACTGGAGGGAGAACGC

Serum Collection and ELISA

Serum was collected once a week. The whole blood of mice was collected 100 μL by orbital blood sampling method in 1.5 mL anticoagulant tubes, centrifuged at 2000 RPM for 10 min, and the serum was sucked out and stored at -20℃ for use. For ELISA, the collected serum was balanced to room temperature and diluted twice with the dilution of the sample in the kit, and 20× washing buffer was diluted 1:20 with distilled water. The concentrations of AFP, AST, ALT, GNAS, PTCH, NPM1 and PAX5 in serum were determined respectively. The specific operation steps complied with the instructions for ELISA kit in The Department of Education. Finally, the OD value of each well was measured at the wavelength of 450 nm with a microplate reader.

The BLI System Monitored the Migration and Distribution of Bone-Marrow Derived Macrophages in Mice

Macrophages were collected and counted, and 5×10^6 macrophages were mixed in 300 μL PBS, and injected into the hydrodynamic modeling mice for 6 weeks by tail vein, with 5 mice in each group. Migration of bone-marrow derived macrophages was monitored by Xenogen Spectrum imaging at 30 min, 24 h, and 48 h after macrophage infusion. To ensure the imaging quality, the abdomen of mice was depilated to minimize the loss of light signals. Before imaging, mice were intraperitoneally injected with D-fluorescein (150 mg/kg) and anesthetized with isoflurane for 3 min. Imaging was performed to monitor the migration and distribution of macrophages in mice. Vivisection was performed 48 h after macrophage infusion, and macrophages distribution were assayed by BLI in lymph nodes and tissues including inguinal lymph nodes (ILN), axillary lymph nodes (ALN), nodi lymph nodes (NLN), liver lymph nodes (LivLN), mesenteric lymph nodes (MLN), liver tissue, stomach tissue, small bowel tissue, heart tissue, lung tissue, kidney tissue, pancreas tissue. All data were analyzed and processed by *in vivo* imaging software (Xenogen), and standardized calculations were performed for all the data.

Construction of Arg1EP-Luciferase-GFP/Raw264.7 Engineered Cells

Raw264.7 cells were infected with Arg1EP-Luciferase-GFP lentivirus. Three days after infection, the cells were cultured in DMEM complete medium containing purinomycin 2 μg/mL and 10% FBS. Finally, cells that stably expressed Arg1EP-Luciferase-GFP fusion protein were screened.

In Vitro Detection of the Bioluminescence Activity of Arg1EP-Luciferase-GFP/Raw264.7 Macrophages Induced by Tumor Microenvironment

2×10^5 Arg1EP-Luciferase -GFP/Raw264.7 cells were inoculated in the upper chamber of a 24-well Transwell culture plate, and the lower layers were added with DMEM medium, DMEM medium containing 25 ng/μL IL-4 and 2×10^5 Hepa1-6 cells, respectively. After co-incubation of 24 h, 10 μL 15 mg/mL D-Luciferin was added to each well and gently shaken into the Caliper IVIS Lumina II imaging system to detect the bioluminescence activity of Arg1EP-Lucifere-GFP/Raw264.7 cell lines.

In Vitro Detection of Bioluminescence Activity of Arg1EP-Luciferase -GFP/Raw264.7 Co-Incubated With Hepa1-6 Cells

1×10^5 Arg1EP-Luciferase-GFP/Raw264.7 cells were inoculated in the upper chamber of a 24-well Transwell culture plate, and 2×10^5 Hepa1-6 cells were inoculated in the lower chamber and co-incubated for 6 h, 12 h, 24 h and 48 h for imaging. (Cells at each time point were set with three multiple wells) 10 uL D-Luciferin was added at the concentration of 15 mg/mL to each well. The bioluminescence activity of Arg1 EP-Luciferase-GFP/Raw264.7 cell line was detected by Caliper IVIS Lumina II imaging system.

The Migration and Distribution of Arg1EP-Luciferase-GFP/Raw264.7 in Mice Were Monitored by BLI System

Arg1EP-luciferase-GFP/Raw264.7 was collected and counted, and 5×10^6 Arg1EP-Luciferase-GFP/Raw264.7 was dissolved in

300 μL PBS and mixed. The specific operation was performed using the same method as for monitoring the distribution of macrophages from bone marrow *in vivo*.

Statistical Analysis

Statistical analysis was performed by GraphPad Prism Software®. Two way ANOVA test (Analysis of variance) was used to compare statistical differences among multiple groups. All the experiments were repeated at least three times. Differences among means were considered significant at $p < 0.05$.

RESULTS

A Liver Cancer Model Was Successfully Established by Hydrodynamic Gene Transfection

Single-guide (sg) RNA sequences of Pten and p53 were designed and cloned into the CRISPRv2 vector. The schematic diagram of CRISPR/Cas9 related plasmids is shown in **Figure 1A**. The lentiCRISPR-sgPten and lentiCRISPR-sgp53 germ plasmids were then transfected into the NIH 3T3 mouse embryonic fibroblasts, and the genomic DNA of the stably transfected clones was sequenced. After the plasmids were packed into lentiviruses and transduced into the target cells, gene knockout was confirmed at the predicted site before the proto-spacer adjacent motif (PAM) (**Figure 1D**). Compared to the control group, the expressions of Pten and p53 were significantly reduced in the knockout group (**Figure 1C**). In order to achieve the knockout of tumor suppressor genes as well as the long-term stable expressions of oncogenes, the Sleeping Beauty (SB) transposon system was amplified with c-Met and △90-β-catenin genes (**Figure 1B**). The transposon system consisted of a transposon expression vector and a plasmid with an SB transposon box, which contained a promoter, the target gene and a polyA sequence with inverted repeat (IR) sequences. The SB transposable enzyme recognized IRs in the transposable box and catalyzed their translocation from one genomic site to another. The transposon system was transfected into NIH 3T3 cells and the overexpression of c-Met and △90-β-catenin was confirmed by Western blot (**Figure 1C**).

The pCMV/SB, pT3-EF1a-c-Met, pT3-△90-β-catenin, lentiCRISPR-sgPten and lentiCRISPR-sgp53 plasmids were mixed in equal proportions and injected into the tail vein of mice under high pressure. The control groups were injected with pCMV/SB and pT3-EF1α. Hydrodynamic transfection is shown in **Figure 1E**. To further monitor cancer progression, random mice were euthanized every two weeks and their liver tissues were analyzed. In the sixth week, the livers of normal mice appeared healthy with a brown tinge and smooth surface, while those of the experimental group were significantly enlarged, looked dark red due to an excessive amount of blood, and had a filamentous surface with numerous white diffuse tumor masses (**Figure 1F**). Tumor continued to grow and the liver filled the entire abdomen by the eighth week, resulting in abdominal swelling. Hematoxylin and eosin (H&E) staining of the malignant liver tissues showed a three-fold to four-fold increase in the size of hepatocytes and obvious crowding of hepatocyte cords. In addition, the hepatocytes had multiple cytoplasmic vacuoles and vacuolar degeneration along with inclusion bodies in the nucleus. Some cells had enlarged or multiple nuclei, and a number of enlarged nuclei were irregular, which was a sign of obvious regeneration of liver cells. In contrast, the hepatocytes in the normal liver were aligned (**Figure 1F**). Oil red staining showed a significant increase in the number of lipid droplets in the hepatocytes and hepatic steatosis (**Figure 1G**). CK19 was commonly expressed in bile duct epithelial cells, and absent from the mature liver cells. Tumor proliferating antigen Ki67 was highly expressed in numerous human solid tumors, which was a marker of cell proliferation and was correlated with patient prognosis. Numerous CK19 and Ki67 positive oval cells were observed in the liver tumors, indicating active proliferation of tumor cells and a high degree of malignancy (**Figure 1G**).

Serological Detection of Primary Mouse Hepatocellular Carcinoma

There was no significant change in the survival rates in the first eight weeks after injection, but 50% HCC mice died after eight weeks, all mice with the HCC died at the third month (**Figure 2A**). The ratio of liver weight to body weight was also markedly higher in the tumorigenic group than in the control group (**Figure 2B**). The liver cancer model was verified six weeks later by analyzing serum AFP levels. AFP is synthesized in large amounts in fetal hepatocytes and undifferentiated liver cancer cells, and is therefore a specific marker of liver cancer. The AFP levels were significantly higher in the experimental versus the control group, indicating the success of liver cancer modeling (**Figure 2C**). In addition, the serum levels of alanine aminotransferase (ALT) and glutamic-oxaloacetic aminotransferase (AST) were increased significantly in the liver cancer group, suggesting considerable liver damage (**Figure 2D**).

The production of TAAs is a humoral immune response in HCC patients, autoantibodies to TAAs may be detected even in patients' serum at early-stage (16–18). Wang et al. identified antibodies specific to 11 TAAs, including TP53, MSH2, mouse G-protein α subunit (GNAS), Pten, mouse pile-protein 5 (PAX5), GNA11, mouse patching homologue (PTCH), IDH1, SRSF2, mouse nucleolar phosphate 1 (NPM1) and Survivin by screening the sera of a large number of liver cancer patients. In our study, it was found that the PTCH, PAX5, GNAS and NPM1 mRNAs were significantly upregulated in liver tumors compared to healthy liver tissues (**Figure 2E**). Moreover, the sera of tumor-bearing mice showed increased titers of antibodies specific to NPM1, GNAS, PAX5 and PTCH at weeks 2, 4 and 6. In contrast, no significant changes were observed in the control group (**Figure 2F**).

Detection of Proto-Oncogenes Overexpression and Tumor Suppressor Genes Knock Down in Liver Tumors

To find out whether liver tumorigenesis resulted from oncogene overexpression and loss of tumor suppressor genes, the expressions of corresponding proteins in liver tissues were

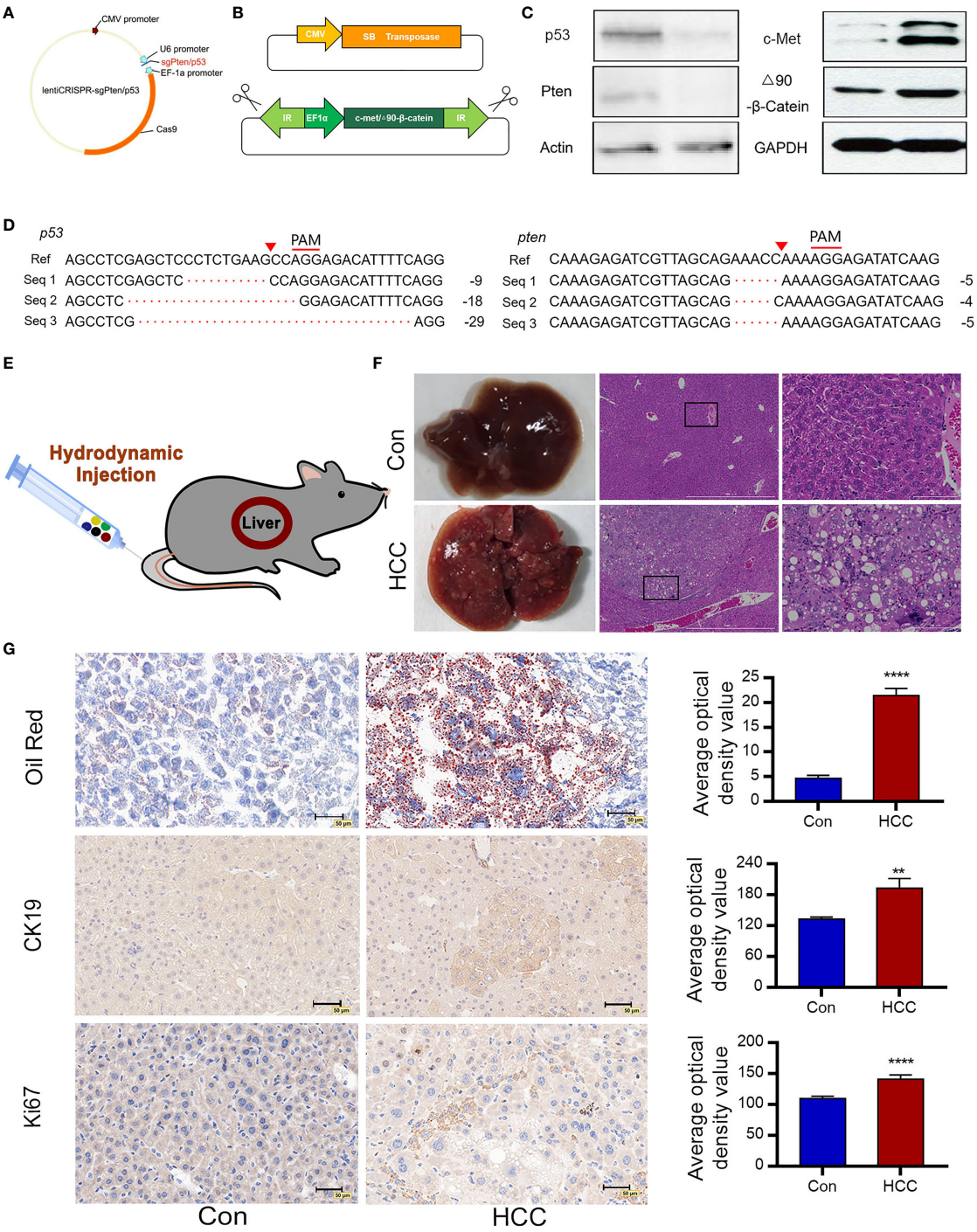


FIGURE 1 | Construction of plasmids and induction of liver tumors following hydrodynamic gene transfection. **(A)** Construction of lentiCRISPRv2/sgRNA plasmids expressing Cas9 and sgPten/p53. **(B)** Schematic diagram of SB transposase-associated plasmids. The SB transposase transfers a gene in the transposon cassette into a host genome. **(C)** Cell genomic DNA sequence of 3T3 cells analyzing deletion mutations of p53 and Pten. Seq 1-3 represent cell genomic DNA read sequences of NIH 3T3 cells knockdown by CRISPR/Cas9. The proto-spacer adjacent motif (PAM) sequences is denoted by red lines. Red triangles indicate the predicted DNA cleavage sites. **(D)** Immunoblots showing loss-of-function of p53 and Pten and overexpressions of c-Met and △90-β-catenin. Actin and GAPDH are used as control. **(E)** Saline containing five plasmids are injected into the tail vein of mice *via* HDI (n=10 for each group). **(F)** Representative macroscopic images of liver tumors observed in C57 mice 6 weeks post-injection and H&E-stained tumor tissue sections indicate large nuclei, multinuclear hepatocytes and reticulated cytoplasm of HCC cells (n=10). **(G)** Oil red O staining demonstrates hepatic steatosis. Ki67 and CK19 immunostaining represent proliferating cells. **P<0.01, ****P<0.0001.

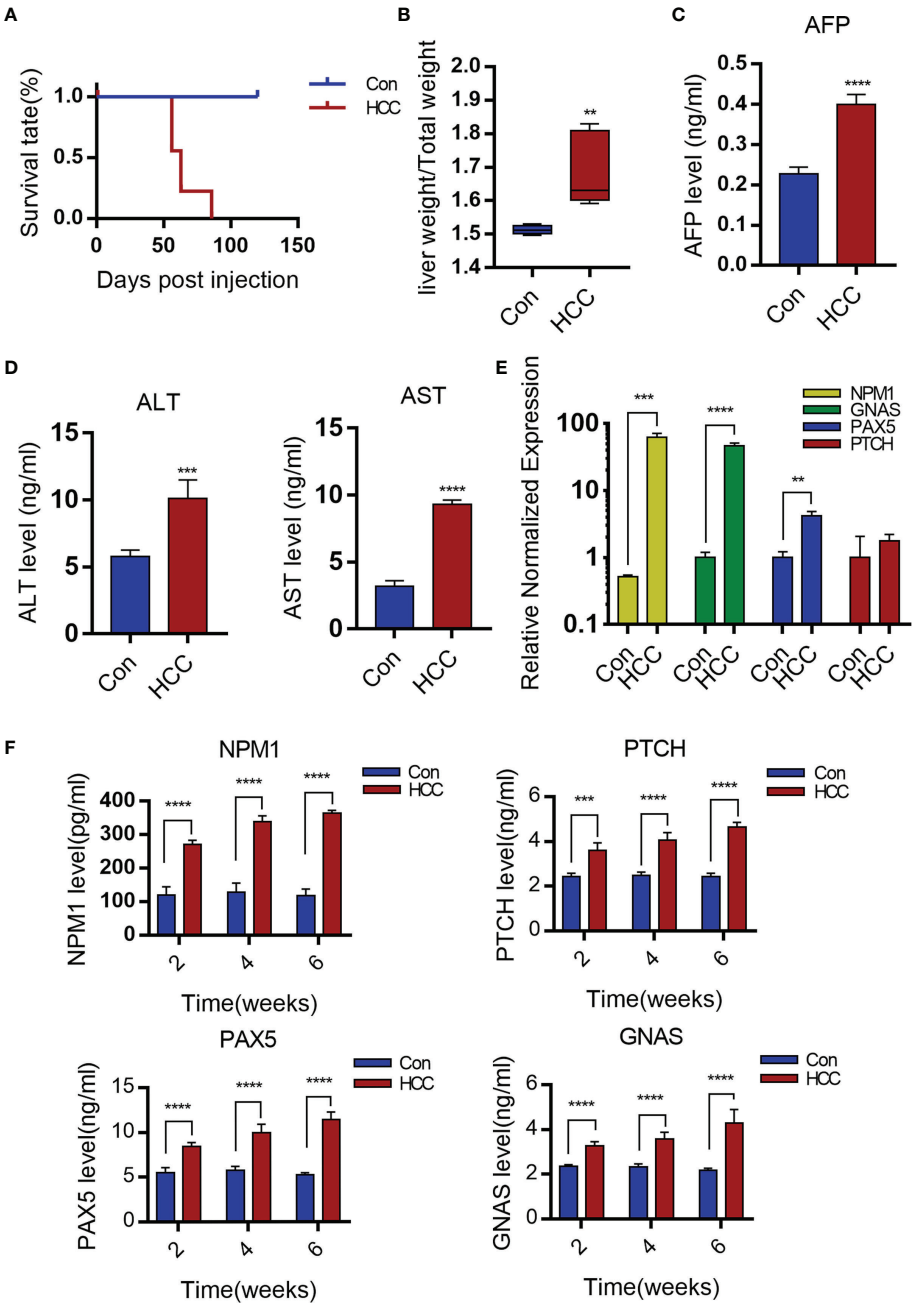


FIGURE 2 | Serological detection of primary mouse hepatocellular carcinoma. **(A)** The survival rate of mice after establishment of liver cancer model. **(B)** Liver weight relative to the whole-body weight of mice 6-week post injection. Data represent mean ± S.D. n=5. **(C, D)** AFP, ALT and AST levels of serum 6 weeks post- injection. Mean ± S.D; n=3, Statistical significance determined with t test; **P<0.05, ***P<0.001, ****P<0.0001; compared with the control. **(E)** Expression levels of GNAS, NPM1, PAX5 and PTCH genes in mouse liver tissue. Mean ± S.D; n=3, Statistical significance is determined with t test; ** P<0.01, ***P<0.001, ****P<0.0001; Compared with the control. **(F)** Serum levels of NPM1, GNAS, PAX5 and PTCH are time-dependence increases. Mean ± S.D; n=3, Statistical significance is determined with t test; ***P<0.001, ****P<0.0001.

analyzed. Immunohistochemistry (IHC) was used to observe the Sleeping Beauty translocation enzyme incorporated into mice liver cells that expressed corresponding proteins in the genome. Negative signals suggested suppressor genes knocked-down in the circled carcinoma area of the two upper panels. Besides, positive signals were observed in the cytoplasma of hepatoma cells, which showed c-Met and △90-β-catenin overexpressions in the circled carcinoma area of the two bottom panels (C-carcinoma, L-liver). In contrast, the control mice did not show any antibody positive staining (**Figure 3A**). On the basis of serial

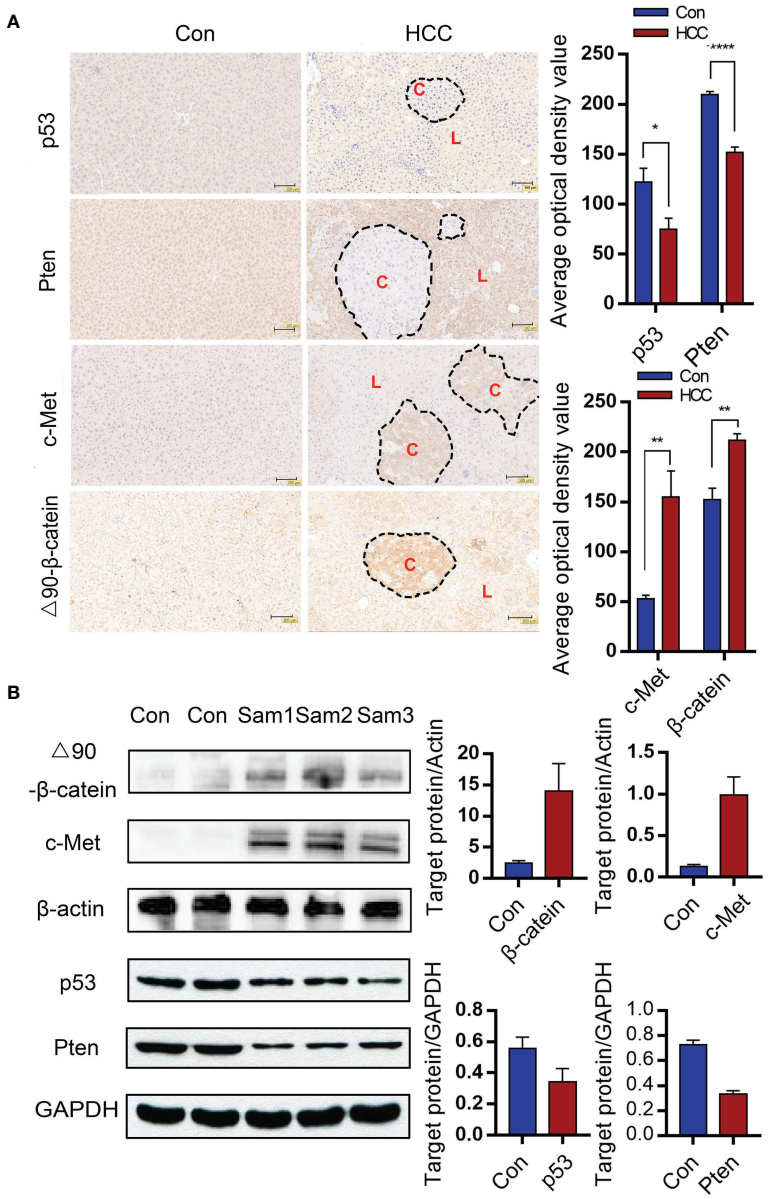


FIGURE 3 | Gene alterations in the liver tumor nodules. **(A)** Immunohistochemical (IHC) of liver tumor tissues indicating knockout of Pten and p53 and overexpressions of c-Met and △90-β-catenin. (C-carcinoma, L-liver) (Magnification: 100x) Mean ± S.D; n=3, Statistical significance determined with t test; *P<0.05, **P<0.01, ****P<0.0001. **(B)** Immunoblot showing overexpressions of c-Met and △90-β-catenin protein and decreased levels of p53 and Pten proteins in the mouse liver. *P<0.05, ** P<0.01, ****P<0.0001.

sections analysis, it was showed that some tumors stained positively for c-Met and β-catenin as well as negatively for Pten and p53 in similar locations, which suggested that part of the transformed hepatocytes had 4 intended changes (**Supplementary Figure 1**). The protein expressions of p53, Pten, c-Met and △90-β-catenin were verified by Western blot. Grayscale statistics were performed on the results as shown in **Figure 3B**. The deletion of endogenous p53 and Pten by the CRISPR system was confirmed by sequencing the DNA extracted from liver nodules 6 weeks after modeling. However, p53 and Pten genes were intact in the livers of control mice (**Supplementary Figure 2**).

Arg1 Was Highly Expressed in the Primary Liver Tumor Nodules

The infiltration of immune cells in liver tumors was analyzed by immunostaining the tissue sections with specific markers. The Kupffer cells were the most abundant, followed by T cells and dendritic cells (**Figure 4A** and **Supplementary Figure 3**). Arg1

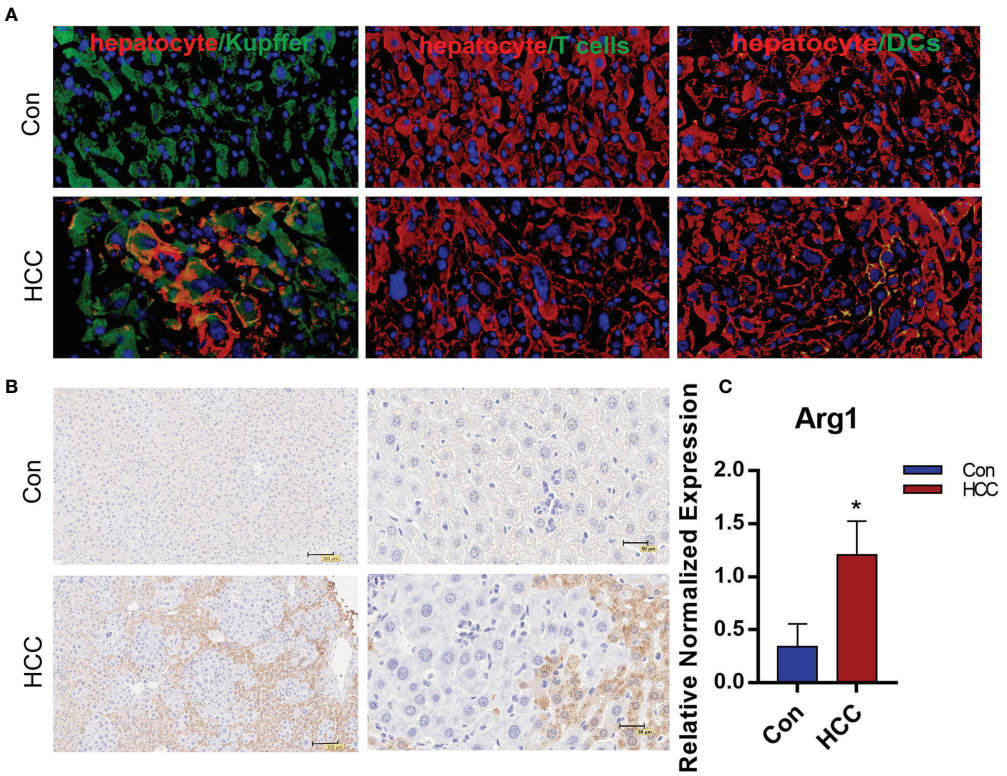


FIGURE 4 | The distribution of immune cells and expression of Arg1 in liver tumors. **(A)**. Distribution of macrophages, T cells and DCs in tumor tissues. The macrophages labeled with Alexa Fluor594, hepatocytes with FITC and the nuclei counterstained with DAPI in the first panel. T cells and DCs with FITC, hepatocytes with Alexa Fluor594, the nuclei counterstained with DAPI in the second and third panels. **(B)** Arg1 expression in the mouse liver. **(C)** Arg1 mRNA levels in Kupffer cells. Mean ± S.D; n=3, Statistical significance is determined with t test; *P<0.05.

was highly expressed around tumor tissues, but was hardly observed in healthy liver tissues (**Figure 4B**). The Kupffer cells were isolated by liver perfusion, which confirmed the significantly higher expression of Arg1 in tumor-bearing mice compared to the controls (**Figure 4C**). These results indicated that the Kupffer cells played an important role in the development of primary liver cancer. To add to our evidence, bone marrow-derived macrophages (BMDM) were adoptively transferred from Luc+ reporter mice into the control and tumor-bearing mice by tail vein infusion, and their distribution and migration were monitored by *in vivo* BLI. As shown in **Figure 5A**, intense bioluminescence was observed in the lungs within 30 min of infusion in both the control and tumor-bearing mice. After 24h, bioluminescence in the experimental mice was concentrated in the liver, whereas only small intestines showed weak bioluminescence in the control mice. The Luc+ macrophages remained in the liver of tumor-bearing mice after 48h, while those in the control group migrated to the spleen and small intestines (**Figure 5A**). *Ex vivo* BLI of the inguinal lymph nodes (ILN), axillary lymph nodes (ALN), submaxillary lymph nodes (NLN), hepatic lymph nodes (LivLN), mesenteric lymph nodes (MLN), and other organs or tissues after 48h also confirmed that the macrophages were mainly distributed in the livers of tumor-bearing mice, whereas those in the control group mainly homed to the secondary lymph nodes and the spleen (**Figures 5B, C**). Bioluminescence statistics of major organs are shown in **Figure 5D**, indicating that the bioluminescence intensity of the liver was the strongest. These findings suggested that the infiltration of macrophages into the liver tumor microenvironment led to their polarization to the Arg1-expressing M2 phenotype.

Macrophage Sensors Were Constructed to Monitor Liver Tumor Development *In Vitro*

To test the above hypothesis, Raw264.7 macrophages were transducted with Arg1EP-Luciferase-GFP lentivirus and incubated with Hepa1-6 cells. The bioluminescence intensity of Arg1EP-Luciferase-GFP/Raw264.7 cells also increased significantly when cultured in the presence of IL-4 compared to the untreated controls, while bioluminescence intensity of Arg1 macrophages co-incubated with Hepa1-6 was the strongest in the tumor microenvironment. (**Figure 6A** and **Supplementary Figure 4**). The bioluminescence intensity of the macrophages increased within 24h of co-culture and peaked at 48h (**Figure 6B**). Likewise, Arg1 mRNA levels increased significantly in the Arg1EP-Luciferase-GFP/Raw264.7 cells after 24h of co-culture with Hepa1-6 cells, as well as following IL-4 induction (0.1678 ± 0.02247), compared to the control group (P <0.05; **Figure 6C**).

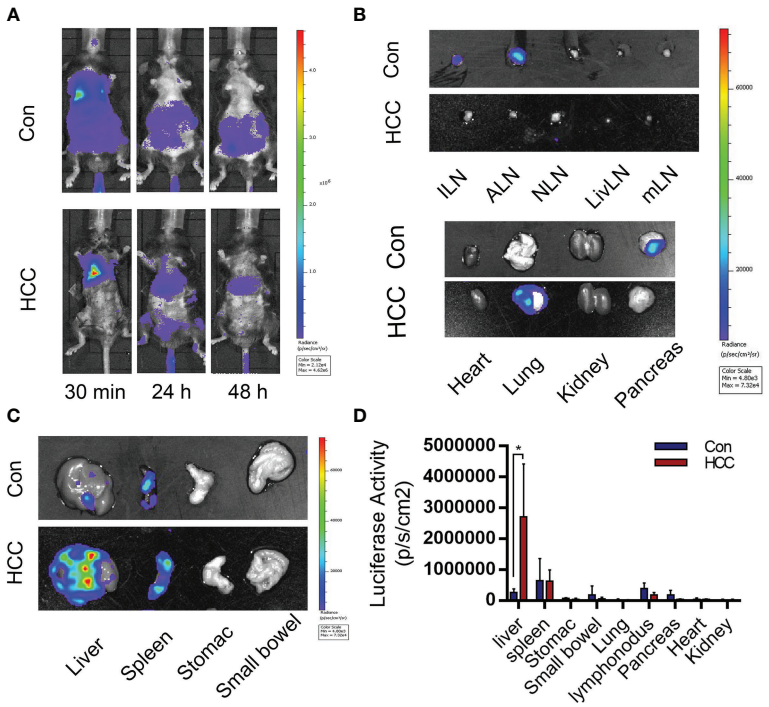


FIGURE 5 | Distribution of bone marrow-derived macrophages (BMDMs) in the control and tumor-bearing mouse model. **(A)** BMDMs' migration from 30 min to 48 h after cell injection. **(B, C)** BMDMs are mainly distributed in the tumor-bearing mice liver and spleen. **(D)** Statistical analysis of the luciferase activity in major organs. Mean ± S.D; n=3, Statistical significance is determined with t test; *P<0.05.

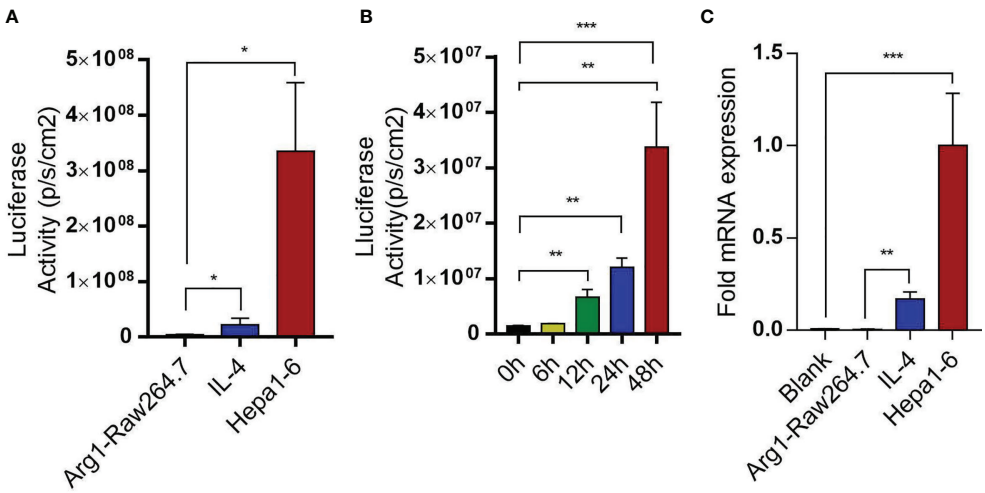


FIGURE 6 | Arg1EP-Luciferase-GFP/Raw264.7 macrophage sensors in tumor microenvironment. **(A)** Luciferase activity of macrophage sensors in different tumor microenvironments. **(B)** In the condition of co-incubation of Hepa1-6 showing increased luciferase activity of macrophage sensors within 48 h. **(C)** Changes in Arg1 mRNA levels after IL-4 inducein and in tumor microenvironment. Mean ± S.D; n=3, Statistical significance is determined with t test; *P<0.05, ***P<0.001.

Macrophage Sensors Were Used to Monitor Liver Tumor Development

On a prior basis, the macrophage sensor was used to monitor the hepatocellular carcinoma model. Arg1EP-Luciferase-GFP/Raw264.7 cells were injected into the tumor-bearing mice for 6 weeks through the tail vein. The macrophages were tracked by *in vivo* BLI and the intensity was used to assess the progression of tumors (**Figure 7A**). As shown in **Figure 7B**, the macrophages that

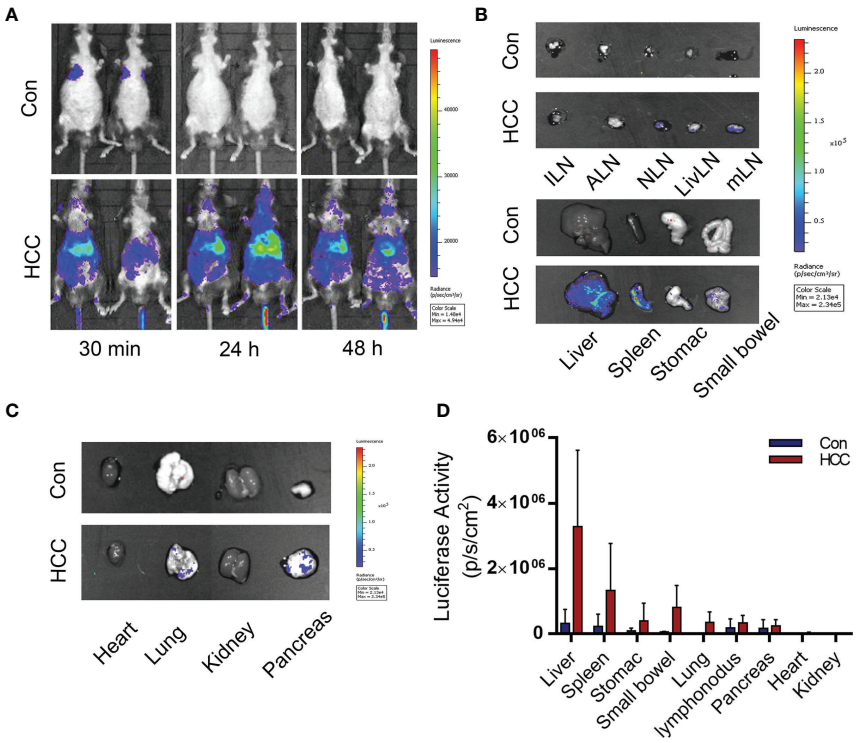


FIGURE 7 | Distribution of Arg1EP-Luciferase-GFP/Raw264.7 in the control and tumor-bearing mouse model. **(A)** Distribution of Arg1EP-Luciferase-GFP/Raw264.7 in mice from 30 min to 48 h point after cell injection. **(B, C)** Arg1EP-Luciferase-GFP/Raw264.7 mainly distributed in the liver and spleen of tumor-bearing mice. **(D)** Statistical analysis of the luciferase activity in major organs.

had been injected into the tumor-bearing mice initially migrated to the lungs and livers within 30 min before settled in the livers by 24 h, and remained in the livers even after 48 h. In contrast, the macrophages in the control group were accumulated in the lungs within 30 min, but were not detected after 24 h. The mice were euthanized 48 h after injection, and their tissues (see previous section) were imaged. As shown in **Figures 7B, C**, the bioluminescence signals were concentrated in the livers and spleens of the tumor-bearing mice, but no signals were detected in any of the tissues of control mice. The quantitative results of light signal in major organs showed that HCC significantly enhanced the M2 polarization of macrophages, with the strongest signal detected in the liver, indicating the macrophage sensor can indeed reflect the concurrence of liver cancer. In addition, a moderate signal intensity could also be detected in the spleen of HCC mice, though dramatically lower than that in the liver. We speculate that those extrahepatic macrophages may be polarized in the liver and then migrated to other organs, such as the spleen. Secondly, HCC may also elicit systemic changes of internal environment, leading to M2 polarization of macrophage sensors outside the liver and thus being monitored. Anyway, the highest signal in the liver suggests that the macrophage sensors are moderately specific to follow HCC progression (**Figure 7D**).To summarize, sensors of M2 polarization of macrophage sensors can be used to track tumor growth, as well as the formation of the tumor microenvironment.

DISCUSSION

A primary liver cancer model was established in mice by simultaneously knocking out Pten and p53, and overexpressing c-Met and △90-β-catenin by hydrodynamic gene transfection. This model was able to simulate the initiation and development of human primary liver cancer, but was technically unchallenging. In addition, this novel liver cancer model could be used as a promising tool for screening diagnostic markers and therapeutic targets. Tward et al. established a liver cancer model by overexpressing both c-Met and △90-β-catenin in the Sleeping Beauty transposase system, and detected tumor formation after 3 months (19). Likewise, Xue et al. induced HCC in FVB mice after 3 months by knocking out Pten and p53 using the CRISPR/Cas9 system (20). To the best of our knowledge, this was the first time that a primary mouse liver cancer model had been established by simultaneously overexpressing c-Met and △90-β-catenin, and knocking out Pten and p53, which could observe extensive tumor growth within 6 weeks. C-Met is a tyrosine kinase that is widely expressed in HCC cell lines and activates △90-β-catenin, which is frequently mutated in various cancers. Mutations and aberrant overexpression of △90-β-catenin is closely related to tumor development and prognosis (21). In addition, c-Met and the mutated Δ90-β-catenin synergistically promote HCC (22).

The tumor suppressor Pten inhibits the PI3K/AKT pathway, which regulates cell proliferation, apoptosis, protein synthesis and glucose metabolism, and promotes tumor progression (23). Pten is mutated or deleted in multiple cancers, and loss of Pten in hepatocytes leads to lipid accumulation (24). The tumor suppressor p53 plays a key role in inducing cellular senescence and death (25). Deletion of p53 is the most common genetic aberration in liver cancer, and promotes HCC tumorigenicity and lung metastasis *via* mTOR/Pten/Akt pathway activation. However, neither sgPten nor sgp53 can induce tumor formation alone, whereas simultaneous mutation of both can significantly accelerate liver cancer progression in mice (10).

Liver carcinogenesis was also confirmed in terms of pathological and serological indices. The hydrodynamic gene transfection technology simulates the spontaneity of primary liver cancer, which accelerates tumor induction. Previous studies have shown that early-stage liver cancer is characterized by the appearance of anti-TAA antibodies in the serum, which are potentially useful biomarkers for early diagnosis. Consistent with this finding, we detected a time-dependent increase in the serum levels of NPM1, GNAS, PAX5 and PTCH in tumor-bearing mice. Therefore, it is possible to determine the course of cancer development according to the changes of these serum biomarkers.

Early imaging diagnosis has limited effectiveness. For instance, CT is relatively expensive and is of limited diagnostic value, whereas PET can detect nothing more than tumors larger than 200 mm^3 and measuring more than 7 mm in diameter. In addition, MRI is time consuming, has low spatial resolution, and is therefore not effective for early tumor detection. BLI is a versatile and sensitive technique for longitudinal monitoring of cancer progression using cells that express luciferase reporter genes. Gambhir et al. used an arginase promoter to drive reporter gene expressions in colorectal and breast mouse tumor models, and the engineered macrophages after polarization could be detected by BLI and luciferase measured in the blood (26). Our study built a new approach to monitor hepatocellular carcinoma models using an Arg1 luciferase reporter plasmid. It is valuable to further studies regarding new drugs and treatment strategies of HCC, especially in evaluating the efficacy of immunotherapy in treating HCC.

The primary mouse liver cancer model established *via* hydrodynamic gene transfection technology is simple to operate, inexpensive and can better fit the occurrence and development of human liver cancer. It can not only provide a reliable animal model for the screening of markers for the early detection of HCC, but also facilitate the research on the mechanism of HCC and the development of anti-tumor drugs. In the early stage of tumor formation, it is impossible to intuitively observe the tumorigenesis of a mouse liver. By combining this imaging approach with engineered immune cell, the formation of the tumor microenvironment of a mouse liver can be determined in the early stage of tumor formation. As a result, the damage caused by vivisection can be prevented and the mouse liver cancer model used more efficiently, which will contribute to drug research and development while facilitating the treatment of primary liver cancer.

CONCLUSION

In our study, a primary mouse liver cancer model was established within six weeks by overexpressing the proto-oncogenes c-Met and △90-β-catenin, and knocking out the tumor suppressors Pten and p53 by HDI. Antibodies specific to TAAs such as PTCH, PAX5, GNAS and NPM1 increased in the mouse liver and serum in a time-dependent manner, and therefore can be used as early immunodiagnostic markers of HCC. A reporter plasmid was also constructed to track the Arg1-overexpressing M2 macrophages during liver carcinogenesis through BLI.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The animal study was reviewed and approved by Experimental Animal Ethics Committee of the Academy of Military Medical Sciences 2020-680.

AUTHOR CONTRIBUTIONS

LL and LZ contributed to conception and design of the study. YZ, QZ, SS and MH organized the database. XY, XW, and PM performed the statistical analysis. XC wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Grant 2018ZX10302205 from the National Science and Technology Major Projects of China.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.794101/full#supplementary-material

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2016. *CA Cancer J Clin* (2016) 66(1):7–30. doi: 10.3322/caac.21332
2. Yu SJ. A Concise Review of Updated Guidelines Regarding the Management of Hepatocellular Carcinoma Around the World: 2010-2016. *Clin Mol Hepatol* (2016) 22(1):7–17. doi: 10.3350/cmh.2016.22.1.7
3. Salvatore V, Gianstefani A, Negrini G, Allegretti G, Galassi M, Piscaglia F. Imaging Diagnosis of Hepatocellular Carcinoma: Recent Advances of Contrast-Enhanced Ultrasonography With SonoVue®. *Liver Cancer* (2016) 5(1):55–66. doi: 10.1159/000367748
4. Charrière B, Maulat C, Suc B, Muscari F. Contribution of Alpha-Fetoprotein in Liver Transplantation for Hepatocellular Carcinoma. *World J Hepatol* (2016) 8(21):881–90. doi: 10.4254/wjh.v8.i21.881
5. Miller BJ. Use of Imaging Prior to Referral to a Musculoskeletal Oncologist. *J Am Acad Orthop Surg* (2019) 27(22):e1001–8. doi: 10.5435/jaaos-d-19-00109
6. Hochman MG, Wu JS. MR Imaging of Common Soft Tissue Masses in the Foot and Ankle. *Magn Reson Imaging Clin N Am* (2017) 25(1):159–81. doi: 10.1016/j.mric.2016.08.013
7. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al. Sorafenib in Advanced Hepatocellular Carcinoma. *N Engl J Med* (2008) 359(4):378–90. doi: 10.1056/NEJMoa0708857
8. Wong JSL, Kwok GGW, Tang V, Li BCW, Leung R, Chiu J, et al. Pilimumab and Nivolumab/Pembrolizumab in Advanced Hepatocellular Carcinoma Refractory to Prior Immune Checkpoint Inhibitors. *J Immunother Cancer* (2021) 9:e001945. doi: 10.1136/jitc-2020-001945
9. Tao J, Xu E, Zhao Y, Singh S, Li X, Couchy G, et al. Modeling a Human Hepatocellular Carcinoma Subset in Mice Through Coexpression of Met and Point-Mutant β-Catenin. *Hepatology* (2016) 64(5):1587–605. doi: 10.1002/hep.28601
10. Liu Y, Qi X, Zeng Z, Wang L, Wang J, Zhang T, et al. CRISPR/Cas9-Mediated P53 and Pten Dual Mutation Accelerates Hepatocarcinogenesis in Adult Hepatitis B Virus Transgenic Mice. *Sci Rep* (2017) 7(1):2796. doi: 10.1038/s41598-017-03070-8
11. Plass C. Editorial: The International Journal of Cancer - Quo Vadis? *Int J Cancer* (2019) 145(12):3193. doi: 10.1002/ijc.32604
12. Mezzanotte L, van 't Root M, Karatas H, Goun EA, Löwik CWGM. *In Vivo* Molecular Bioluminescence Imaging: New Tools and Applications. *Trends Biotechnol* (2017) 35(7):640–52. doi: 10.1016/j.tibtech
13. Murdoch C, Giannoudis A, Lewis CE. Mechanisms Regulating the Recruitment of Macrophages Into Hypoxic Areas of Tumors and Other Ischemic Tissues. *Blood* (2004) 104(8):2224–34. doi: 10.1182/blood-2004-03-1109
14. Jayasingam SD, Citartan M, Thang TH, Mat Zin AA, Ang KC, Ch'ng ES. Evaluating the Polarization of Tumor-Associated Macrophages Into M1 and M2 Phenotypes in Human Cancer Tissue: Technicalities and Challenges in Routine Clinical Practice. *Front Oncol* (2020) 9. doi: 10.3389/fonc
15. Colegio OR, Chu N-Q, Szabo AL, Chu T, Rhebergen AM, Jairam V. Functional Polarization of Tumour-Associated Macrophages by Tumour-Derived Lactic Acid. *Nature* (2014) 513(7519):559–63. doi: 10.1038/nature13490
16. Macdonald IK, Parsy-Kowalska CB, Chapman CJ. Autoantibodies: Opportunities for Early Cancer Detection. *Trends Cancer* (2017) 3(3):198–213. doi: 10.1016/j.trecan.2017.02.003
17. Zhu Q, Liu M, Dai L, Ying X, Ye H, Zhou Y, et al. Using Immunoproteomics to Identify Tumor-Associated Antigens (TAAs) as Biomarkers in Cancer Immunodiagnosis. *Autoimmun Rev* (2013) 12(12):1123–8. doi: 10.1016/j.autrev.2013.06.015
18. Paradis V, Bedossa P. In the New Area of Noninvasive Markers of Hepatocellular Carcinoma. *J Hepatol* (2007) 46(1):9–11. doi: 10.1016/j.jhep.2006.10.006
19. Tward AD, Jones KD, Yant S, Cheung ST, Fan ST, Chen X, et al. Distinct Pathways of Genomic Progression to Benign and Malignant Tumors of the Liver. *Proc Natl Acad Sci USA* (2007) 104(37):14771–6. doi: 10.1073/pnas.0706578104
20. Xue W, Chen S, Yin H, Tammela T, Papagiannakopoulos T, Joshi NS, et al. CRISPR-Mediated Direct Mutation of Cancer Genes in the Mouse Liver. *Nature* (2014) 514(7522):380–4. doi: 10.1038/nature13589
21. Zeng G, Apte U, Micsenyi A, Bell A, Monga SP. Tyrosine Residues 654 and 670 in Beta-Catenin Are Crucial in Regulation of Met-Beta-Catenin Interactions. *Exp Cell Res* (2006) 312(18):3620–30. doi: 10.1016/j.yexcr.2006.08.003
22. Monga SP, Mars WM, Pediaditakis P, Bell A, Mulé K, Bowen WC, et al. Hepatocyte Growth Factor Induces Wnt-Independent Nuclear Translocation of Beta-Catenin After Met-Beta-Catenin Dissociation in Hepatocytes. *Cancer Res* (2002) 62(7):2064–71.
23. Vivanco I, Sawyers CL. The Phosphatidylinositol 3-Kinase AKT Pathway in Human Cancer. *Nat Rev Cancer* (2002) 2(7):489–501. doi: 10.1038/nrc839
24. Di Cristofano A, Pandolfi PP. The Multiple Roles of PTEN in Tumor Suppression. *Cell* (2000) 100(4):387–90. doi: 10.1016/s0092-8674(00)80674-1
25. Bykov VJN, Eriksson SE, Bianchi J, Wiman KG. Targeting Mutant P53 for Efficient Cancer Therapy. *Nat Rev Cancer* (2018) 18(2):89–102. doi: 10.1038/nrc.2017.109
26. Aalipour A, Chuang HY, Murty S, D'Souza AL, Park SM, Gulati GS, et al. Engineered Immune Cells as Highly Sensitive Cancer Diagnostics. *Nat Biotechnol* (2019) 37(5):531–9. doi: 10.1038/s41587-019-0064-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Cao, Zhang, Zhou, Sun, He, Wang, Ma, Yang, Lv and Zhan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

METHODS
published: 24 February 2022
doi: 10.3389/fonc.2022.742701

CHAIMELEON Project: Creation of a Pan-European Repository of Health Imaging Data for the Development of AI-Powered Cancer Management Tools

Luis Martí Bonmatí[1], Ana Miguel[1], Amelia Suárez[2], Mario Aznar[2], Jean Paul Beregi[3], Laure Fournier[3], Emanuele Neri[4], Andrea Laghi[5], Manuela França[6], Francesco Sardanelli[7], Tobias Penzkofer[8], Phillipe Lambin[9], Ignacio Blanquer[10], Marion I. Menzel[11,12], Karine Seymour[13], Sergio Figueiras[14], Katharina Krischak[15], Ricard Martínez[16], Yisroel Mirsky[17], Guang Yang[18†] and Ángel Alberich-Bayarri[19†]*

[1] Medical Imaging Department, La Fe University and Polytechnic Hospital & Biomedical Imaging Research Group Grupo de Investigación Biomédica en Imagen (GIBI2[30]) at La Fe University and Polytechnic Hospital and Health Research Institute, Valencia, Spain, [2] Matical Innovation SL, Madrid, Spain, [3] Collège des enseignants en radiologie de France, Paris, France, [4] Diagnostic Radiology 3, Department of Translational Research, University of Pisa, Pisa, Italy, [5] Medicina Traslazionale e Oncologia, Sant Andrea Sapienza Rome, Rome, Italy, [6] Department of Radiology, Centro Hospitalar Universitário do Porto, Porto, Portugal, [7] Servizio di Diagnostica per Immagini, "Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) Policlinico San Donato, Milanese, Italy, [8] Department of Radiology, CHARITÉ-Universitätsmedizin Berlin, Berlin, Germany, [9] Department of Precision Medicine, Maastricht University, Maastricht, Netherlands, [10] Computing Science Department, Universitat Politècnica de València, València, Spain, [11] GE Healthcare, München, Germany, [12] Department of Physics, Technical University of Munich, Garching, Germany, [13] Medexprim, Labège, France, [14] Bahia Software S.L.U., Coruña, Spain, [15] European Institute for Biomedical Imaging Research, EIBIR gemeinnützige GmbH, Vienna, Austria, [16] Departamento de Derecho Constitucional, Ciencia Política y Administración, Universitat de València, València, Spain, [17] Software and Information Systems Engineering, Ben-Gurion University of the Negev, Beer Sheva, Israel, [18] National Heart and Lung Institute, Imperial College London, London, United Kingdom, [19] Quantitative Imaging Biomarkers in Medicine, QUIBIM SL, Valencia, Spain

The CHAIMELEON project aims to set up a pan-European repository of health imaging data, tools and methodologies, with the ambition to set a standard and provide resources for future AI experimentation for cancer management. The project is a 4 year long, EU-funded project tackling some of the most ambitious research in the fields of biomedical imaging, artificial intelligence and cancer treatment, addressing the four types of cancer that currently have the highest prevalence worldwide: lung, breast, prostate and colorectal. To allow this, clinical partners and external collaborators will populate the repository with multimodality (MR, CT, PET/CT) imaging and related clinical data. Subsequently, AI developers will enable a multimodal analytical data engine facilitating the interpretation, extraction and exploitation of the information stored at the repository. The development and implementation of AI-powered pipelines will enable advancement towards automating data deidentification, curation, annotation, integrity securing and image harmonization. By the end of the project, the usability and performance of the repository as a tool fostering AI experimentation will be technically validated, including a validation subphase by world-class European AI developers, participating in Open Challenges to the AI Community. Upon successful validation of the repository, a set of

OPEN ACCESS

Edited by:
Bo Gao, Affiliated Hospital of Guizhou Medical University, China

Reviewed by:
Jessica Mary Winfield, Royal Marsden NHS Foundation Trust, United Kingdom
Karim Lekadir, University of Barcelona, Spain

***Correspondence:**
Luis Martí Bonmatí
marti_lui@gva.es

[†]These authors have contributed equally to this work and share last authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *16 July 2021*
Accepted: *28 January 2022*
Published: *24 February 2022*

Citation:
Bonmatí LM, Miguel A, Suárez A, Aznar M, Beregi JP, Fournier L, Neri E, Laghi A, França M, Sardanelli F, Penzkofer T, Lambin P, Blanquer I, Menzel MI, Seymour K, Figueiras S, Krischak K, Martínez R, Mirsky Y, Yang G and Alberich-Bayarri Á (2022) CHAIMELEON Project: Creation of a Pan-European Repository of Health Imaging Data for the Development of AI-Powered Cancer Management Tools. Front. Oncol. 12:742701. doi: 10.3389/fonc.2022.742701

selected AI tools will undergo early *in-silico* validation in observational clinical studies coordinated by leading experts in the partner hospitals. Tool performance will be assessed, including external independent validation on hallmark clinical decisions in response to some of the currently most important clinical end points in cancer. The project brings together a consortium of 18 European partners including hospitals, universities, R&D centers and private research companies, constituting an ecosystem of infrastructures, biobanks, AI*/in-silico* experimentation and cloud computing technologies in oncology.

Keywords: radiology, artificial intelligence-AI, cancer imaging, cancer management, quantitative imaging biomarkers, image harmonization

INTRODUCTION

The use of Artificial Intelligence (AI) on health data is generating promising tools to assist clinicians in cancer management, as an increasing number of health imaging-based AI approaches are proving to have vast potential to become useful clinical tools in different areas of application (1). These include recurrence and survival prediction using multidimensional heterogeneous data (2) prediction of tumor molecular features and association with tumor spread (3, 4), stratification of patients based on risk (5), and prediction of treatment response (6) among many others.

Despite these major advancements, the development of imaging-based AI tools relies on the availability of large, quality-controlled datasets (7), which currently still remains a major challenge. The generation of these imaging biobanks is a resource-intensive endeavor, facing multiple technical and operational difficulties such as image and data harmonization, data curation and annotation, image pre-processing and annotation, as well as various legal and ethical restrictions (8–13). As a result, the quantity, quality, and representativeness of datasets still remain major limiting factors in the development of predictive cancer management tools.

Despite these limitations, several health imaging repositories have been created to date (14–16), such as the Cancer Imaging Archive (TCIA) which is one of the most renowned amongst those focusing on cancer imaging (17). Albeit of huge potential, the vast majority of these repositories have been created as stand-alone entities, being currently not in a position to become interoperable with similar existing initiatives. As such, the need for the creation of a fully FAIR (Findable, Accessible, Interoperable, Reusable), GDPR compliant, European imaging repository still stands (18).

To address the lack of data availability as well as the interoperability limitation of currently existing initiatives, the CHAIMELEON project aims to set-up and populate a cancer imaging repository facilitating access to large, high-quality sets of anonymized data. This will be achieved through the creation of a distributed data repository that will be made interoperable with other existing repositories and biobanks, enabling secure share and reuse of data as an intuitive sustainable single-access point resource for the community of developers working on AI-powered cancer management solutions. The repository will use a controlled access policy, whereby registered users will have access to datasets upon acceptance of the terms and conditions of use. These conditions, albeit still under definition by the governance bodies, will include the contracting of non-identification commitments as well as others related to the purpose of use of the data. Since imaging datasets contain images acquired at different centers with different scanners (cross-vendor/cross-institution datasets), quantitative image features, parameters, values and ranges extracted from images acquired at one center may not be reproducible from once center to another. This is due to a lack of consistency of medical images, as they generate from different equipment vendors, models and software versions. To ensure the reproducibility of quantitative imaging biomarkers (QIB) and allow scientific reuse of retrospective imaging data from multicenter acquisitions, CHAIMELEON has set the development and testing of imaging data harmonization protocols as one of its main objectives. Different harmonization approaches have been proposed, including a disruptive one for the generation of synthetic images adjusted to a common harmonization framework, ensuring that the authenticity and integrity of each synthetic coherent image is properly secured.

The project will involve the setup of the IT infrastructure, the creation of protocols for legal compliance, the development of tools for agile data ingestion and curation and processing pipelines for imaging data annotation and harmonization, as well as methodologies and tools for enhanced interpretability of AI models among others. Since the repository will be targeted to AI developers as end-users, it will not be designed as a simple data warehouse but as a complete AI-powered solution that will provide integrated quality data. To do so, AI-powered pipelines for data annotation, multicentric data harmonization, integrity securing of synthetic AI-generated images and clinical decision prediction will be implemented, aiming to enhance the interpretability of the tested AI models. An overview of the project execution steps is summarized in **Figure 1**.

Ethics, integrity and compliance with data protection regulatory frameworks will be integral and critical to the project, guaranteeing regulatory compliance in all aspects of the work performed both at the technical and clinical level. To ensure this, all project actions will be carefully guided and supervised by legal experts on data privacy.

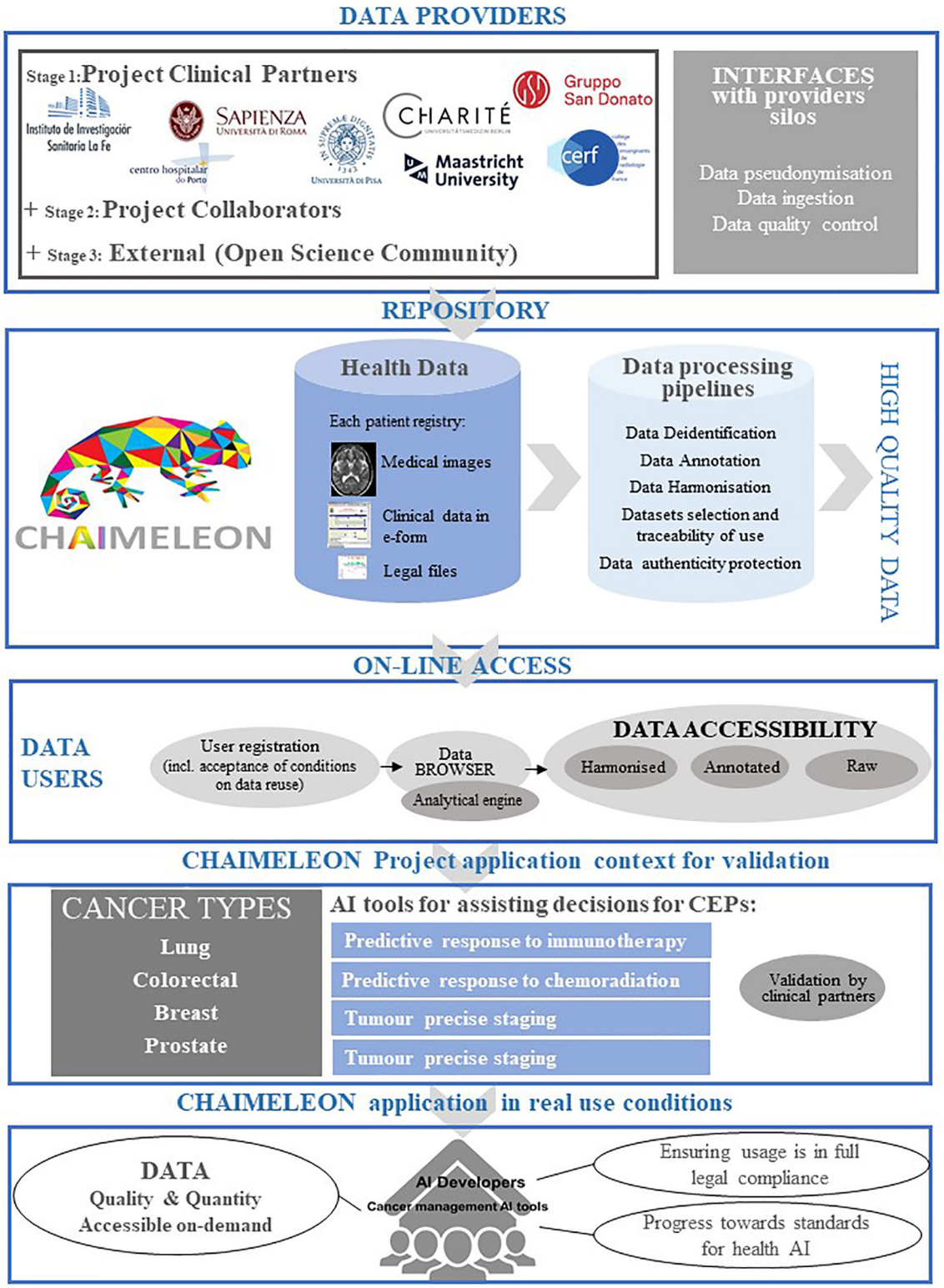


FIGURE 1 | Project overview.

The successful execution of this project will make multiple resources available to the research community. These will include data infrastructures, legal protocols, new AI models and tools, and methodologies for performance validation and enhanced interpretability of AI solutions. All these resources will be ultimately validated in the context of lung, breast, colorectal and prostate cancer. Given that these are such complex and heterogeneous diseases, this will allow to demonstrate the versatility and scalability of the repository and tools across a range of heterogeneous cancer applications. The final aim is to design the repository as a transversal resource with the required versatility and scalability to incorporate data for other types of cancer and imaging modalities in the near future.

MATERIALS AND METHODS

Project Consortium

To bring this project to fruition, an interdisciplinary consortium was recruited, bringing together experts in the fields of IT systems, automated health data management, data privacy and legal compliance, and researchers with experience in the design, set-up and management of imaging repositories and radiomic features. The CHAIMELEON consortium involves 18 partners from 9 European countries (Spain, Germany, France, Austria, United Kingdom, Israel, Italy, The Netherlands and Portugal) constituting a pan-European ecosystem of knowledge, infrastructures, biobanks and technologies in oncology, AI/*In-silico* experimentation and cloud computing. The partner institutions include 9 hospitals, 3 universities, and 6 R&D centers and private companies. An advisory board consisting of a recognized group of experts in the fields of oncology and AI applied to cancer management has been designated to give general advice and guidance to the consortium.

Project Timeline and Management Strategy

The CHAIMELEON project is a 4-year long, EU funded project that started in September 2020 and will finalize in August 2024. Currently in its second year of execution, the project results already include a first complete version of the repository design, a selected set of standards to be used to ensure the repository's interoperability and the first proposal for its legal operating model among others. To ensure its correct implementation, the project has set 13 different milestones to be met over its entire duration. A final, populated version of the repository is expected to be available to the public on project month 34 (June 2023). A summary of relevant project milestones indicating dataset availability of the repository can be found in **Table 1**.

Distributed Architecture and IT Infrastructure of the Repository

To facilitate its scalability and promote cooperative work with the rest of the scientific community, the repository is cloud-based and built upon open standards. Likewise, the use of free, open-access services has been prioritized, to keep maintenance and running costs to a minimum while implementing a robust software infrastructure. In particular, CHAIMELEON uses European Open Science Cloud (EOSC) services to join the projects currently contributing to this initiative. **Table 2** summarizes CHAIMELEON's main design and infrastructure features.

The CHAIMELEON repository architecture has been set as a hybrid type of architecture whereby local data warehouses and tools deployed within hospitals streamline the process of data collection and curation, while a central repository allows (1) management and annotation of anonymized data (2) AI model training, and (3) use of data processing pipelines. Local tools deployed at hospitals are being built on the *Medexprim Suite*™ framework to allow cohort selection, image extraction, collection, extraction and mapping of clinical data, as well as data curation, quality control, anonymization and transfer to the central repository (**Figure 2**). This set of tools is currently undergoing constant customization in order to adjust to the specific IT environment of each data provider site.

The storage and processing platform (central repository) uses a set of horizontal technologies that enable the distributed storage of medical images and associated clinical data, along with a processing environment where different applications can run. In particular, the storage has been implemented through

TABLE 1 | Project milestones.

Milestone ID	Description	Due date (months)
M1	Initial repository design available for regulatory clearance. Repository's legal operational model established.	12
M2	Start of data collection at data provider sites, with clearance for data to be incorporated into the CHAIMELEON repository.	13
M3	Completion of the repository design phase and the verification of the repository's compliance with GDPR.	18
M4	First repository prototype released, fully interfaced with data provider sites	24
M5	Start of the repository's technical validation phase Stage 1 – Internal by project partners	30
M6	Start of the repository's technical validation phase Stage 2 – External validation *via* open challenges to the AI community	31
M7	End of the repository's technical validation phase Stage 1 – Internal validation completed and documented	34
M8	Execution of the repository's technical validation phase Stage 2 – External validation *via* open challenges	34
M9	Start of the repository's data expansion stage – addition of new datasets provided by external collaborators. Legal clearance and IT interfacing with selected centers.	37
M10	End of the repository's technical validation phase Stage 2 – External validation *via* open challenges	38
M11	Start of the clinical validation phase - observational studies for AI-based solutions developed/refined using the repository start	41
M12	End of the clinical validation phase	46
M13	Assessment of observational studies finalized.	48

TABLE 2 | Summary of CHAIMELEON's main design and infrastructure features.

Feature	Description
Distributed infrastructure	In the first phase of the project, data will be centralized, after it has been collected, curated and anonymized by a set of tools deployed locally. In the second phase of the project, we will explore a distributed architecture, where the architecture will be composed of a central index and multiple physical repositories (local indexes), which may be either regional, national or hospital-based data warehouses connected to the hospital's PACS and EHR/RIS. Repositories will be connected using encrypted communications and standards for interoperability, such as DICOM-TLS or DICOM web. Federated Learning approaches and distributed data exploration solutions will be explored.
Single-entry point for pan-European users	CHAIMELEON will be designed to facilitate AI developers access to any relevant curated datasets, independently of their origin.
Publicly available, upon user registration	The registration process will include requirements for the researchers to sign acceptance of the conditions of use and access to the data. These will include commitments related to the purposes of data use and contracting of non-identification commitments.
Types of roles	Different entities and physical persons under different roles will be key parties to the repository, including data providers, entities providing infrastructure or services (primary data users), and researchers willing to access data for research purposes (data users). Roles will be carefully defined and assigned the applicable rights and obligations.
Powered with automatic tools, human refined	The latest machine learning advancements on data ingestion, curation, quality control, annotation, segmentation and harmonization will be incorporated into CHAIMELEON. During the project execution, extensive human resources will be devoted to the supervision and refinement of the automation tools. As technologies evolve, the repository will steadily progress towards less human supervision and more automated processes.
Pseudonymized and anonymized data	The Repository will have two levels of de-identification. The first one will be pseudonymization at local premises, in order to preserve traceability and enable potential linkage to other biobanks (e.g., Pathological or genetic). The second, at the central repository level, will be complete anonymization, meaning the data will no longer be identifiable, even indirectly.

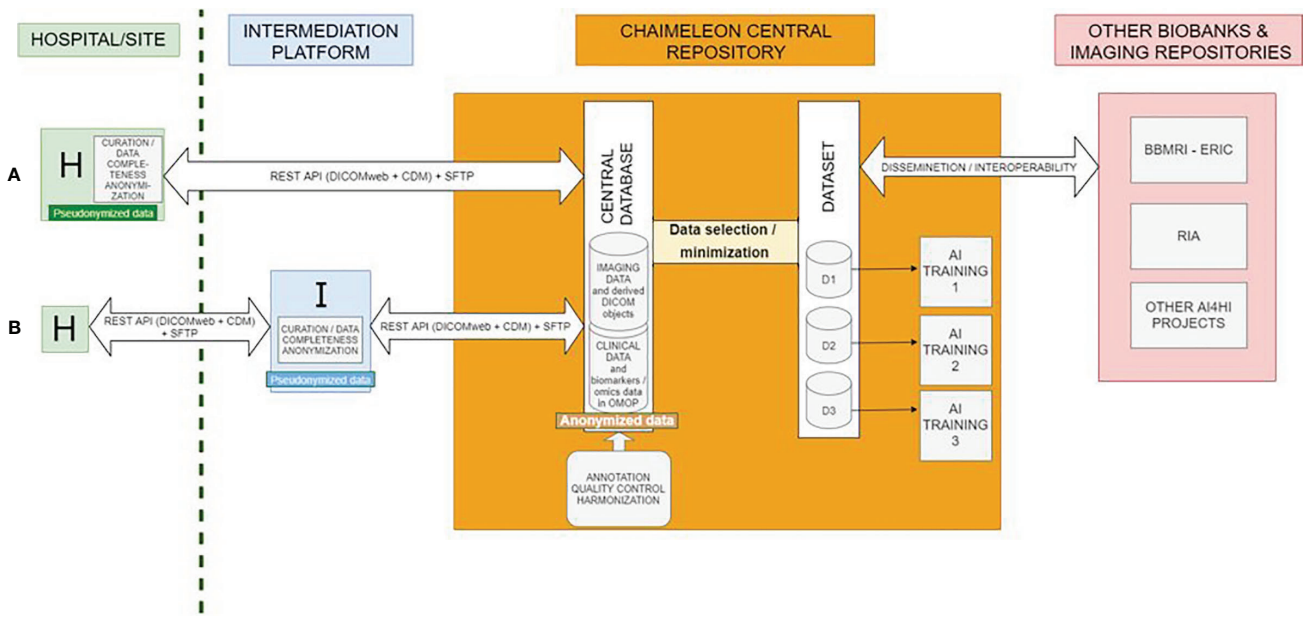


FIGURE 2 | Distributed architecture and IT infrastructure. **(A)** and **(B)** refer to two different types of hospitals based on their capacity to curate, complete, and anonymize data prior to their ingestion into the central repository. **(A)** Data processing is done on site. **(B)** Data processing is done via an intermediation platform.

CEPH (19) running on an on-premise cloud infrastructure implemented through OpenStack (20). The application and services run embedded on Docker (21) containers orchestrated by a Kubernetes (22) platform. Images are stored as DICOM files on a filesystem hierarchy, using access control lists to manage authorization. Clinical data is stored on a Mongo DB (23) through standardized e-forms and authentication is based on OpenID (24). A set of services has been developed to map access permissions, mount selected data on virtual workstation consoles and persist access and data processing actions on a Blockchain registry. Some data processing applications including one for image harmonization are currently being developed by project partners as one of the main challenges to this project. The overall platform architecture and list of processing applications is illustrated in **Figure 3**. In line with making the repository data FAIR, data ingestion processes ensure the incorporation of relevant information and make sure datasets are searchable by different criteria such as type of disease, imaging modality, or patient´s gender and age. Registered users (data requesters adhering to the data usage policy/license) will have on-line,

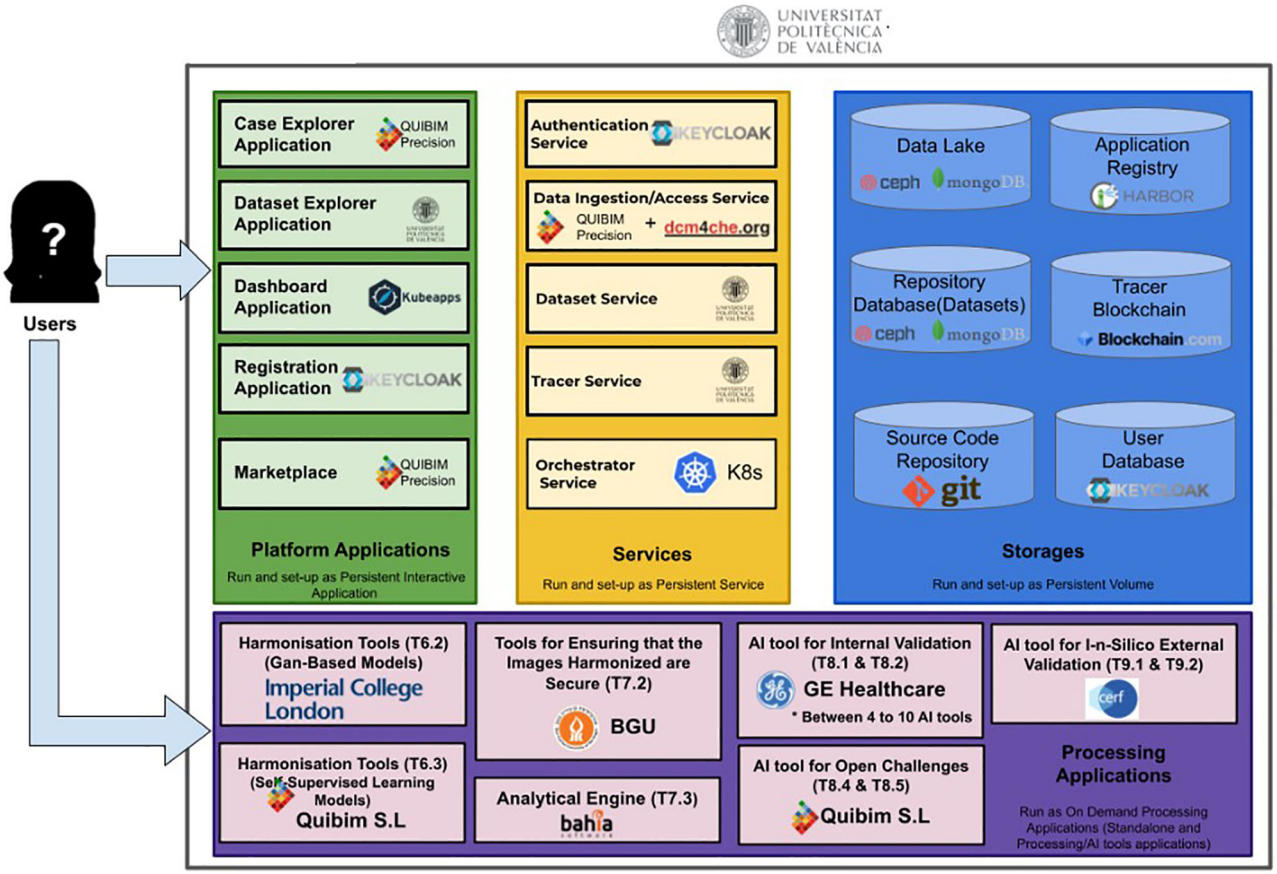


FIGURE 3 | High-level architecture of the central repository and technologies used.

controlled access to anonymized data only. Accessibility to pseudonymized data, on the other hand, will be limited to authorized repository managers locally in order to enable linkage to other related biobanks and the use of standards for metadata exchange and annotation. Data usage policies and licenses are currently being defined for users to commit to the reuse data for research purposes mainly.

The repository design has been built under the principles of legal compliance and privacy and security by design and by default. To do so, the consortium's data protection officer has performed a risk analysis and data protection impact assessment, identifying security, ethics, and regulatory compliance risks, as well as those associated with the re-identification of patients and the potential impact of the use of AI.

The long-term sustainability of the repository is being promoted by design, as it is being established as a controlled access cloud-based repository hosting reusable methodologies and protocols which will be interoperable with similar existing initiatives. Other means for the continuity of the repository will be assessed over the project lifetime, including creating synergies with new repositories and biobanks and incorporating data for other types of cancer and imaging modalities. Methodologies and protocols, including image processing tools, automatic annotation pipelines, and tools promoting interpretability of the AI models will be made available for reuse by similar research initiatives in the health imaging community.

Repository Population and Image Harmonization Strategies

The data to be made accessible in the repository will include imaging data (complete radiological studies for a given case) in DICOM format, linked to the correspondent e-form including relevant clinical features on the patient´s profile (age, gender, ethnicity, symptoms, comorbidities, etc.), tumor data (pathological, molecular and genetic), treatment details and outcome.

These data will be provided by the clinical partner institutions in the consortium which will populate the repository with retrospective cases during the project lifetime over two data recruitment periods. All cases will correspond to real world data, meaning the study subjects will be patients diagnosed in the participant hospitals for the 4 targeted types of cancer, collected through the routine delivery of healthcare with no enrolment conditions. Local imaging protocols will be used all over the data recruitment period. All cases included will be fully closed cases, that is, those for which all the required data, as specified in the

project requirements, are already available for a given patient. More specifically, both clinical and imaging data will be collected over the period comprised between the date of diagnosis and a 12 or 24 month follow up, depending on the type of cancer.

It is estimated that de-identified medical imaging data and related clinical data for a total of 11.500 cases of lung cancer, 6.000 of breast cancer, 6.000 of colorectal cancer and 10.000 of prostate cancer approximately will be accessible from the repository. **Table 3** shows a summary of the type of data expected to be accessible in CHAIMELEON per type of cancer.

To further accelerate the Repository´s population and to evidence its scalability and adaptability, during project Stage 2 (years 3 and 4) external collaborator hospitals from 5 additional countries will join the data provision efforts. These independent cases will be used for the external validation of the AI tools within the repository architecture. With both internal and external cases incorporated, the total size forecasted for the repository by project end is of nearly 40.000 cases, corresponding to approximately 20 million images (25).

To address the data inhomogeneity derived from the multiple sources the images will originate from, computational data harmonization methods such as self-supervised learning and GAN based models will be developed, in an aim to merge the data from the different sources into a single coherent data set by modifying data formats, terminologies, and measurement units (12, 13, 26). The project will consider reusing the results of the Imaging Biomarker Standardisation Initiative (27). Computational data harmonization methods, such as ComBat including its variations (BM-ComBat and QN-ComBat), and Distance-Weighted Discrimination, will also be explored (11).

Lastly, and so as to ensure that CHAIMELEON will offer only high-quality datasets, a data submission process will be implemented, involving the performance of quality assurance tests and the assignment of data set quality scores upon data submission onto the central repository. A detailed description of the quality check process is described in the **Supplementary Material**.

Data Models and FAIR Principles

CHAIMELEON's approach to image metadata is oriented at enabling data interoperability following FAIR data principles. As such, the data model to be implemented will refer to the current DICOM-MIABIS joint model (28), which proposes a first integration of the international DICOM standards into the MIABIS (Minimum Information About Biobank Data Sharing) core model. While MIABIS aims to standardize data elements used to describe biobanks and samples, the DICOM fields will be used to describe heterogeneous information across datasets, such as imaging protocols, modalities, sequences, scanners, and labels. Complete DICOM information (excluding patient information) will be available for further description of imaging collections. Therefore, the model will provide for the inclusion of metadata relating to the acquisition parameters, MRI and CT. Pseudo-anonymization processes, where applicable, will be programmed for the conservation of essential metadata such as patient preparation protocol. In this manner, this project further highlights the current need of datasheets for datasets (29)

Regarding the Common Data Model (CDM) to be used for clinical data across clinical centers, the use of the OMOP (Observational Medical Outcomes Partnership) CDM (30) was agreed upon, which is based on standard ontologies such as SNOMED CT] (31) or ICD-10 (32). As such, the local clinical data warehouse deployed at each hospital is built using the OMOP CDM, and the set of necessary clinical variable terms to collect has been standardized according to this standard. As of project month 16, a total of 930 terms (97%) have been successfully adapted to match the standard vocabularies with currently available concepts. The CHAIMELEON project is committed to contributing to the evolution of standards rather than developing proprietary formats. For this reason, an upgrade of the OMOP CDM has been suggested to include the 24 remaining terms (3%) given their relevance in current clinical practice.

As curated, annotated, and enriched datasets are being constituted, the MIABIS model will be used to label these collections and promote their reuse. Findability, on the other hand, will be provided through the use of persistent identifiers that will be assigned to the datasets and persisted on the Blockchain. Accessibility will be ensured through the use of an open Application Programming Interface (API) to query the dataset service. Reusability will then be enhanced by the provision of virtual processing environments.

Development of AI Tools for Cancer Management

The project will contribute to the development, refining, testing and early clinical validation of AI tools targeted to reproducibly assist clinicians in the precise estimation of some of the currently most relevant Clinical End Points (CEPs) in cancer (**Table 4**) (33).

TABLE 3 | Types of datasets to be accessible from the CHAIMELEON repository.

Type of cancer	Imaging Data	Estimated number of cases	
		Training phase	Validation phase
Lung cancer	CT/PET/CT	7000	4500
Breast cancer	Mammography, Digital breast tomosynthesis, Ultrasound and MRI	3500	2500
Colorectal cancer			
Colon	CT	2334	1667
Rectum	MRI	1167	833
Prostate cancer	MRI	6000	4000

TABLE 4 | Clinical end points to be addressed in CHAIMELEON for the four targeted types of cancer.

Type of Cancer	Current therapies	CEPs
Lung	Immunotherapy	Predicting patients with a positive response to immunotherapy
Colorectal	Surgery/neoadjuvant chemotherapy	(Rectal cancer) Prediction of patients with a positive response to chemoradiation and classification in different treatment response sub-groups. (Colon cancer) Identification of patients at higher risk of distant metastases at an early timepoint.
Breast	Surgery, radiation and systemic therapy	Diagnostic performance and cancer staging.
Prostate	Wide range due to heterogeneity	Early Staging/Grading

The tools include image preparation and harmonization, tissue segmentation, radiomics data extraction, treatment allocation and prognosis prediction (7, 34–37). Once these are tested, Open Challenges will be organized promoting other world-class developers to use CHAIMELEON resources to train their own models.

This project will follow a methodological approach of continuous learning, allowing a smooth update of the models including new data annotations and training to progressively improve performance over time. It is expected that this continuous learning will provide the repository infrastructure with large scalability in terms of algorithm performance and management of datasets.

Although many AI models that predict factors related to both disease phenotyping and treatment effects have recently been published (38–45), important challenges remain in the standardization of the criteria for evaluation of model performance, reproducibility and clinical utility. The principal challenge remains the optimal collection and integration of diverse multimodal data sources in such quantitative manner that delivers unambiguous clinical predictions. In CHAIMELEON, we aim to undertake ground-breaking research on the AI space leading to a new paradigm in the investigation of imaging biomarkers at multi-center studies and clinical trials, overcoming the problem of reproducibility in QIBs.

The cancer management solutions to be developed in the context of this project include those assisting radiologists in image processing and analysis, impacting diagnosis and follow up capabilities, helping predict tumor behavior, as well as aiding patient's stratification, therapy allocation and tumor response to treatment.

Technical and Clinical Validation of AI Solutions

By the end of the project, the performance of the repository as tool fostering AI experimentation will be validated. This technical validation phase will assess the usefulness of the repository for accelerating experimentation of AI solutions and contributing to better training and testing of the AI models. This process will occur in two subsequent steps. First, AI developers of the consortium will undertake training and testing of a selected set of their proprietary AI tools using the data provided by consortium partners. Secondly, an external validation subphase will be done by other world-class European AI developers, initiative that will be articulated *via* Open Challenges to the AI community.

Lastly, a clinical validation phase will assess how the technically validated tools can assist clinicians in addressing the selected CEPs for lung, breast, colorectal and prostate cancer. To do so, observational *In-silico* studies will be designed by cancer experts to assess the capacity of these tools to aid clinical decision-making in cancer management in terms of diagnosis, prognosis, treatment selection and treatment follow-up. Cases provided by external collaborator centers will be used for this clinical validation phase to ensure the reproducibility of the results on real world cases.

EXPECTED RESULTS

The CHAIMELEON repository along with its related AI-powered tools are being designed to impact the management of the four most prevalent types of cancer worldwide. Due to the social and economic burden these imply, we expect the outcome of this project to have an EU-wide impact both at the social and economic level. Upon successful validation on how the proposed AI tools can assist clinicians in daily decision making, we expect the repository infrastructure, legal operational model, analysis tools and web-based user interfaces to have the potential to be adapted to the management of other types of cancer (46).

The project will contribute to the current state of the art of AI for health imaging by defining a framework for legitimate access to anonymized imaging and related clinical data provided by hospitals in different European nations, making these more openly accessible across the EU for secondary use in research. From image acquisition to image evaluation and QIB reporting, our work is aligned with exciting research on the use of AI for improving image quality and interpretation. At the technical level, we hope to contribute to the advancement in robustness of AI systems against malicious attacks, interpretability of AI-based models, and validation of AI tools in clinical observational studies. We also aim to impact the fields of standardization of radiological procedures for image acquisition and analysis, as well as harmonization for extraction of reproducible imaging biomarkers.

From a legal perspective, this project will contribute to the creation of ethical standards for the use of health imaging data in the context of AI tool development. All in all, we expect CHAIMELEON to generate resources that facilitate faster and

more successful development of AI-based solutions for cancer management, while promoting actions to foster the evolution of European laws in the reuse of health data for research purposes (47). By doing so, we expect to have a major mid to long-term impact in the European health imaging sector, increasing trust in AI solutions amongst healthcare professionals, patients, and stakeholders in both industry and academia.

DISCUSSION

The EU-funded CHAIMELEON project aims to set up one of the most ambitious health imaging repositories across the EU, contributing to major advancements in the field of cancer management at a global scale. Once implemented, it will have the potential to generate innovative products and services beyond the direct outcomes planned for the project. The developed methodologies and protocols will pave the way for the use of this type of repository, not just in other fields of biomedical research but also in any other disciplines where public interest is the main driver.

Upon successful demonstration of the disruptive analysis approaches used in this project, such as those related to the improvement of the interpretability of AI models, these will have the potential to be used in other currently prevalent pathologies, such as cardiovascular, neurological or psychiatric diseases. The developed AI-powered tools may be used as clinical support systems for complex diagnosis, and contribute to new diagnostic approaches based on telemedicine or second opinion.

To further ensure future interoperability across repositories of the same nature, the CHAIMELEON is taking part in the AI for Health Imaging (AI4HI) league, a collaborative network of similar projects funded under the same topic. The main goal to this collaboration is to ensure the long-term sustainability of these kind of repositories, and to promote cooperation and data sharing among users.

Further information on the main project objectives, partners and contributors, work progress and latest updates on project results can be found on the project website (chaimeleon.eu) and social media platforms.

DATA AVAILABILITY STATEMENT

The data collected for this project will be real world patient data, collected retrospectively through the routine delivery of healthcare with no enrollment conditions. This is a purely observational, non-interventional study. Requests to access these datasets should be directed to LM, marti_lui@gva.es.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comité de Ética de la Investigación con Medicamentos (CEIM), Hospital Universitario y Politécnico La Fe. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

AM and LM wrote the manuscript. AS and MA conceived the project idea and wrote the proposal for funding with the support of LM. The rest of authors contributed to the conception of the project idea and the project proposal preparation.

FUNDING

CHAIMELEON has been funded by as a Horizon 2020 project (RIA, topic DT-TDS-05-2020-AI for Health Imaging; call SC1-FA-DTS-2019-1, under Grant Agreement No. 952172).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.742701/full#supplementary-material

REFERENCES

1. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial Intelligence in Cancer Imaging: Clinical Challenges and Applications. *CA A Cancer J Clin* (2019) 69:caac.21552. doi: 10.3322/caac.21552
2. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine Learning Applications in Cancer Prognosis and Prediction. *Comput Struct Biotechnol J* (2015) 13:8–17. doi: 10.1016/j.csbj.2014.11.005
3. Li Z, Wang Y, Yu J, Guo Y, Cao W. Deep Learning Based Radiomics (DLR) and Its Usage in Noninvasive IDH1 Prediction for Low Grade Glioma. *Sci Rep* (2017) 7:5467. doi: 10.1038/s41598-017-05848-2
4. Forghani R, Chatterjee A, Reinhold C, Pérez-Lara A, Romero-Sanchez G, Ueno Y, et al. Head and Neck Squamous Cell Carcinoma: Prediction of Cervical Lymph Node Metastasis by Dual-Energy CT Texture Analysis With Machine Learning. *Eur Radiol* (2019) 29:6172–81. doi: 10.1007/s00330-019-06159-y
5. MERCURY Study Group. Diagnostic Accuracy of Preoperative Magnetic Resonance Imaging in Predicting Curative Resection of Rectal Cancer: Prospective Observational Study. *BMJ* (2006) 333:779. doi: 10.1136/bmj.38937.646400.55
6. Thoeny HC, Ross BD. Predicting and Monitoring Cancer Treatment Response With Diffusion-Weighted MRI. *J Magn Reson Imaging* (2010) 32:2–16. doi: 10.1002/jmri.22167
7. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: Extracting More Information From Medical Images Using Advanced Feature Analysis. *Eur J Cancer* (2012) 48:441–6. doi: 10.1016/j.ejca.2011.11.036
8. European Society of Radiology (ESR). ESR Position Paper on Imaging Biobanks. *Insights Imaging* (2015) 6:403–10. doi: 10.1007/s13244-015-0409-x
9. Diaz O, Kushibar K, Osuala R, Linardos A, Garrucho L, Igual L, et al. Data Preparation for Artificial Intelligence in Medical Imaging: A Comprehensive Guide to Open-Access Platforms and Tools. *Phys Med* (2021) 83:25–37. doi: 10.1016/j.ejmp.2021.02.007

10. Mollura DJ, Culp MP, Pollack E, Battino G, Scheel JR, Mango VL, et al. Artificial Intelligence in Low- and Middle-Income Countries: Innovating Global Health Radiology. *Radiology* (2020) 297:513–20. doi: 10.1148/radiol.2020201434
11. Fortin J-P, Parker D, Tunç B, Watanabe T, Elliott MA, Ruparel K, et al. Harmonization of Multi-Site Diffusion Tensor Imaging Data. *NeuroImage* (2017) 161:149–70. doi: 10.1016/j.neuroimage.2017.08.047
12. Gao Y, Liu Y, Wang Y, Shi Z, Yu J. A Universal Intensity Standardization Method Based on a Many-To-One Weak-Paired Cycle Generative Adversarial Network for Magnetic Resonance Images. *IEEE Trans Med Imaging* (2019) 38:2059–69. doi: 10.1109/TMI.2019.2894692
13. Tor-Diez C, Porras AR, Packer RJ, Avery RA, Linguraru MG. Unsupervised MRI Homogenization: Application to Pediatric Anterior Visual Pathway Segmentation. In: M Liu, P Yan, C Lian, X Cao, editors. *Machine Learning in Medical Imaging Lecture Notes in Computer Science*. Cham: Springer International Publishing. (2020) p. 180–8. doi: 10.1007/978-3-030-59861-7_19
14. Martí-Bonmatí L, Alberich-Bayarri Á, Ladenstein R, Blanquer I, Segrelles JD, Cerdá-Alberich L, et al. PRIMAGE Project: Predictive *In Silico* Multiscale Analytics to Support Childhood Cancer Personalised Evaluation Empowered by Imaging Biomarkers. *Eur Radiol Exp* (2020) 4:22. doi: 10.1186/s41747-020-00150-9
15. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med* (2015) 12: e1001779. doi: 10.1371/journal.pmed.1001779
16. Bamberg F, Kauczor H-U, Weckbach S, Schlett CL, Forsting M, Ladd SC, et al. Whole-Body MR Imaging in the German National Cohort: Rationale, Design, and Technical Background. *Radiology* (2015) 277:206–20. doi: 10.1148/radiol.2015142272
17. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *J Digit Imaging* (2013) 26:1045–57. doi: 10.1007/s10278-013-9622-7
18. Gabelloni M, Faggioni l, Borgheresi R, Restante G, Shortrede J, Tumminello L, et al. Bridging Gaps Between Images and Data: A Systematic Update on Imaging Biobanks. *Eur Radiol* (2021). doi: 10.1007/s00330-021-08431-6
19. Bugawan TL, Klitz W, Blair A, Erlich HA. High-Resolution HLA Class I Typing in the CEPH Families: Analysis of Linkage Disequilibrium Among HLA Loci: High-Resolution HLA Class I Typing in the CEPH Families. *Tissue Antigens* (2000) 56:392–404. doi: 10.1034/j.1399-0039.2000.560502.x
20. Solano A, Dormido R, Duro N. Sánchez J. A Self-Provisioning Mechanism in OpenStack for IoT Devices. *Sensors* (2016) 16:1306. doi: 10.3390/s16081306
21. Digan W, Countouris H, Barritault M, Baudoin D, Laurent-Puig P, Blons H, et al. An Architecture for Genomics Analysis in a Clinical Setting Using Galaxy and Docker. *GigaScience* (2017) 6:1–9. doi: 10.1093/gigascience/gix099
22. Poniszewska-Marańda A, Czechowska E. Kubernetes Cluster for Automating Software Production Environment. *Sensors* (2021) 21:1910. doi: 10.3390/s21051910
23. Holik M, Horvath T, Oujezsky V, Munster P, Tomasov A, Valach S. MongoDB Database as Storage for GPON Frames. *Sensors* (2020) 20:6208. doi: 10.3390/s20216208
24. Ma W, Sartipi K, Sharghigoorabi H, Koff D, Bak P. OpenID Connect as a Security Service in Cloud-Based Medical Imaging Systems. *J Med Imag* (2016) 3:026501. doi: 10.1117/1.JMI.3.2.026501
25. Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the Sample Size Required for Developing a Clinical Prediction Model. *BMJ* (2020) 368:m441. doi: 10.1136/bmj.m441
26. Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, et al. Preparing Medical Imaging Data for Machine Learning. *Radiology* (2020) 295:4–15. doi: 10.1148/radiol.2020192224
27. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-Based Phenotyping. *Radiology* (2020) 295:328–38. doi: 10.1148/radiol.2020191145
28. Scapicchio C, Gabelloni M, Forte SM, Alberich LC, Faggioni L, Borgheresi R, et al. DICOM-MIABIS Integration Model for Biobanks: A Use Case of the EU PRIMAGE Project. *Eur Radiol Exp* (2021) 5:20. doi: 10.1186/s41747-021-00214-4
29. Gebru T, Morgenstern J, Vecchione B, Wortman Vaughan J, Wallach H, Daumé HIII, et al. Datasheets for Datasets. *Communications of the ACM* (2021) 64(12):86–92. doi: 10.1145/3458723
30. Rubin DL, Kahn CE. Common Data Elements in Radiology. *Radiology* (2017) 283:837–44. doi: 10.1148/radiol.2016161553
31. Lee D, de Keizer N, Lau F, Cornet R. Literature Review of SNOMED CT Use. *J Am Med Inform Assoc* (2014) 21:e11–9. doi: 10.1136/amiajnl-2013-001636
32. Hirsch JA, Leslie-Mazwi TM, Nicola GN, Oklu R, Schoppe KA, Silva E, et al. The ICD-10 System: A Gift That Keeps on Taking: Table 1. *J NeuroIntervent Surg* (2015) 7:619–22. doi: 10.1136/neurintsurg-2014-011321
33. McLeod C, Norman R, Litton E, Saville BR, Webb S, Snelling TL. Choosing Primary Endpoints for Clinical Trials of Health Care Interventions. *Contemp Clin Trials Commun* (2019) 16:100486. doi: 10.1016/j.conctc.2019.100486
34. Huang H, Yang G, Zhang W, Xu X, Yang W, Jiang W, et al. A Deep Multi-Task Learning Framework for Brain Tumor Segmentation. *Front Oncol* (2021) 11:690244. doi: 10.3389/fonc.2021.690244
35. Liu Y, Zheng H, Liang Z, Miao Q, Brisbane WG, Marks LS, et al. Textured-Based Deep Learning in Prostate Cancer Classification With 3T Multiparametric MRI: Comparison With PI-RADS-Based Classification. *Diagnostics* (2021) 11:1785. doi: 10.3390/diagnostics11101785
36. Walls GM, Osman SOS, Brown KH, Butterworth KT, Hanna GG, Hounsell AR, et al. Radiomics for Predicting Lung Cancer Outcomes Following Radiotherapy: A Systematic Review. *Clin Oncol* (2021) 21:S0936655521003721. doi: 10.1016/j.clon.2021.10.006
37. Morin O, Vallières M, Braunstein S, Ginart JB, Upadhaya T, Woodruff HC, et al. An Artificial Intelligence Framework Integrating Longitudinal Electronic Health Records With Real-World Data Enables Continuous Pan-Cancer Prognostication. *Nat Cancer* (2021) 2:709–22. doi: 10.1038/s43018-021-00236-2
38. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial Intelligence in Digital Pathology — New Tools for Diagnosis and Precision Oncology. *Nat Rev Clin Oncol* (2019) 16:703–15. doi: 10.1038/s41571-019-0252-y
39. Geras KJ, Mann RM, Moy L. Artificial Intelligence for Mammography and Digital Breast Tomosynthesis: Current Concepts and Future Perspectives. *Radiology* (2019) 293:246–59. doi: 10.1148/radiol.2019182627
40. Patel SK, George B, Rai V. Artificial Intelligence to Decode Cancer Mechanism: Beyond Patient Stratification for Precision Oncology. *Front Pharmacol* (2020) 11:1177. doi: 10.3389/fphar.2020.01177
41. Sakamoto T, Furukawa T, Lami K, Pham HHN, Uegami W, Kuroda K, et al. A Narrative Review of Digital Pathology and Artificial Intelligence: Focusing on Lung Cancer. *Transl Lung Cancer Res* (2020) 9:2255–76. doi: 10.21037/tlcr-20-591
42. Zhu W, Xie L, Han J, Guo X. The Application of Deep Learning in Cancer Prognosis Prediction. *Cancers* (2020) 12:603. doi: 10.3390/cancers12030603
43. Tătaru OS, Vartolomei MD, Rassweiler JJ, Virgil O, Lucarelli G, Porpiglia F, et al. Artificial Intelligence and Machine Learning in Prostate Cancer Patient Management—Current Trends and Future Perspectives. *Diagnostics* (2021) 11:354. doi: 10.3390/diagnostics11020354
44. Trivizakis E, Papadakis G, Souglakos I, Papanikolaou N, Koumakis L, Spandidos D, et al. Artificial Intelligence Radiogenomics for Advancing Precision and Effectiveness in Oncologic Care (Review). *Int J Oncol* (2020) 57:43–53. doi: 10.3892/ijo.2020.5063
45. Neri E, Del Re M, Paiar F, Erba P, Cocuzza P, Regge D, et al. Radiomics and Liquid Biopsy in Oncology: The Holons of Systems Medicine. *Insights Imaging* (2018) 9:915–24. doi: 10.1007/s13244-018-0657-7
46. Lekadir K, Osuala R, Gallin C, Lazrak N, Kushibar K, Tsakou G, et al. FUTURE-AI: Guiding Principles and Consensus Recommendations for Trustworthy Artificial Intelligence in Medical Imaging. *1 ArXiv210909658* (2021), 1–36.
47. Mitchell C, Ploem C. Legal Challenges for the Implementation of Advanced Clinical Digital Decision Support Systems in Europe. *J Clin Transl Res* (2018) 3:424–30.

Conflict of Interest: Authors AS, MA were employed by Matical Innovation SL. MM was employed by GE Healthcare. KS was employed by Medexprim. SF way employed by Bahia Software S.L.U. AA-B way employed by QUIBIM SL.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer KL declared a past co-authorship with several of the authors LM, PL to the handling editor.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of

the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Bonmatí, Miguel, Suárez, Aznar, Beregi, Fournier, Neri, Laghi, França, Sardanelli, Penzkofer, Lambin, Blanquer, Menzel, Seymour, Figueiras, Krischak, Martínez, Mirsky, Yang and Alberich-Bayarri. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 02 March 2022
doi: 10.3389/fonc.2022.805263

RDAU-Net: Based on a Residual Convolutional Neural Network With DFP and CBAM for Brain Tumor Segmentation

Jingjing Wang[1], Zishu Yu[1], Zhenye Luan[1], Jinwen Ren[1], Yanhua Zhao[2] and Gang Yu[1]**

[1] College of Physics and Electronics Science, Shandong Normal University, Jinan, China, [2] Obstetrics and Gynecology, Tengzhou Xigang Central Health Center, Tengzhou, China

OPEN ACCESS

Edited by:
Kyung Hyun Sung, UCLA Health System, United States

Reviewed by:
Jue Jiang, Memorial Sloan Ketter Cancer Center, United States
Weiwei Zong, Henry Ford Health System, United States

***Correspondence:**
Yanhua Zhao
13561119458@163.com
Gang Yu
yugang2011@foxmail.com

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *30 October 2021*
Accepted: *14 January 2022*
Published: *02 March 2022*

Citation:
Wang J, Yu Z, Luan Z, Ren J, Zhao Y and Yu G (2022) RDAU-Net: Based on a Residual Convolutional Neural Network With DFP and CBAM for Brain Tumor Segmentation. Front. Oncol. 12:805263. doi: 10.3389/fonc.2022.805263

Due to the high heterogeneity of brain tumors, automatic segmentation of brain tumors remains a challenging task. In this paper, we propose RDAU-Net by adding dilated feature pyramid blocks with 3D CBAM blocks and inserting 3D CBAM blocks after skip-connection layers. Moreover, a CBAM with channel attention and spatial attention facilitates the combination of more expressive feature information, thereby leading to more efficient extraction of contextual information from images of various scales. The performance was evaluated on the Multimodal Brain Tumor Segmentation (BraTS) challenge data. Experimental results show that RDAU-Net achieves state-of-the-art performance. The Dice coefficient for WT on the BraTS 2019 dataset exceeded the baseline value by 9.2%.

Keywords: brain tumor segmentation, U-Net, attention mechanism, dilation feature pyramid, deep learning

INTRODUCTION

Tumors that grow in the skull are commonly referred to as brain tumors and include primary brain tumors, which occur in the brain parenchyma, and secondary brain tumors, which metastasize to the skull from other parts of the body. According to the World Health Organization (WHO) classification criteria, brain tumors are classified into four grades: grade I, astrocytoma; grade II, oligodendroma gliomas; grade III, anaplastic glioma; and grade IV, glioblastoma multiforme (GBM) (1). The lower the grade of the tumor, the less malignant it is, and the better the prognosis is. As a result, early diagnosis of brain tumors is very important for treatment.

Magnetic resonance imaging (MRI) is considered a standard technique due to its satisfactory soft-tissue contrast and wide availability (2). MRI is a noninvasive imaging technique that uses magnetic resonance phenomena to obtain electromagnetic signals from the human body and reconstruct information about the body as a type of tomography. MRI is available in a variety of imaging sequences. These imaging sequences can produce MRI images with distinctive features that can reflect the anatomical morphology of the human body.

In current clinical practice, brain tumors are labeled manually by physicians, which is time-consuming. Moreover, brain tumors are similar to normal brain tissues in terms of morphology and

intensity; hence, manual labeling by physicians suffers from subjective variability and lacks reproducibility (3). Therefore, accurate automatic segmentation of brain tumors in T1, T1-c, T2, and FLAIR is essential for quantitative analysis and evaluation of brain tumors (4).

Progress on Image Segmentation Methods

In recent years, deep neural network (DNN)-based methods have achieved high performance for brain tumor segmentation (5–8). Convolutional neural networks (CNNs) (9) have achieved great success in many research areas, such as image recognition (10–12), image segmentation (13–15), and natural language processing (16, 17). In (18), a fully convolutional neural network (FCN) for image pixel-level image classification was proposed, which solves the problem of semantic-level image segmentation with input data of arbitrary size. Ronneberger et al. (19) proposed the U-Net framework with a skip connection module connecting the encoder and decoder. In contrast to FCNs, U-Net fuses shallow and deep features and has produced impressive results in medical image segmentation. Inspired by U-Net, Attention U-Net (20) and ResU-Net (21) were proposed and used for medical image segmentation. In Attention U-Net, an attention mechanism is added to the skip connection part. This module generates gating information to readjust the weight coefficients of features at various spatial locations. In ResU-Net, each convolutional layer is replaced with a residual convolutional layer, thereby avoiding gradient disappearance in backpropagation in deep network structures.

Attention Mechanism

Attention mechanisms were first introduced in natural language processing (22–25). Currently, attention mechanisms are also widely used in deep learning to enhance feature extraction (26–28). Hu et al. (29) proposed plug-and-play squeeze-and-excitation (SE) attention, which learns feature relationships to obtain contextual information on channel dimensions by global average pooling. Wang et al. (30) proposed ECANet, which uses a local cross-channel interaction strategy without downscaling and adaptive selection of one-dimensional convolutional kernels. In addition to these single-channel attention mechanisms, there are several dual-attention mechanisms. For example, Fu et al. (31) proposed the dual attention mechanism network (DANet) to improve the accuracy of network segmentation by capturing feature dependencies based on the spatial and channel dimensions of the self-attention mechanism and summing the outputs of the two modules. Woo et al. (32) proposed a convolutional block attention module (CBAM) to enhance useful information and suppress useless information by a tandem channel attention mechanism and a spatial attention mechanism.

Although ResU-Net uses a residual module to mitigate the problem of vanishing network gradients, it still suffers from the following problems: (1) Multiscale features have an important role, but ResU-Net does not extract features from images of various sizes, and thus, a substantial amount of detailed information is lost. (2) The skip connection cascades the shallow features of the decoder part and the corresponding depth features to achieve feature fusion, but the shallow features of the encoder contain considerable redundant information, which, in turn, affects the segmentation results.

In this paper, we propose RADU-Net, which is an improved version of ResU-Net that is inspired by the attention mechanism. Our contributions are mainly as follows.

1. We insert the 3D CBAM dual attention mechanism in each residual module to alleviate the problem of gradient disappearance or explosion as the network structure deepens and obtain the feature information of the image more accurately.
2. We add the dilated feature pyramid module with the 3D CBAM dual attention mechanism between the encoder and the decoder as a solution to the problem that the traditional U-Net network does not extract multi-scale features of images to obtain feature maps of different sizes.
3. We insert a 3D CBAM block after the skip connection in each layer to improve the extraction of channel information and spatial information to reduce the redundant information of low-level features.

Via these modifications, RADU-Net solves the above problems and improves the overall segmentation accuracy of brain tumors.

METHODS

RDAU-Net

This paper proposes a 3D convolutional neural network, namely, RDAU-Net, for the brain tumor segmentation task. The 3D CNN considers more comprehensive spatial context information and achieves more accurate performance than the 2D CNN in image segmentation. **Figure 1** illustrates the complete structure of RDAU-Net. RDAU-Net includes an encoder part and a decoder part. The input of the encoder part includes 2 purple 3D convolutional layers with a convolutional kernel size of 3 × 3 × 3, 5 3D RA blocks of various sizes and 4 3D convolutional layers with a step size of 2 as the downsampling layer. The RA block is that the orange CBAM block is added after each convolutional layer of the residual block for feature extraction. Between the encoder and the decoder is the DA block, which is the CBAM block that is added after each dilated convolution layer. The decoder part is symmetric to the encoder and contains 4 deconvolution layers. The purple arrow between the encoder and decoder is the skip connection, and a CBAM block is inserted after each skip connection layer. The pink block is the concatenation layer. The network ends with a 3D convolutional layer and a gray block of sigmoid function layers. (3,3,3,4,32) in the first layer indicates that the convolution kernel size of this layer is 3 × 3 × 3, the number of input features is 4 and the number of output features is 32.

3D CBAM Block

To extract more accurate image feature information, we use CBAM (32) as the attention module of the network. We transform the 2D CBAM attention module into a 3D CBAM

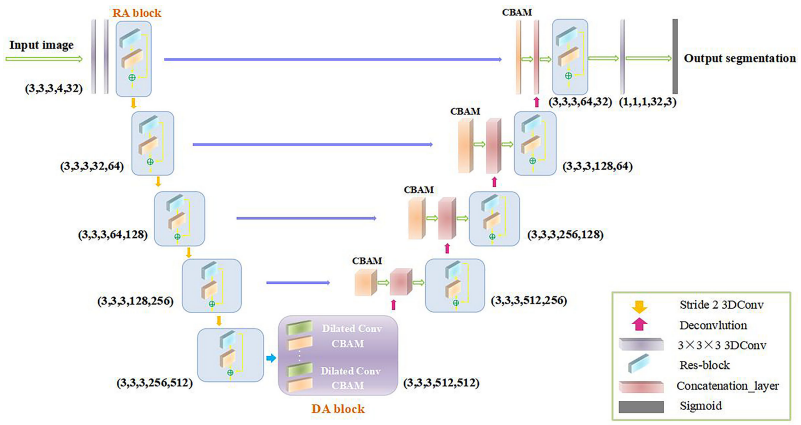


FIGURE 1 | Architecture of the proposed RDAU-Net.

attention module. CBAM ties together channel attention and spatial attention to increase the weights of useful features in the channel and useful features in the space, as illustrated in **Figure 2A**. First, the deep blue module in **Figure 2A**, which is denoted as $F \in R^{W\times H\times D\times C}$, is input into the channel attention module. The channel attention module as shown in **Figure 2B** consists of a 3D global max pooling module (the pink module in **Figure 2B**), a 3D global average pooling module (the yellow module in **Figure 2B**), and a shared MLP (multilayer perception) that consists of a 3D neural network. The input feature map F is subjected to 3D global max pooling and 3D global average pooling operations and to MLP. The two outputs of MLP are added elementwise. Then, after the sigmoid activation function, we obtain the weight $M_c(F)$ after the channel attention module.

$$M_C(F) = \sigma(MLP(AvgPool3D(F) + MLP(MaxPool3D(F)) = \sigma(W_1(W_0(F^C_{avg}) + W_1(W_0(F^C_{\max})) \quad (1)$$

where $W_0 \in R^{C/r\times C}$, $W_1 \in R^{C/C\times r}$, and σ is the sigmoid activation function. r takes a value of 16, namely, the channel C is changed to $C/16$ during max-pooling and average pooling to reduce the number of parameters.

Then, $M_c(F)$ is multiplied with the input feature map F to obtain the output feature map F' of the channel attention module, and the formula is as follows:

$$F' = M_C(F) \otimes F \quad (2)$$

where $M_c(F)$ denotes the output weight after the channel attention module and ⊗ denotes element-by-element multiplication.

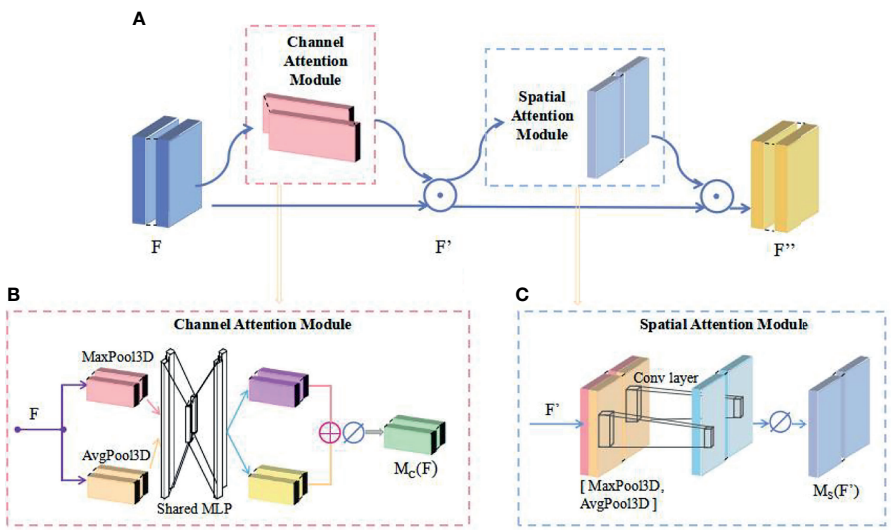


FIGURE 2 | Structure of a 3D CBAM block. **(A)** The overview of 3D CBAM; **(B)** Structure of the Channel Attention Module; **(C)** Structure of the Spatial Attention Module.

F' is used as the input feature map for the spatial attention model as shown in **Figure 2C**, and the channels are compressed by 3D global max pooling and 3D global average pooling in the channel dimension. The two extracted feature maps F^S_{avg} and $F^S_{\max}$ are subjected to a channel-based merging operation to obtain a 2-channel feature map, which is subsequently downscaled into a single channel (the shallow blue module in **Figure 2C**) by a 7×7 convolution operation before application of the sigmoid function to generate the output weight $M_S(F')$ in the blue part of the spatial attention module.

$$M_S(F'') = \sigma(f^{7\times7}([AvgPool3D(F''); MaxPool3D(F'')])) = \sigma(f^{7\times7}([F^S_{avg}; F^S_{\max}])) \quad (3)$$

Finally, $M_S(F')$ is multiplied by F' to obtain the final output feature map F'' of the yellow module, as expressed in Equation (3).

$$F'' = M_S(F'') \otimes F'' \quad (4)$$

where $M_S(F')$ denotes the output weight after spatial attention and $\otimes$ denotes element-by-element multiplication.

RA Block

We propose a residual block with a 3D CBAM, namely, an RA block that is composed of two 3×3×3 convolutional layers for the pink module, two normalization layers for the purple module, two activation layers for the yellow module, and an attention layer for the green module, as illustrated in **Figure 3**. Batch normalization (BN) is sensitive to the batch size because the mean and variance are calculated on a single batch. The instance normalization (IN) operation is performed within a single sample and does not depend on the batch. The leaky rectified linear unit (LR) is a variant of ReLU with a variation in response to input fractions of less than 0, thereby mitigating the sparsity of ReLU and alleviating some of the problems of neuronal death that are caused by ReLU. Therefore, we substitute the instance normalization (IN) and leaky rectified linear unit (LR) functions for the popular batch normalization (BN) and rectified linear unit (ReLU) functions, respectively. The RA block effectively improves the extraction of image feature information by adding a CBAM after the final convolution layer of the residual block, thereby improving the segmentation accuracy of the network.

DA Block

Our proposed DA block is shown as a deep blue block in **Figure 1**, which is composed of multiple green blocks of parallel 3×3×3 dilated convolution layers with various expansion rates and the orange CBAM blocks, and the multiple parallel feature maps are finally summed to obtain the output feature map of the purple block, as illustrated in **Figure 4**. Multiscale features are important for the segmentation of brain tumors. Therefore, we propose the DA block for efficiently obtaining feature maps of various sizes through levels of dilated convolutional layers in the spatial pyramid model.

Dice Loss Function

The BraTS dataset is extremely unbalanced, and convolutional neural networks are very sensitive to unbalanced datasets. Therefore, we use the dice loss function to solve the problem. The dice function is expressed as follows:

$$L = 1 - \frac{2 * \sum_{n=1}^{N} p_{pred} * p_{true}}{\sum_{n=1}^{N} p^2_{pred} + p^2_{true} + \epsilon} \quad (5)$$

where P_{pred} is the decoder output, P_{ture} is the segmentation mask. where summation is voxel-wise, and ϵ is a small constant to avoid zero division.The process of RDAU-Net implementation is as follows. The 3D brain tumor data of four modalities are convolved twice by two 3×3×3 layers to increase the number of features in the initial filters. Then, the extracted feature maps are input into the RA block for feature extraction. The CBAM is used to exploit useful information of the input features. The feature maps that are extracted by the first RA block are downsampled by a convolutional layer with a step size of 2. After four rounds of convolution and downsampling, the extracted feature maps are input into the DA block for feature extraction with various feature sizes. The DA block combines multiple dilated convolutional layers in parallel and incorporates a CBAM dual attention mechanism

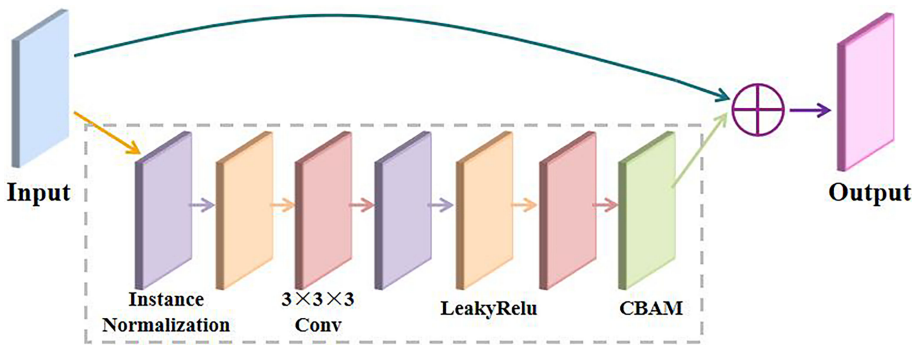


FIGURE 3 | Structure of an RA block.

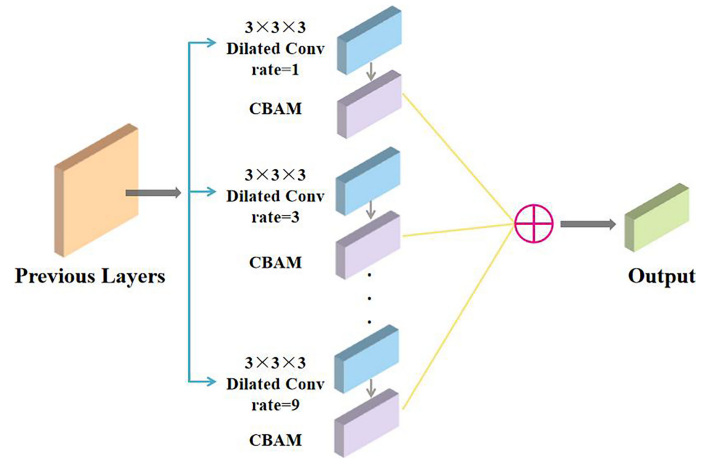


FIGURE 4 | Structure of a DA block.

behind each convolutional layer. This attention pyramid pooling module effectively obtains feature maps of various sizes through levels of expanded convolutional layers in the spatial pyramid model while extracting useful information on channels and spaces to increase the tumor segmentation accuracy. Then, the feature maps that are extracted from the DA block are upsampled by deconvolution. The upsampled feature maps are connected by a skip connection with the feature maps that have undergone feature weighting by the CBAM block in the corresponding layer in the encoder. Our proposed method inserts a CBAM block after each skip connection layer to improve the extraction of channel information and spatial information through tandem channel attention and spatial attention as a way to reduce the redundant information of low-level features. Attention is produced when the skip connection connects the feature maps that are extracted by the encoder directly to the corresponding layers of the decoder. After upsampling, the final prediction result of the network is output through the sigmoid function. Finally, an image of the same size as the input image is generated.

EXPERIMENTS

Evaluation Indicators

We use the Dice coefficient, Hausdorff distance, sensitivity, and specificity evaluation metrics to evaluate our experimental results.

The Dice coefficient (Dice) is defined as:

$$DSC = \frac{2 \times TP}{FP + FN + 2 \times TP} \tag{6}$$

Sensitivity and specificity are defined as:

$$Sensitivity = \frac{TP}{TP + FN} \tag{7}$$

$$Specificity = \frac{TP}{TP + FP} \tag{8}$$

where TP, FP, and FN indicate the true positive, false positive, and false negative values, respectively.

The Hausdorff distance indicates the maximum mismatch between the predicted edge of the tumor segmentation result and the ground-truth boundary.

$$HD(T,S) = \max\left\{\sup_{x\in X} \inf_{y\in Y} d(x,y), \sup_{y\in Y} \inf_{x\in X} d(x,y)\right\} \tag{9}$$

where sup and inf denote the upper and lower boundaries, respectively, of the brain tumor region; x and y are points on the tumor surface, where $x \in T$ and $y \in P$; and d() is the distance function.

Experimental Details

The experiments are carried out on a workstation that is configured with an Intel® Xeon(R) CPU E5-2620 v4 @ 2.10 GHz × 32 and equipped with two 12 GB TITAN Xp graphics cards. The proposed network is tested under the environment of TensorFlow-gpu==1.10.0[1] and Keras==2.2.0[2]. For training, the input preprocessed image has an image block size of 4 × 128 × 128 × 128. Since the computational complexity of the 3D image block is too high, we set the batch size to 2. The experiments are performed using the Adam optimization method, and the initial learning rate is set to $Lr_{init} = 1^{\cdot}10^{-4}$. The learning rate is reduced by 50% after 15 epochs if the validation loss is not improving, and we regularize with an 12 weight decay of 10^{-5}. The size of the output prediction segmentation result is also 128×128×128. Finally, we reshape the data to a size of 240 × 240 × 155 by using the Nilearn package[3].

Dataset

Our proposed method is validated on the BraTS 2018 and 2019 (33–36) datasets. The BraTS 2018 dataset contains 285 glioma cases, which correspond to 210 HGG patients and 75 LGG patients. The

[1] https://github.com/TensorFlow/TensorFlow

[2] https://github.com/keras-team/keras

[3] https://github.com/nilearn/nilearn

validation set contains 66 MRs of patients with unknown tumor grade without real labels. The BraTS 2019 dataset contains 335 glioma cases, which correspond to 259 HGG patients and 76 LGG patients. The validation set contains 125 MRs of patients with unknown tumor grade without real labels. The dataset contains data offour modalities, namely, T1, T2, TIC, and FLAIR, and each MRI is a 3D image of size 240×240×155. The task of the segmentation challenge was to segment three tumor subregions: 1) the whole tumor (WT), 2) the tumor core (TC), and 3) the enhancing tumor (ET). **Figure 5** shows the modalities for a case in the BraTS 2018 training dataset and the ground truth.

Data Preprocessing

MRI scans will often show intensity heterogeneity due to variations in the magnetic field. The variations in these mappings are called bias fields, and bias fields can cause problems for classifiers. We used N4ITK for bias field correction of the images, and N4ITK (36) is a modification of the N3 bias field correction method (37). Moreover, the images differ in terms of contrast among these four modes. Therefore, we normalize using the z-score method, namely, we subtract the intensity of each pixel from the average intensity of all pixels in each multimodal image and divide by the standard deviation, while GT is not normalized. Finally, we change the size from 240 × 240× 155 to 128 × 128 × 128 to reduce the number of parameters of the network.

Ablation Study

To determine whether the CBAM dual attention module is effective in enhancing the segmentation performance of the network, we perform ablation experiments.

These experiments are conducted with and without the use of the attention module. The scores for the four evaluation metrics on the datasets of the BraTS 2018 and 2019 challenges are obtained separately. As presented in **Table 1**, the Dice coefficient and Hausdorff distance of the network in which the dual attention mechanism is utilized improved across the board on both the BraTS 2018 and 2019 challenge datasets, especially the Dice score of WT, which improved by 1% on both datasets, and the Hausdorff distance of ET, which decreased by 0.2 mm and 0.11 mm, respectively. In addition, the sensitivity of the model in which the attention mechanism is utilized improved by 0.5% and 0.2%, respectively, on ET. In general, adding CBAMs to the network can effectively improve the performance of the network.

Histograms of various evaluation metrics on the datasets of the BraTS 2018 and 2019 challenges are presented in **Figure 6**. According to these results, inclusion of the CBAM in the network can effectively improve the segmentation accuracy of the model.

Figure 7 shows the visual segmentation results of our proposed approach on the BraTS 2018 challenge data training set. From **Figure 7**, we find that our network model can segment various regions of the tumor, especially the ET parts of the tumor, but there are segmentation errors in small places compared with the ground truth. In summary, by comparing our segmentation results with those of RDU-Net (without attention) and the ground truth, we find that our proposed model obtains satisfactory segmentation results with the RA block and the DA block.

Results on the BraTS Challenge Training Data

We compare the segmentation performance of our approach with those of other typical deep network methods, as shown in **Figure 8**, on the BraTS 2018 challenge training dataset. The original data in **Figure 8** from top to bottom are T1-c; the segmentation results of

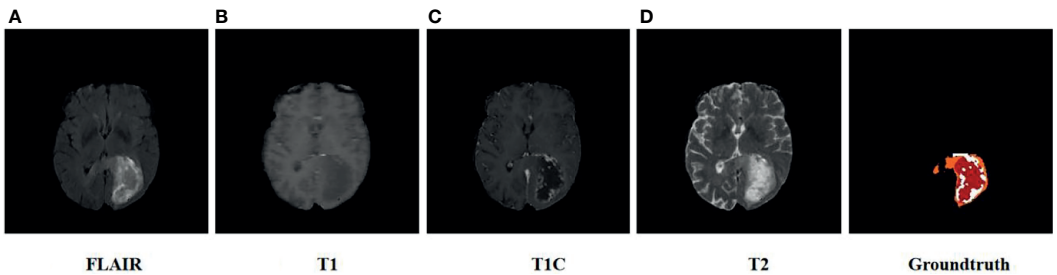


FIGURE 5 | Schematic diagram of the BRATS 2018 training dataset. From left to right are **(A)** FLAIR, **(B)** T1, **(C)** T1-c, and **(D)** ground truth, where the colors of the ground truth represent categories of tumor segmentation: orange represents edematous regions, red represents enhancing tumors, and white represents necrotic and nonenhancing tumors.

TABLE 1 | Segmentation results of our proposed network with and without attention mechanism on BraTS 2019 validation set using Dice, Hausdorff distance, specificity, and sensitivity metrics.

Models	Hausdorff (mm) ↓			Dice ↑			Specificity ↑			Sensitivity ↑		
	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET
RDU-Net (without CBAM)	6.9705	6.1778	3.0002	0.8895	0.9018	0.8222	0.9894	0.9969	0.9958	0.9169	0.9031	0.8714
RDAU-Net	6.8692	6.0015	2.9980	0.9001	0.9061	0.8311	0.9892	0.9975	0.9966	0.9172	0.9028	0.8730

Results are specified for WT, whole tumor; ET, enhancing tumor; and TC, tumor core.

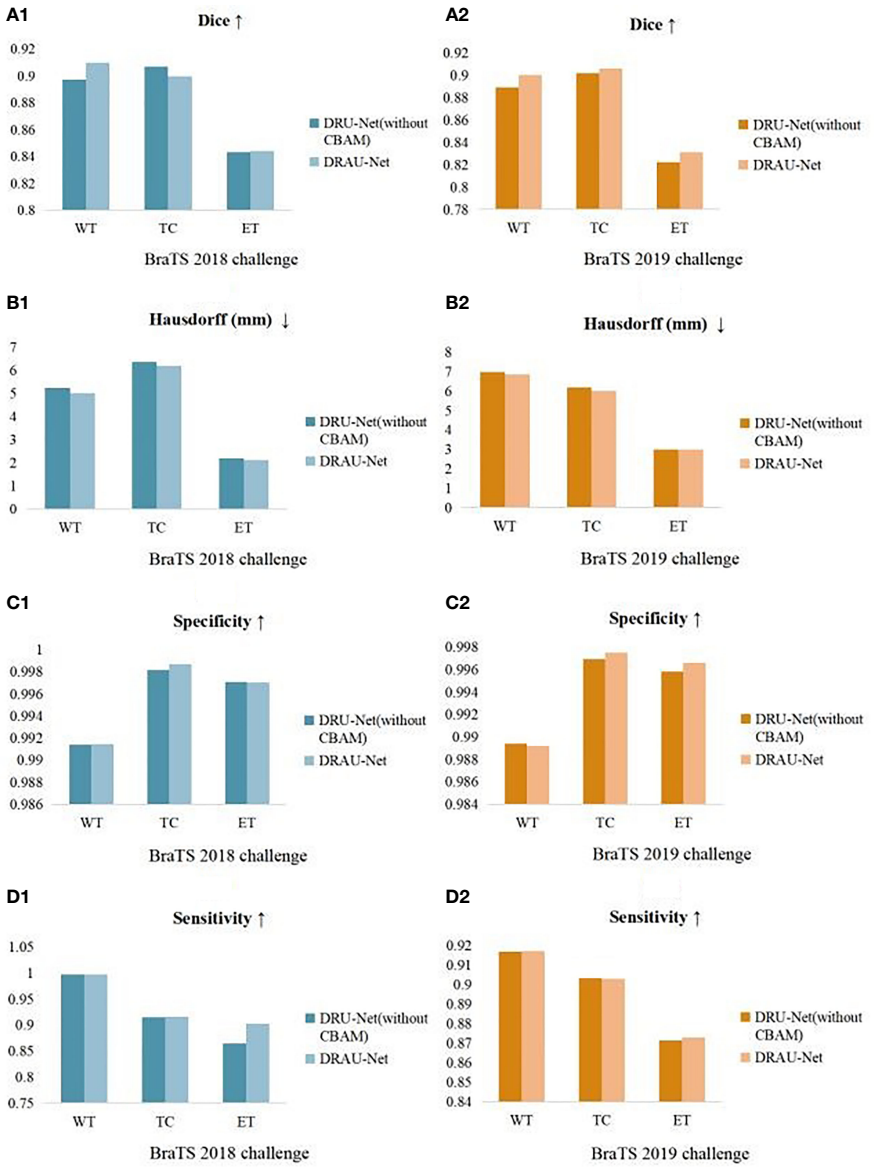


FIGURE 6 | Histograms of the considered evaluation metrics on the datasets of the BraTS 2018 and 2019 challenges. **(A–D)**: Dice, Hausdorff (mm), specificity, and sensitivity. Blue indicates the BraTS 2018 challenge dataset and orange indicates the BraTS 2019 challenge dataset.

the S3DU-Net (38) model, the AGResU-Net (39) model, and our RDAU-Net method; and the ground truths. By comparing these methods in **Figure 8**, we observe that the segmentation results of our proposed method, namely, RDAU-Net, are the closest to the ground truth on WT, TC, and ET and that the segmentation results are significantly better than those of the remaining two methods.

Comparison With State-of-the-Art Methods

Performance Comparison on the BraTS 2018 Challenge Dataset

The first-place winner of the Multimodal Brain Tumor Segmentation Challenge (BraTS) 2018 was Myronenko et al., who trained their model using large image blocks with a size of 160×192×128. As presented in **Table 2**, the Myronenko et al. (40) method has a higher overall Dice score than other mainstream methods, but our method outperforms Myronenko's method in terms of Dice score by 0.26%, 4.7%, and 3% for WT, TC, and ET, respectively. Our method also outperforms Myronenko's method in terms of Hausdorff distance, which is reduced by 0.7 mm and 1.7 mm for TC and ET, respectively. The No New-Net (41) method won second place with only a few minor changes to U-Net. As presented in **Table 2**, our method outperforms their method overall in segmentation. The Dice scores on TC and ET are 5% and 3.5% higher than theirs, respectively; the Hausdorff

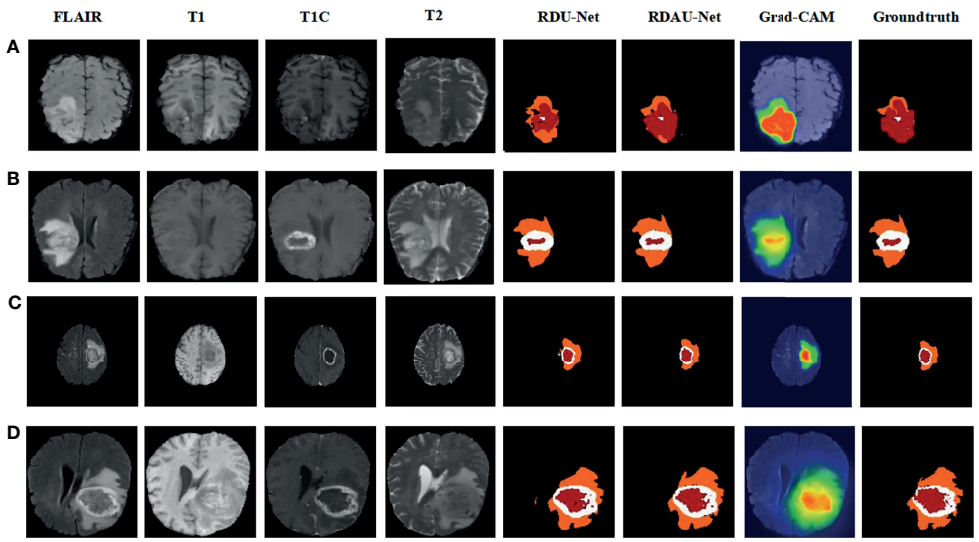


FIGURE 7 | Segmentation results of the proposed method and RDU-Net (without attention) on the BRATS 2018 datasets, the Grad-CAM result and the ground truths. Rows **(A–D)**: Brats18_TCIA01_387_1_93, Brats18_TCIA01_231_1_83, Brats18_CBICA_APR_1_105, Brats18_CBICA_AUN_1_81, where 105, 101, 78, and 73 represent the 105th, 101st, 78th, 73rd slices, respectively, of the MRI data. The colors represent the following categories: ■ edema, ■ necrosis, and □ enhancing core.

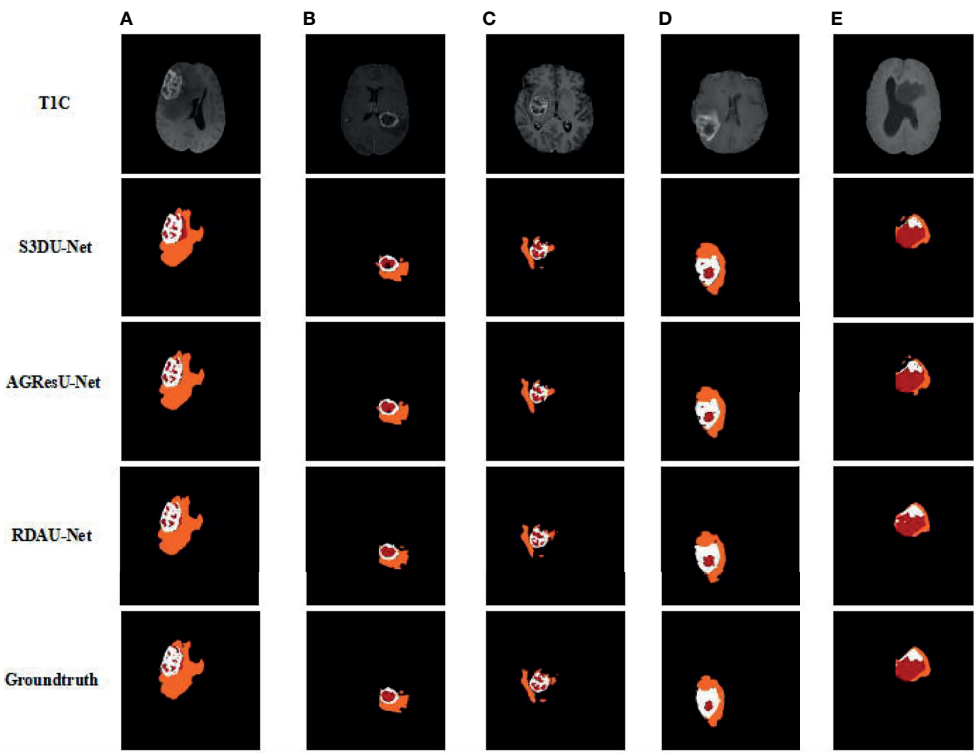


FIGURE 8 | Results of brain tumor segmentation in MRI images using the proposed approach and the ground truths. Columns **(A–E)** correspond to cases in the Brats 2018 challenge dataset: Brats18_2013_12_1_89, Brats18_CBICA_AQV_1_87, Brats18_TCIA02_117_1_69, Brats18_TCIA04_361_1_90, and Brats18_TCIA05_444_1_84, where 89, 87, 69, 90, and 84 represent the 89th, 87th, 69th, 90th, and 84th slices, respectively, of the MRI data. The colors represent the following categories: ■ edema, ■ necrosis, and □ enhancing core.

TABLE 2 | Comparison of performance between our approach and the state-of-the-art methods on BraTS 2018 validation set using Dice, Hausdorff distance, specificity, and sensitivity metrics.

Models	Hausdorff (mm) ↓			Dice ↑			Specificity ↑			Sensitivity ↑		
	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET
Myronenko (40)	4.41	6.84	3.82	0.9068	0.8602	0.8173	–	–	–	–	–	–
No New-Net (41)	4.97	7.04	2.54	0.9083	0.8544	0.8101	**0.9955**	0.9986	0.9979	0.9187	0.8444	0.8312
C-A-Net (42)	**4.1724**	6.5445	2.7162	**0.9095**	0.8651	0.8136	0.9951	0.9968	**0.9983**	0.9142	0.8683	0.8135
AGResU-Net (39)	5.62	8.36	3.57	0.872	0.808	0.772	–	–	–	–	–	–
S3DU-Net (38)	4.7165	7.7478	4.4321	0.8935	0.8309	0.7493	0.9927	0.9981	0.9976	**0.9290**	0.8161	0.7849
Our	5.0206	**6.1829**	**2.1260**	0.9094	**0.9071**	**0.8446**	0.9922	**0.9987**	0.9975	0.9161	**0.8699**	**0.8805**

Results are specified for WT, whole tumor; ET, enhancing tumor; and TC, tumor core. The results are marked in bold.

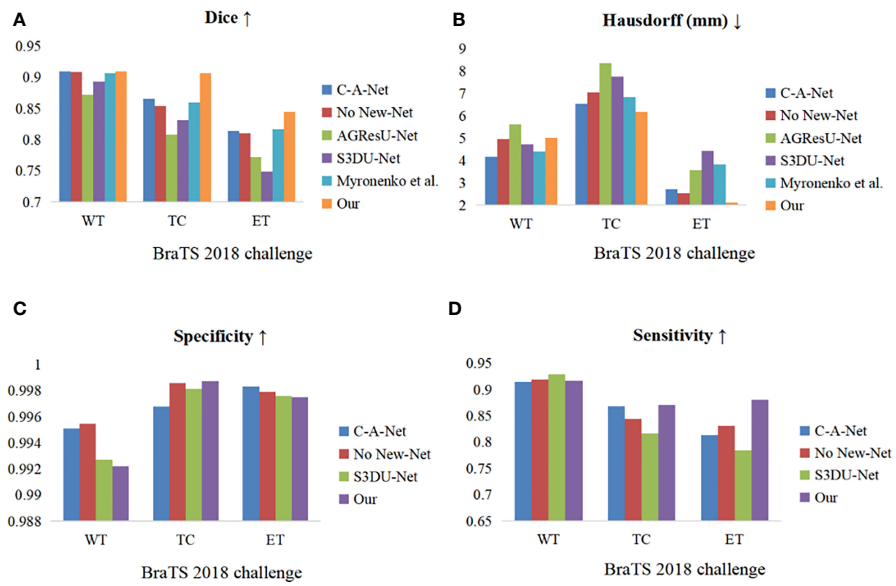


FIGURE 9 | Comparative histograms of evaluation indicators for four methods on the BraTS 2018 validation set. **(A–D)** Dice, Hausdorff (mm), specificity, and sensitivity. The colors correspond to the methods.

TABLE 3 | Comparison of performance between our approach and the state-of-the-art methods on BraTS 2019 validation set using Dice, Hausdorff distance, specificity, and sensitivity metrics.

Models	Hausdorff (mm) ↓			Dice ↑			Specificity ↑			Sensitivity ↑		
	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET
Two-stage cascaded U-Net (43)	**4.4438**	5.8620	3.2055	**0.9082**	0.8632	0.8020	0.9943	0.9974	0.9984	0.9237	0.8622	0.8038
Hamhanghala et al. (44)	6.9	8.4	4.6	0.8965	0.7901	0.7665	0.9939	0.9976	**0.9985**	0.9132	0.7771	0.7688
DDU-Nets46]	4.874	8.013	3.376	0.898	0.793	0.780	0.994	0.996	0.998	0.903	0.808	0.791
Myronenko et al. (46)	5.89	6.562	3.921	0.8900	0.8340	0.8000	–	–	–	–	–	–
3D U-Net (47)	7.357	**5.667**	5.994	0.807	0.894	0.737	**0.996**	0.995	0.998	0.826	0.897	0.766
Our	6.8692	6.0015	**2.9980**	0.9001	**0.9061**	**0.8311**	0.9892	**0.9975**	0.9966	**0.9172**	**0.9028**	**0.8730**

Results are specified for WT, whole tumor; ET, enhancing tumor; and TC, tumor core. The results are marked in bold.

distances on TC and ET are 0.9 mm and 0.4 mm shorter than theirs, respectively; and the sensitivity scores on TC and ET are 3% and 5% better than theirs, respectively. C-A-Net (42) is a single-channel multitask network that combines multiple CNN structures. According to **Table 2**, our network outperforms the C-A-Net method on all metrics except the Hausdorff distance score on WT, and the Dice scores on TC, and ET are improved by 5%, and 3.5%, respectively. Our sensitivity score on ET is nearly 5% higher than that of the C-A-Net method. The AGResU-Net method (39) integrates the residual module and

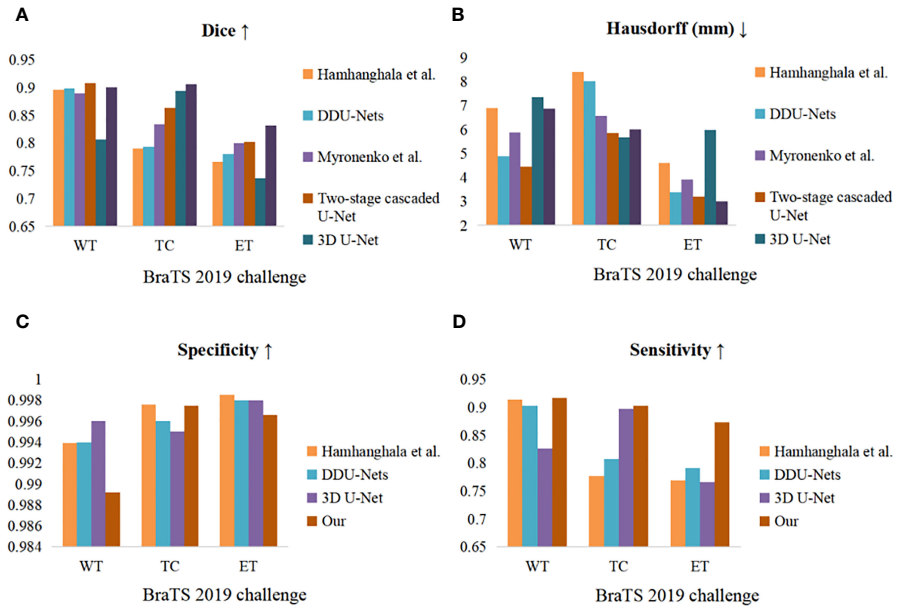


FIGURE 10 | Comparative histograms of evaluation indicators for four methods on the BraTS 2019 validation set. **(A–D)** Dice, Hausdorff (mm), specificity, and sensitivity. The colors correspond to the methods.

attention gates in the original U-Net. As presented in **Table 2**, our method outperforms the network on all metrics, especially the Hausdorff distance on TC: our method obtains a Hausdorff distance that is 2.1 mm shorter than that obtained by the AGResU-Net method, and our method obtains Dice scores that are improved by 9% and 7% on TC and ET, respectively. The S3DU-Net method (38) is a module in which the convolutional block in U-Net is changed to a 3D convolution with three parallel branches. It scores well on the sensitivity metric on WT, but our network outperforms that network overall. Compared with the S3DU-Net method, our method improves the Dice score by 7% and 9% on TC and ET, respectively; reduces the Hausdorff distance by 1.5 mm and 2.3 mm on TC and ET, respectively; and improves the sensitivity by 5% and 9% on TC and ET, respectively. The histograms in **Figure 9** compare several methods in terms of various metrics. According to the comparison in **Figure 9**, our method is highly competitive.

Performance Comparison on the BraTS 2019 Challenge Dataset

For the BraTS 2019 challenge dataset, RDAU-Net still achieved the best results in terms of the Hausdorff distance metric in the tumor core region and the Dice coefficients of TC and ET and performed very competitively in terms of other metrics. The two-stage cascaded U-Net (43) won first place in the BraTS 2019 challenge. Compared with the champion team, the Dice scores and the sensitivity scores on ET and TC of our method are almost 3% and 4.3% higher and 7% and 4% higher, respectively. The Hausdorff distance of the ET subregion that was obtained using our method is approximately 0.2 mm shorter. Hamhanghala et al. (44) used generative adversarial networks to expand the data. Although the training sample was expanded with synthetic data, this approach did not yield more prominent results. DDU-Nets (45) contain three models of distributed dense connectivity. DDU-Nets are not particularly effective overall, although they slightly outperform our method in terms of Hausdorff distance scores on WT and ET. Myronenko et al. (46) further improved the loss function based on 2018 by introducing the focal loss and using eight 32 G video cards for training. As presented in **Table 3**, our network still outperforms Myronenko et al.'s method on all other metrics, except for a slightly lower Hausdorff distance score on WT than their method. The 3D U-Net (47) is designed mainly to handle block diagrams. From **Table 3**, our method outperforms the 3D U-Net method on all metrics except for the Dice score on TC and the specificity scores on WT and ET, which are slightly lower than those of the 3D U-Net method. The histograms in **Figure 10** clearly compare the considered methods in terms of various metrics. According to the comparison, our method is still very competitive, even on the newer dataset.

CONCLUSIONS

We propose a new method, namely, RDAU-Net that is based on an improved ResU-Net for brain tumor segmentation in MRI. We add DA blocks to expand the receptive field and obtain image information of various sizes and insert a CBAM block after each skip connection layer to improve the extraction of channel information and spatial information to reduce the

redundant information of low-level features. By conducting experiments on the BraTS 2018 and BraTS 2019 datasets, we find that the use of an RA block instead of a convolutional layer, the inclusion of a DA block in the network, and the insertion of 3D CBAM blocks can effectively improve the performance of the network. RDAU-Net has more obvious advantages than the SOTA method. However, the performance of the method in the WT region still has substantial room for improvement, and we hope to solve this problem through postprocessing of the network. In conclusion, the method has greater advantages in segmenting subregions of brain tumors and can be effectively applied to clinical research.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://doi.org/10.7937/K9/TCIA.2017.KLXWJJ1Q.

AUTHOR CONTRIBUTIONS

JW provided guidance on the content of the article as well as the writing of the article. ZY writes articles and conducts experiments. ZL and JR completed the work of collecting references. YZ provided guidance on medical knowledge and medical image knowledge. GY provided guidance on the experiments done in the article. All authors are approved for publication.

FUNDING

This work was jointly supported by the Shandong University Science and technology project (No.J17KA0 82, No. J16LN21), the National Natural Science Foundation of China (No.61401259), the China Postdoctoral Science Foundation (No.2015M582128).

REFERENCES

1. Siegel R, Naishadham D, Jemal A. Cancer Statistics, 2013. *CA Cancer J Clin* (2013) 63(1):11–30. doi: 10.3322/caac.21166
2. Yang D. Standardized MRI Assessment of High-Grade Glioma Response: A Review of the Essential Elements and Pitfalls of the RANO Criteria. *Neuro-Oncology Pract* (2014) 3(1):59–67. doi: 10.1093/nop/npv023
3. Reznek RH. CT/MRI of Neuroendocrine Tumours. *Cancer Imaging* (2006) 6:163–77. doi: 10.1102/1470-7330.2006.9037
4. Bauer S, Wiest R, Nolte L, Reyes M. A Survey of MRI-Based Medical Image Analysis for Brain Tumor Studies. *Phys Med Biol* (2013) 58(13):97–129. doi: 10.1088/0031-9155/58/13/R97
5. Pereira S, Pinto A, Alves V, Silva C. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE Trans Med Imaging* (2016) 35(5):1240–51. doi: 10.1109/TMI.2016.2538465
6. Nie D, Wang L, Adeli E, Lao C, Lin W, Shen D. 3-D Fully Convolutional Networks for Multimodal Isointense Infant Brain Image Segmentation. *IEEE Trans Cybern* (2019) 49(3):1123–36. doi: 10.1109/TCYB.2018.2797905
7. Chen S, Ding C, Liu M. Dual-Force Convolutional Neural Networks for Accurate Brain Tumor Segmentation. *Pattern Recognit* (2019) 88:90–100. doi: 10.1016/j.patcog.2018.11.009
8. Zhuge Y, >Ning H, Mathen P, Cheng J, Krauze A, Camphausen K, et al. Automated Glioma Grading on Conventional MRI Images Using Deep Convolutional Neural Networks. *Med Phys* (2020) 47(7):3044–53. doi: 10.1002/mp.14168
9. Zhong Z, Fan B, Duan J, Wang L, Ding K, Xiang S, et al. Discriminant Tensor Spectral-Spatial Feature Extraction for Hyperspectral Image Classification. *IEEE Geosci Remote Sens Lett* (2015) 12(5):1028–32. doi: 10.1109/LGRS.2014.2375188
10. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In: *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc. San Diego, CA, USA: arXiv* (2015). p. 1–14.
11. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification With Deep Convolutional Neural Networks. *Commun ACM* (2017) 60(6):84–90. doi: 10.1145/3065386
12. Srigurulekha K, Ramachandran V. Food Image Recognition Using CNN. In: *2020 Int. Conf. Comput. Commun. Informatics, ICCCI 2020* (2020). Coimbatore, India: IEEE. doi: 10.1109/ICCCI48352.2020.9104078
13. Qamar S, Jin H, Zheng R, Ahmad P. 3d Hyper-Dense Connected Convolutional Neural Network for Brain Tumor Segmentation. In: *Proc. - 2018 14th Int. Conf. Semant. Knowl. Grids, SKG 2018*. Guangzhou, China: IEEE (2018). p. 123–30. doi: 10.1109/SKG.2018.00024
14. Wang G, Zuluaga M, Li W, Pratt R, Patel P, Aertsen M, et al. DeepIGeoS: A Deep Interactive Geodesic Framework for Medical Image Segmentation. *IEEE Trans Pattern Anal Mach Intell* (2019) 41(7):1559–72. doi: 10.1109/TPAMI.2018.2840695
15. Dolz J, Gopinath K, Yuan J, Lombaert H, Desrosiers C, Ben A. HyperDense-Net: A Hyper-Densely Connected CNN for Multi-Modal Image Segmentation. *IEEE Trans Med Imaging* (2019) 38(5):1116–26. doi: 10.1109/TMI.2018.2878669
16. Wang W, Gang J. Application of Convolutional Neural Network in Natural Language Processing. In: *Proc. 2018 Int. Conf. Inf. Syst. Comput. Aided Educ. ICISCAE 2018*. Changchun, China: IEEE. (2019). p. 64–70. doi: 10.1109/ICISCAE.2018.8666928
17. Jin N, Wu J, Ma X, Yan K, Mo Y. Multi-Task Learning Model Based on Multi-Scale CNN and LSTM for Sentiment Classification. *IEEE Access* (2020) 8:77060–72. doi: 10.1109/ACCESS.2020.2989428
18. Long J, Shelhamer E, Darrell T, Berkeley U. Fully Convolutional Networks for Semantic Segmentation. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, America: IEEE (2015). p. 640–51. doi: 10.1109/CVPR.2015.7298965
19. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *Med. Image Comput. Comput. Interv. , MICCAI*. Munich, Germany: arXiv. (2015). p. 234–41. doi: 10.1007/978-3-319-24574-4_28
20. Oktay O, Schlemper J, Le Folgoc L, Lee M, Heinrich M, Misawa K, et al. Attention U-Net: Learning Where to Look for the Pancreas. *arXiv(Midl)* (2018) 1–10.
21. Shankaranarayana SM, Ram K, Mitra K, Sivaprakasam M. Joint Optic Disc and Cup Segmentation Using Fully Convolutional and Adversarial Networks. *Lect Notes Comput Sci* (2017) 10554 LNCS:168–76. doi: 10.1007/978-3-319-67561-9_19
22. Bahdanau D, Cho KH, Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate. In: *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. San Diego, CA, USA: arXiv. (2015). p. 1–15.
23. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is All You Need. *Adv Neural Inf Process Syst* (2017) 5999–6009.
24. Galassi A, Lippi M, Torroni P. Attention in Natural Language Processing. *IEEE Trans Neural Networks Learn Syst* (2020) 32(10):4291–308. doi: 10.1109/tnnls.2020.3019893
25. Jiang L, Sun X, Mercaldo F, Santone A. DECAB-LSTM: Deep Contextualized Attentional Bidirectional LSTM for Cancer Hallmark Classification. *Knowledge-Based Syst* (2020) 210:106486–95. doi: 10.1016/j.knosys.2020.106486

26. de S. Correia A, Colombini EL. *Attention, Please! A Survey of Neural Attention Models in Deep Learning. arXiv* (2021).
27. Niu Z, Zhong G, Yu H. A Review on the Attention Mechanism of Deep Learning. *Neurocomputing* (2021) 452:48–62. doi: 10.1016/j.neucom.2021.03.091
28. Shao Y, Lan J, Liang Y, Hu J. Residual Networks With Multi-Attention Mechanism for Hyperspectral Image Classification. *Arab J Geosci* (2021) 14 (4):252–70. doi: 10.1007/s12517-021-06516-6
29. Hu J, Shen L, Sun G. Squeeze-And-Excitation Networks. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Salt Lake City, Utah: IEEE (2018). p. 7132–41. doi: 10.1109/CVPR.2018.00745
30. Wang Q, Wu B, Zhu P, Li P, Zuo W, Hu Q. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Seattle, Washington: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) IEEE (2020). p. 11531–9. doi: 10.1109/CVPR42600.2020.01155
31. Fu J, Liu J, Tian H, Li Y, Bao Y, Fang Z, et al. Dual Attention Network for Scene Segmentation. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Long Beach, CA: IEEE (2019). p. 3141–9. doi: 10.1109/CVPR.2019.00326
32. Woo S, Park J, Lee J, Kweon I. CBAM: Convolutional Block Attention Module. *arXiv* (2018) 1:3–19. doi: 10.1007/978-3-030-01234-2
33. Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby J, et al. Advancing The Cancer Genome Atlas Glioma MRI Collections With Expert Segmentation Labels and Radiomic Features. *Sci Data* (2017) 4(March):1–13. doi: 10.1038/sdata.2017.117
34. Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M, Crimi A, et al. Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge. *arXiv* (2018) 1–49.
35. Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby JS, et al. Segmentation Labels and Radiomic Features for the Pre-Operative Scans of the TCGA-GBM Collection. *Cancer Imaging Arch* (2017).
36. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK: Improved N3 Bias Correction. *IEEE Trans Med Imaging* (2010) 29 (6):1310–20. doi: 10.1109/TMI.2010.2046908
37. Larsen CT, Eugenio Iglesias J, Van Leemput K. N3 Bias Field Correction Explained as a Bayesian Modeling Method. *Lect Notes Comput Sci* (2014) 8677:1–12. doi: 10.1007/978-3-319-12289-2_1
38. Zhang J, Jiang Z, Dong J, Hou Y, Liu B. Attention Gate ResU-Net for Automatic MRI Brain Tumor Segmentation. *IEEE Access* (2020) 8:58533–45. doi: 10.1109/ACCESS.2020.2983075
39. Chen W, Liu B, Peng S, Sun J, Qiao X. *S3D-UNet: Separable 3d U-Net for Brain Tumor Segmentation Wei*. Granada, Spain: Springer International Publishing (2018). doi: 10.1007/978-3-030-11726-9_32
40. Myronenko A. 3d MRI Brain Tumor Segmentation Using Autoencoder Regularization. *Int MICCAI Brainlesion Workshop* (2018) 11384:311–20. doi: 10.1007/978-3-030-11726-9
41. Isensee F, Kickingereder P, Wick W, Bendszus M, Maier-Hein K. No New-Net. *Lect Notes Comput Sci* (2019) 11384:234–44. doi: 10.1007/978-3-030-11726-9_21
42. Zhou C, Chen S, Ding C, Tao D. Learning Contextual and Attentive Information for Brain Tumor Segmentation, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer International Publishing (2019) 11384:497–507. doi: 10.1007/978-3-030-11726-9
43. Jiang Z, Ding C, Liu M, Tao D. Two-Stage Cascaded U-Net: 1st Place Solution to Brats Challenge 2019 Segmentation Task. *Lecture Notes Comput Sci* (2020) 11992:231–41. doi: 10.1007/978-3-030-46640-4_22
44. Hamghalam M, Lei B, Wang T. Brain Tumor Synthetic Segmentation in 3d Multimodal MRI Scans. *Lect Notes Comput Sci* (2020) 11992:153–62. doi: 10.1007/978-3-030-46640-4_15
45. Zhang H, Li J, Shen M, Wang Y, Yang G. DDU-Nets: Distributed Dense Model for 3D MRI Brain Tumor Segmentation. *Lect Notes Comput Sci* (2020) 11993:208–17. doi: 10.1007/978-3-030-46643-5_20
46. Myronenko A, Hatamizadeh A. Robust Semantic Segmentation of Brain Tumor Regions From 3D MRIs. *Lect Notes Comput Sci* (2020) 11993:82–9. doi: 10.1007/978-3-030-46643-5_8
47. Wang F, Jiang R, Zheng L, Meng C, Biswai B. 3d U-Net Based Brain Tumor Segmentation and Survival Days Prediction. *Lecture Notes Comput Sci* (2020) 11992:131–41. doi: 10.1007/978-3-030-46640-4_13

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wang, Yu, Luan, Ren, Zhao and Yu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 10 March 2022
doi: 10.3389/fonc.2022.830910

Comprehensive Risk System Based on Shear Wave Elastography and BI-RADS Categories in Assessing Axillary Lymph Node Metastasis of Invasive Breast Cancer—A Multicenter Study

Huiting Zhang[1†], Yijie Dong[1†], Xiaohong Jia[1], Jingwen Zhang[1], Zhiyao Li[2], Zhirui Chuan[2], Yanjun Xu[3], Bin Hu[4], Yunxia Huang[5], Cai Chang[5], Jinfeng Xu[6], Fajin Dong[6], Xiaona Xia[7], Chengrong Wu[7], Wenjia Hu[8], Gang Wu[8], Qiaoying Li[9], Qin Chen[10], Wanyue Deng[10], Qiongchao Jiang[11], Yonglin Mou[12], Huannan Yan[12], Xiaojing Xu[13], Hongju Yan[13], Ping Zhou[14], Yang Shao[14], Ligang Cui[15], Ping He[15], Linxue Qian[16], Jinping Liu[16], Liying Shi[17], Yanan Zhao[18], Yongyuan Xu[18], Yanyan Song[19], Weiwei Zhan[1*] and Jianqiao Zhou[1*]*

[1] Department of Ultrasound, Ruijin Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China, [2] Department of Medical Ultrasound, Yunnan Cancer Hospital & The Third Affiliated Hospital of Kunming Medical University, Kunming, China, [3] Department of Ultrasound in Medicine, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai Institute of Ultrasound in Medicine, Shanghai, China, [4] Department of Ultrasound, Minhang Hospital, Fudan University, Shanghai, China, [5] Department of Ultrasonography, Fudan University Shanghai Cancer Center, Shanghai Medical College, Fudan University, Shanghai, China, [6] Department of Ultrasound, Shenzhen People's Hospital, The Second Clinical Medical College, Jinan University, and The First Affiliated Hospital, Southern University of Science and Technology, Shenzhen, China, [7] Department of Ultrasound Medicine, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China, [8] Department of Ultrasound, People's Hospital of Henan Province, Zhengzhou, China, [9] Department of Ultrasound Diseases, Tangdu Hospital, Four Military Medical University, Xi'an, China, [10] Department of Ultrasound, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, China, [11] Department of Ultrasound, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China, [12] Department of Ultrasound, General Hospital of Northern Theater Command, Shenyang, China, [13] Department of Ultrasound, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou, China, [14] Department of Ultrasound, The Third Xiangya Hospital of Central South University, Changsha, China, [15] Department of Ultrasound, Peking University Third Hospital, Beijing, China, [16] Department of Ultrasound, Beijing Friendship Hospital, Capital Medical University, Beijing, China, [17] Department of Ultrasound, Affiliated Hospital of Guizhou Medical University, Guizhou, China, [18] Department of Ultrasound, Second Affiliated Hospital of Zhejiang University, School of Medicine, Hangzhou, China, [19] Department of Biostatistics, Institute of Medical Sciences, Shanghai Jiaotong University School of Medicine, Shanghai, China*

Purpose: To develop a risk stratification system that can predict axillary lymph node (LN) metastasis in invasive breast cancer based on the combination of shear wave elastography (SWE) and conventional ultrasound.

Materials and Methods: A total of 619 participants pathologically diagnosed with invasive breast cancer underwent breast ultrasound examinations were recruited from a multicenter of 17 hospitals in China from August 2016 to August 2017. Conventional ultrasound and SWE features were compared between positive and negative LN metastasis groups. The regression equation, the weighting, and the counting methods

OPEN ACCESS

Edited by:
Kyung Hyun Sung, UCLA Health System, United States

Reviewed by:
Kaustav Bera, Maimonides Medical Center, United States
Wenwu Ling, Sichuan University, China

***Correspondence:**
Jianqiao Zhou
zhousu30@126.com
Yanyan Song
syyssmu@126.com
Weiwei Zhan
shanghairuijinus@163.com

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *07 December 2021*
Accepted: *14 February 2022*
Published: *10 March 2022*

Citation:
Zhang H, Dong Y, Jia X, Zhang J, Li Z, Chuan Z, Xu Y, Hu B, Huang Y, Chang C, Xu J, Dong F, Xia X, Wu C, Hu W, Wu G, Li Q, Chen Q, Deng W, Jiang Q, Mou Y, Yan H, Xu X, Yan H, Zhou P, Shao Y, Cui L, He P, Qian L, Liu J, Shi L, Zhao Y, Xu Y, Song Y, Zhan W and Zhou J (2022) Comprehensive Risk System Based on Shear Wave Elastography and BI-RADS Categories in Assessing Axillary Lymph Node Metastasis of Invasive Breast Cancer—A Multicenter Study. Front. Oncol. 12:830910. doi: 10.3389/fonc.2022.830910

were used to predict axillary LN metastasis. The sensitivity, specificity, and the areas under the receiver operating characteristic curve (AUC) were calculated.

Results: A significant difference was found in the Breast Imaging Reporting and Data System (BI-RADS) category, the "stiff rim" sign, minimum elastic modulus of the internal tumor and peritumor region of 3 mm between positive and negative LN groups ($p < 0.05$ for all). There was no significant difference in the diagnostic performance of the regression equation, the weighting, and the counting methods ($p > 0.05$ for all). Using the counting method, a 0–4 grade risk stratification system based on the four characteristics was established, which yielded an AUC of 0.656 (95% CI, 0.617–0.693, $p < 0.001$), a sensitivity of 54.60% (95% CI, 46.9%–62.1%), and a specificity of 68.99% (95% CI, 64.5%–73.3%) in predicting axillary LN metastasis.

Conclusion: A 0–4 grade risk stratification system was developed based on SWE characteristics and BI-RADS categories, and this system has the potential to predict axillary LN metastases in invasive breast cancer.

Keywords: breast neoplasms, lymphatic metastasis, ultrasonography, elasticity imaging techniques, risk assessment

INTRODUCTION

Female breast cancer, the first cause of the death in malignant tumors (1), is the most common neoplasm in 20–59-year-old women; early diagnosis can reduce 40% of deaths (2–5). According to the American Joint Committee on Cancer, tumor size (T), node (N), and distant metastasis (M) have served as the global standard for conveying disease status among clinicians (6). Sentinel lymph node biopsy (SLNB) is regarded as the standard method for identifying the staging and determining the clinical arrangement. However, SLNB was recommended only for T1 or T2 tumors (7, 8), and multiple factors contribute to its limited pathological findings, suggesting that tedious and time-consuming SLNB is not optimal (9). Besides, SLNB may cause complications like arm numbness or upper limb edema in 3.5%–10.9% patients because of the increasing anesthesia time (10). Thus, an accurate and non-invasive examination for evaluating axillary LN status is expected before surgery or clinical treatment. However, showing a great difference among various studies, the sensitivity and specificity of axillary ultrasound in detection of LN metastasis ranged from 45.2% to 92.7% and from 40.5% to 93.9%, respectively (11–14).

For breast ultrasound examinations, as the suspicious imaging features increase, the Breast Imaging Reporting and Data System (BI-RADS) categories of the breast tumor increase accordingly. Tumors in BI-RADS category 4 or 5 indicate a risk of malignancy, and a detailed assessment of the SLN status is required (15). Shear wave elastography (SWE) is a novel ultrasonography based on the speed of shear wave velocity in different tissues, which can provide quantitative (that is kPa or m/s) and qualitative information (such as "stiff rim" sign) in differentiation of breast tumors (10, 12, 16). It has been reported that LN metastasis of breast cancer was associated with breast lesion elastic modulus (17). Besides, the "stiff rim" sign, usually appearing in malignant breast lesions with increased peritumoral stiffness, was regarded as the infiltration of cancer cells into the interstitial tissues or a desmoplastic reaction (16), which is an independent prognostic factor predicting tumor recurrence and patient death (18, 19). Thus, it is reasonable to hypothesize that the "stiff rim" sign may be potentially associated with LN metastasis. However, to our best knowledge, no study has detected the relationship between SWE information of peritumor tissues and axillary LN metastasis.

Hence, our study aims at assessing the risk of breast cancer LN metastasis in ultrasound examination by combining quantitative and qualitative SWE features, as well as BI-RADS categories.

MATERIALS AND METHODS

Patients

This prospective study was approved by the institutional ethics committee, and written informed consent to participate was acquired before examinations. From August 2016 to August 2017, there were a total of 689 patients with malignant breast tumors from a multicenter consisting of 17 hospitals in China diagnosed by surgical pathology enrollment, with conventional ultrasound and SWE images acquired before treatment. The pathology of LNs from SLNB or axillary LN dissection was acquired. All macrometastases, micrometastases, and isolated tumor cells were counted as node positive. The exclusion criteria were set as follows: (1) patients with non-invasive malignant tumors according to the pathology; (2) patients who accepted chemotherapy before surgery; and (3) patients who accepted chemotherapy or biopsy before ultrasound examination. In total, we have excluded 70 patients, including 34 ductal carcinoma *in situ*, 11 solid papillary carcinoma, 6 intraductal papillary

carcinoma, 2 cases with neoadjuvant chemotherapy before surgery, 6 cases of secondary tumor, and 11 cases without LN pathology. According to the rules established by the study, only one lesion was evaluated per patient, i.e., the most suspicious lesion in ultrasound examination or the largest one among the same BI-RADS category was selected in patients with multiple masses. Finally, our study included 619 invasive breast cancer participants (age range, 22–91 y; mean age, 52.16 + 11.21 y), including 1 male and 618 female patients.

SWE Image Acquisition and Analysis

All the ultrasound examinations of patients from 17 hospitals used the Resona 7 ultrasound system (Mindray Medical International, Shenzhen, China) equipped with an L11-3 high-frequency probe. Prior to collecting data, all participating radiologists received systematic training in conventional ultrasound and SWE breast examination. Moreover, all radiologists from the multicenter had more than 3 years of experience in breast ultrasound elastography. Following standard conventional ultrasound examination, SWE was performed according to the already well-established method (20). The SWE display scale was set at 140 kPa according to the protocol provided by the manufacturer.

An online measurement of the tumor size and SWE parameters was done instantly by the examining radiologist. In measuring SWE parameters, the outline of the lesion was drawn by tracing, and the function of the "shell" was equipped to acquire the elastic modulus of the peritumoral tissues in regions of 1, 2, and 3 mm outside the boundary of internal tumors. Values of mean elasticity (E_{mean}), maximum elasticity (E_{max}), minimum elastic modulus (E_{min}), standard deviation (E_{sd}) evaluating both internal tumors and peritumoral tissues as well as intratumor ratio between mean elastic modulus of breast lesions, and normal fatty tissue (E_{ratio}) were acquired and recorded. In assessing the location of the breast tumor, it is based on the location of the center of the lump in the 4 quadrants of the breast (upper/lower outer/inner quadrants). Upon completion of the evaluation, both B-mode images and SWE images of each patient were stored in a hard disk of the ultrasound system and subsequently sent to the study center (**Figure 1**).

At the study center, three radiologists with more than 8 years of experience in breast ultrasound evaluated ultrasound images according to consensus principles and proposed BI-RADS classification based on conventional ultrasound. At the same time, the presence of the "stiff rim" sign of the tumor was determined based on the SWE images. SWE images are colored in blue, green, orange, and red to sequentially show the progressive increase in tissue stiffness. When compared with the stiffness of the surrounding normal breast tissues and the interior tumor tissues, the peritumoral region, up to approximately 3 mm outside the boundary of internal tumors, showed increased stiffness (coded in orange or red), then it is judged as positive for the "stiff rim" sign (16).

Statistical Analysis

Statistical analyses were performed using SPSS Version 23.0 (IBM, Armonk, NY, USA) and SAS 9.4(Statistical Analysis System, v.9.4). A p value less than 0.05 was considered for significant differences. The Mann–Whitney U test was used to compare the statistical difference in continuous variables between positive and negative LN groups, and the chi-square test was used to verify whether there was a statistical difference in categorical variables between the two groups. Significant

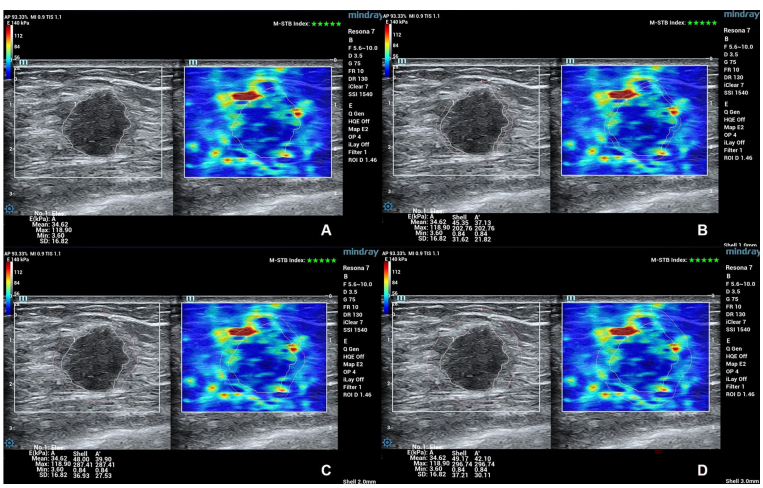


FIGURE 1 | B-mode and SWE images of a patient with invasive breast tumor. The "stiff rim" sign presents on the SWE images. **(A)** SWE images and parameters of outlined tumor. The E_{max}, E_{mean}, E_{min}, and E_{sd} of internal tumor were 34.63, 118.90, 3.60, and 16.82 kPa. **(B)** SWE images and parameters of shell 1 mm. The E_{max}, E_{mean}, E_{min}, and E_{sd} of shell 1 mm were 45.35, 202.76, 0.84, and 31.62 kPa. **(C)** SWE images and parameters of shell 2 mm. The E_{max}, E_{mean}, E_{min}, and E_{sd} of shell 2 mm were 48.00, 287.41, 0.84, and 36.93 kPa. **(D)** SWE images and parameters of shell 3 mm. The E_{max}, E_{mean}, E_{min}, and E_{sd} of shell 3 mm were 49.17, 296.74, 0.84, and 37.21 kPa.

variables were included in the final logistic regression analysis. The selection criteria alpha was set to 0.05. The odds ratio (OR) and 95% confidence interval (CI) for each feature were obtained from the logistic regression model. To predict axillary LN metastasis, the regression equation, the weighting, and the counting methods were used. The OR for the BI-RADS category and each SWE feature from the final logistic regression model was used in the weighting and counting methods. In the weighting method, the ORs were standardized using rounding to ensure appropriate application. In the counting method, the number of features with OR >1 was counted (21). The area under the receiver operating characteristic curve (AUC), sensitivity, specificity, and 95% CI of different variables were calculated, with histologic diagnosis as the gold standard. Finally, the risk system was established by balancing the diagnostic performance and ease of use. Multivariate logistic analysis was used to evaluate the OR of each risk grade. The Cochran–Armitage test was used to determine the change in the probability of axillary LN metastasis with the risk grade increased.

RESULTS

Pathologic Results

A total of 619 invasive breast cancers included 508 cases of invasive ductal carcinoma and 111 cases of other types (**Table 1**). Among them, 174 (28.1%) patients were diagnosed with positive LN metastasis and 445 (71.9%) were negative. The mean size of the tumor was 20.99 ± 9.04 mm (range 4.7–89 mm).

Clinical and BI-RADS Characteristics

Features of categorical variables including staging of tumor, tumor location, and tumor BI-RADS categories were compared between negative and positive LN groups, as shown in **Table 2**. There was a significant difference in constituent ratio between 5 BI-RADS categories ($p < 0.001$). In the 619 masses, 481 lesions were defined as BI-RADS 3-4C, with about 25% (ranging from 22.4% to 27.3%) cases in positive metastasis LN, while in 138 masses of BI-RADS 5, patients with positive LN metastasis were up to 43.5%. There was no significant difference in tumor staging ($p = 0.340$) and location ($p = 0.390$) in the two groups.

TABLE 1 | Pathologic diagnosis in 619 patients.

Histological type	Number
Invasive ductal carcinoma	508
Invasive lobular carcinoma	70
Mucinous breast carcinoma	9
Invasive ductal carcinoma with neuroendocrine differentiation	2
Metaplastic breast carcinoma	2
Invasive adenocarcinoma of breast	1
Invasive apocrine carcinoma	1
Tubular carcinoma	1
Others	
Intraductal carcinoma associated with microinvasive carcinoma	18
Invasive breast carcinoma with micropapillary features	5
Intracystic papillary breast carcinoma with areas of infiltration	1
Intraductal carcinoma associated with metaplastic squamous cell carcinoma	1

Qualitative SWE Characteristics

For the qualitative SWE feature, i.e., the "stiff rim" sign, there was a significant difference between the positive and negative groups for lymph node metastasis ($p < 0.001$). In patients with negative LN metastasis, 68.3% (304/445) cases presented a "stiff rim" sign in SWE images, while in positive LN groups, up to 82.2% (143/174) cases presented a "stiff rim" sign, significantly higher than that of negative LN groups.

Quantitative SWE Characteristics

The characteristics of continuous SWE parameters in internal tumors and peritumor tissues between the two groups are shown in **Table 3**. The E_{mean}, E_{max}, E_{min}, E_{sd}, and E_{ratio} of internal tumors in positive LN groups were similar to that of negative LN groups ($p > 0.05$ for all). The E_{mean}, E_{max}, and E_{sd} of shell 1, 2, and 3 mm as well as E_{min} of shell 1 and 2 mm showed no significant difference between the two groups ($p > 0.05$ for all). However, the E_{min} of internal tumor ($p = 0.005$) and shell 3 mm ($p = 0.007$) was significantly lower in positive LN groups than that in the negative group. The AUC, sensitivity, specificity, 95% CI, and cutoff values are shown in **Table 4**. Only E_{min} of internal tumor ($p < 0.01$) and shell 3 mm ($p < 0.01$) could predict axillary LN metastasis, with cutoff values of 3.31 and 3.52, respectively.

Risk Grade of LN Metastasis and Its Diagnostic Performance

The multiple logistic regression analysis indicated that the following four features were significantly associated with LN metastasis: (1) the breast tumor was assessed as BI-RADS 5; (2) the E_{min} of intrinsic tumor was no more than 3.31 kPa; (3) the E_{min} of shell 3 mm was no more than 3.52 kPa; and (4) the tumor showed a positive "stiff rim" sign in SWE. The OR values of the four features were 2.155, 1.654, 1.564, and 1.900, respectively ($p < 0.05$ for all). In the weighting method, the weight values of all four features were 2 based on the OR values of the above four features (**Table 5**). In the prediction of LN metastasis, the logistic regression equation yielded the highest AUC of 0.659 (95% CI, 0.621–0.697), while the weighting method and the counting method had the same AUC of 0.656 (95% CI, 0.617–0.693). There was no statistical difference among the three ($p > 0.05$ for all) (**Table 6** and **Figure 2**).

Following the principle of balancing the diagnostic performance and ease of use, the counting method was used to establish the risk system for predicting axillary LN metastases. Thus, a 0–4 grade risk stratification system was established based on the four suspicious features mentioned above: The LN metastasis risk grade of the breast cancer is defined from 0 to 4 according to the number of suspicious features, i.e., the mass was judged to be risk grade 4 when all suspicious features were present and grade 0 when no suspicious features were present. The risk grade of LN metastasis for the 619 patients was determined according to the above rule.

Subsequently, we compared the difference of risk grade between patients with positive and negative metastasis LNs

TABLE 2 | Clinical and ultrasound characteristics between negative and positive LN groups.

Parameters	Negative LN (n = 445)	Positive LN (n = 174)	Total (n = 619)	*p* value
Tumor staging				0.340
T1	244 (73.5%)	88 (26.5%)	332	
T2 and T3	201 (70.0%)	86 (30.0%)	287	
Tumor location				0.390
Upper outer quadrant	263 (72.3%)	101 (27.7%)	364	
Upper inner quadrant	106 (75.7%)	34 (24.3%)	140	
Lower outer quadrant	55 (65.5%)	29 (34.5%)	84	
Lower inner quadrant	21 (67.7%)	10 (32.3%)	31	
BI-RADS categories				$p < 0.001$
3	8 (72.7%)	3 (27.3%)	11	
4A	57 (75.0%)	19 (25.0%)	76	
4B	111 (75.0%)	37 (25.0%)	148	
4C	191 (77.6%)	55 (22.4%)	246	
5	78 (56.5%)	60 (43.5%)	138	
"Stiff rim" sign in SWE				$p < 0.001$
Present	304 (68.0%)	143 (32.0%)	447	
Absent	141 (82.0%)	31 (18.0%)	172	

BI-RADS, Breast Imaging Reporting and Data System; LN, lymph node; SWE, shear wave elastography.

TABLE 3 | Quantitative evaluation of SWE in the internal and peritumor tissues between positive and negative LN metastasis groups.

	Negative LN (n = 445)	Positive LN (n = 174)	*p* value
SWE parameters of the tumor			
E_{mean} (kPa)	24.50 (17.88–32.89)	23.66 (18.00–34.07)	0.754
E_{max} (kPa)	115.03 (79.18–165.53)	110.05 (73.31–163.98)	0.743
E_{min} (kPa)	3.64 (1.67–5.41)	2.67 (1.18–4.79)	0.005
E_{sd} (kPa)	15.02 (10.67–21.64)	14.67 (9.35–22.11)	0.344
E_{ratio}	5.61 (4.20–7.31)	5.24 (4.27–7.39)	0.579
Shell of tumor			
1 mm			
E_{mean} (kPa)	34.47 (25.88–47.87)	33.34 (25.75–47.33)	0.564
E_{max} (kPa)	140.54 (97.98–207.07)	131.76 (89.86–194.43)	0.237
E_{min} (kPa)	3.23 (1.05–5.65)	2.70 (1.01–5.35)	0.281
E_{sd} (kPa)	22.92 (15.87–31.86)	21.91 (14.78–30.08)	0.336
2 mm			
E_{mean} (kPa)	36.02 (26.27–47.75)	34.73 (25.13–47.18)	0.523
E_{max} (kPa)	147.05 (105.21–212.23)	146.00 (100.27–199.41)	0.500
E_{min} (kPa)	2.72 (0.90–5.27)	2.32 (0.83–4.39)	0.141
E_{sd} (kPa)	22.63 (16.79–32.33)	22.11 (15.13–30.31)	0.351
3 mm			
E_{mean} (kPa)	35.97 (25.60–45.77)	34.64 (23.77–45.69)	0.495
E_{max} (kPa)	147.68 (104.28–209.37)	146.00 (101.97–202.73)	0.775
E_{min} (kPa)	2.65 (0.98–5.40)	2.03 (0.66–3.55)	0.007
E_{sd} (kPa)	22.23 (16.27–31.31)	21.42 (15.43–28.84)	0.238

All data represent as median with interquartile range in parentheses. E_{mean}, mean elastic modulus; E_{max}, the maximum elastic modulus; E_{min}, minimum elastic modulus; E_{sd}, elastic modulus standard deviation; E_{ratio}, ratio between mean elastic modulus of breast lesions and normal fatty tissue; LN, lymph node; SWE, shear wave elastography.

(**Table 7**), and a significant difference was shown between the two groups ($p < 0.001$). The percentage of patients with positive LNs from grades 0 to 4 was constantly increased, with 7.1% (3/42) for grade 0, 19.2% (30/156) for grade 1, 24.5% (46/188) for grade 2, 35.6% (64/180) for grade 3, and 58.5% (31/53) for grade 4, respectively, which is shown in **Figure 3**. The Cochran–Armitage trend test indicated that the risk of axillary LN metastasis was increased with the increase in grade level ($p < 0.0001$). Results of the logistic analysis of the comprehensive grade showed that the OR of axillary LN metastasis was increased progressively with rising grade. The risk of axillary LN metastasis was 3.095 times higher in patients with risk grade 1 than in patients with risk grade 0 (p = 0.074), and 4.211–18.318 times higher in patients with risk grades 2–4 (p < 0.05 for all). The risk grade model predicted axillary LN metastasis with an AUC of 0.656 (95% CI, 0.617–0.693), a sensitivity of 54.60% (95% CI, 46.9%–62.1%), a specificity of 68.99% (95% CI, 64.5%–73.3%), and a cutoff value of grade 2.

Since risk category 1 was not different from category 0, we also compared the difference of OR values between grade 1 vs. grades 2, 3, and 4 and grades 0–1 vs. grades 2, 3, and 4 in the risk system. Results showed that the risk of axillary LN metastasis was

TABLE 4 | Quantitative SWE evaluation of the tumor and peritumor shell for the predicting of axillary LN metastasis.

	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Cutoff value	AUC (95% CI)	*p* value
SWE parameters of the tumor					
E_{mean} (kPa)	54.02 (46.3–61.6)	51.69 (46.9–56.4)	≤24.16	0.508 (0.468–0.548)	0.756
E_{max} (kPa)	4.60 (2.0–8.9)	89.89 (86.7–92.5)	≤49.43	0.508 (0.468–0.549)	0.745
E_{min} (kPa)	59.77 (52.1–67.1)	55.96 (51.2–60.6)	≤3.31	0.573 (0.533–0.613)	< 0.01
E_{sd} (kPa)	27.59 (21.1–34.9)	80.67 (76.7–84.2)	≤9.91	0.524 (0.484–0.564)	0.352
E_{ratio}	50.57 (42.9–58.2)	56.40 (51.7–61.1)	≤5.24	0.514 (0.474–0.554)	0.582
Shell of tumor					
1 mm					
E_{mean} (kPa)	44.83 (37.3–52.5)	62.02 (57.3–66.6)	≤30.90	0.515 (0.475–0.555)	0.567
E_{max} (kPa)	27.01 (46.3–61.6)	79.78 (75.7–83.4)	≤90.87	0.531 (0.490–0.570)	0.234
E_{min} (kPa)	63.22 (55.6–70.4)	45.84 (41.1–50.6)	≤3.67	0.528 (0.488–0.568)	0.282
E_{sd} (kPa)	40.23 (32.9–47.9)	66.29 (61.7–70.7)	≤18.34	0.525 (0.485–0.565)	0.338
2 mm					
E_{mean} (kPa)	27.59 (21.1–34.9)	77.98 (73.8–81.7)	≤25.51	0.517 (0.476–0.557)	0.525
E_{max} (kPa)	74.14 (67.0–80.5)	32.13 (27.8–36.7)	≤193.19	0.517 (0.477–0.557)	0.498
E_{min} (kPa)	67.82 (60.3–74.7)	42.92 (38.3–47.7)	≤3.38	0.538 (0.498–0.578)	0.134
E_{sd} (kPa)	47.13 (39.5–54.8)	60.22 (55.5–64.8)	≤20.59	0.524 (0.484–0.564)	0.353
3 mm					
E_{mean} (kPa)	27.59 (21.1–34.9)	78.43 (74.3–82.2)	≤24.61	0.518 (0.477–0.558)	0.497
E_{max} (kPa)	82.76 (76.3–88.1)	22.47 (18.7–26.6)	≤218.08	0.507 (0.467–0.547)	0.772
E_{min} (kPa)	75.29 (68.2–81.5)	39.78 (35.2–44.5)	≤3.52	0.570 (0.530–0.609)	< 0.01
E_{sd} (kPa)	37.93 (30.7–45.6)	70.11 (65.6–74.3)	≤17.85	0.530 (0.490–0.570)	0.238

AUC, area under curve; CI, confidence interval; E_{mean}, mean elastic modulus; E_{max}, the maximum elastic modulus; E_{min}, minimum elastic modulus; E_{sd}, elastic modulus standard deviation; E_{ratio}, ratio between mean elastic modulus of breast lesions and normal fatty tissue; LN, lymph node; SWE, shear wave elastography.

TABLE 5 | Odds ratios for the suspicious features between positive and negative LN groups and the corresponding weighting values.

Suspicious features	Logistic regression		
	OR (95% CI)	*p* value[a]	Weighting value
BI-RADS categories			
3–4C	1.00 (reference)		1
5	2.155 (1.431–3.247)	<0.05	2
E_{min}			
>3.31	1.00 (reference)	<0.05	1
≤3.31	1.654 (1.120–2.442)		2
E_{min} of shell 3 mm			
>3.52	1.00 (reference)		1
≤3.52	1.564 (1.019–2.401)	<0.05	2
"Stiff rim" sign in SWE			
Present	1.00 (reference)		1
Absent	1.900 (1.210–2.982)	<0.05	2

BI-RADS, Breast Imaging Reporting and Data System; E_{min}, minimum elastic modulus; LN, lymph node; OR, odds ratios; SWE, shear wave elastography.
[a]*Compared with reference.*

higher in patients with risk grades 3 and 4 than in patients with risk grades 0–1 and grade 1 ($p < 0.05$ for all) (**Table 8**).

DISCUSSION

The risk of LN metastasis constantly exists in breast cancer patients, and it is expected to be provided in an early stage in order to guide the clinical decision and improve the tumor survival. SWE has been used to predict axillary LN metastasis in invasive breast cancer. Previous studies have reported that higher elasticity of the tumor was related to lymph node metastasis of breast cancer (22, 23), and further study has revealed that the mean elastic modulus of internal breast tumor could independently predict axillary LN metastasis (17). Wen et al. reported that the E_{max} and E_{mean} of primary tumors were higher in LN metastasis cases (24). In our study, however, quantitative SWE parameters including E_{mean}, E_{max}, E_{sd}, and E_{ratio} did not show a significant difference between patients with positive and negative LN metastasis groups ($p > 0.05$ for all), indicating that these SWE parameters of intrinsic tumor and peritumor tissues might not provide sufficient information in predicting LN metastasis for our research group. The difference between the results of this study and previous studies may be due to the difference of the study sample and the difference of the ultrasound instruments used.

TABLE 6 | Comparison of the logistic regression equation, the weighting, and the counting methods in predicting the axillary LN metastasis.

Method	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	AUC (95% CI)
Logistic regression equation	59.20 (51.5–66.6)	64.72 (60.1–69.2)	0.659 (0.621–0.697)
The weighting method	54.60 (46.9–62.1)	68.99 (64.5–73.3)	0.656 (0.617–0.693)
The counting method	54.60 (46.9–62.1)	68.99 (64.5–73.3)	0.656 (0.617–0.693)

AUC, area under curve; CI, confidence interval; LN, lymph node.

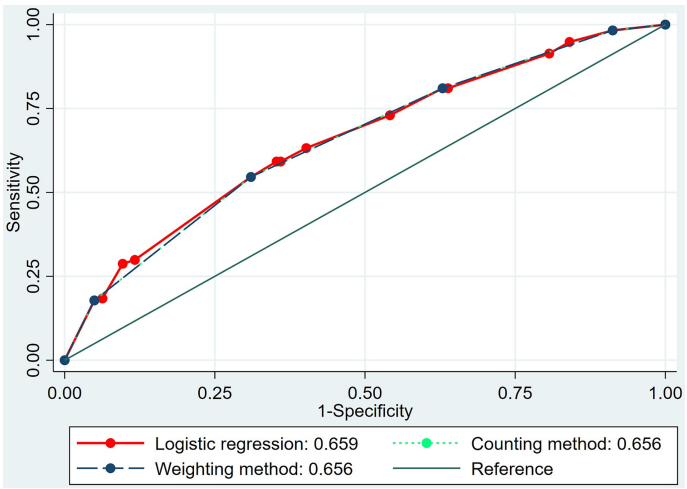


FIGURE 2 | The area under the receiver operating characteristic (ROC) curves of the logistic regression equation, the weighting method, and the counting method in predicting axillary lymph node metastasis of invasive breast cancer.

TABLE 7 | Proportion and odds ratios for each risk grade in negative and positive LN metastasis groups.

Risk grade	Negative LN (n = 445)	Positive LN (n = 174)	OR	*p* value[a]
0	39 (92.9%)	3 (7.1%)	1.00 (reference)	–
1	126 (80.8%)	30 (19.2%)	3.095 (0.896–10.696)	0.074
2	142 (75.5%)	46 (24.5%)	4.211 (1.243–14.271)	0.021
3	116 (64.4%)	64 (35.6%)	7.172 (2.132–24.132)	0.001
4	22 (41.5%)	31 (58.5%)	18.318 (5.016–66.892)	<0.001

LN, lymph node; OR, odds ratios.
[a]*Compared with reference.*

In our study, SWE parameters including the E_{min} of internal tumor and shell 3 mm and the "stiff rim" sign showed a significant difference between patients with positive and negative LN metastasis groups ($p < 0.05$ for all). Zhou et al. have reported that the "stiff rim" sign could select the malignant breast lesions, and possible reasons may contribute to cancer cell infiltration into interstitial tissues or a desmoplastic reaction (16), indicating that a qualitative assessment of peritumoral tissue stiffness may reflect the progress of invasive malignant tumors to some extent. Peritumoral stroma and peritumoral invasion played a critical role in the spreading and metastasis of breast cancer (18, 19, 25, 26) and were also responsible for the formation of the "stiff rim" sign, as well as the attenuation of shear wave energy (16). The low-quality shear waves caused by attenuation may be interpreted as slow-speed shear waves (16). Therefore, the "stiff rim" sign was mostly found in the positive LN metastasis group, and the E_{min} of the positive LN metastasis group was lower than that of the negative LN metastatic group.

There was a significant difference in the constituent ratio between 5 BI-RADS categories ($p < 0.001$). In the 619 masses, 481 lesions were defined as BI-RADS 3-4C, with about 25% (ranging from 22.4% to 27.3%) cases in positive metastasis LN, while in 138 masses of BI-RADS 5, patients with positive LN metastasis were up to 43.5%. Guo et al. revealed that irregular shape and higher color Doppler flow imaging grades were associated with axillary LN metastasis of breast cancer (27). Bae et al. showed that architectural distortion of breast tumors was associated with LN metastasis of breast cancer (28). Previous studies also indicated that a non-circumscribed margin and heterogeneous internal echo were potential predictors of

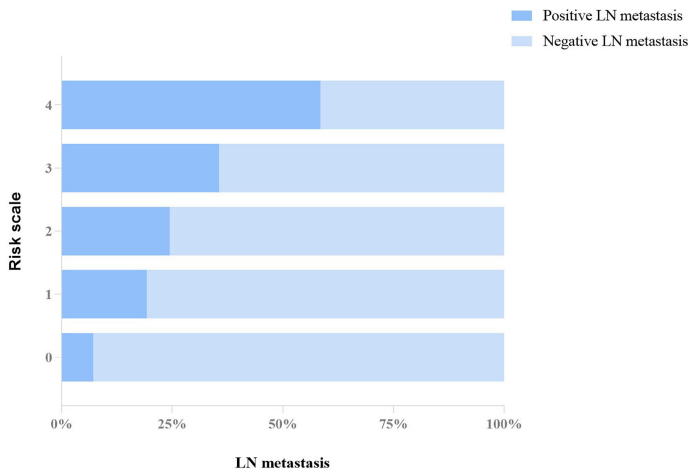


FIGURE 3 | Proportion distribution of positive and negative LN metastasis in each grade of the risk system.

TABLE 8 | Differences of odds ratios between grade 1 and grades 2, 3, 4, as well as between grades 0–1 and grades 2, 3, 4.

Grade group	OR	*p* value[a]
Grade 1 vs. grade 2, 3, 4		
1	1.00 (reference)	–
2	1.361 (0.810–2.286)	0.245
3	2.317 (1.403–3.827)	< 0.01
4	5.918 (3.010–11.636)	< 0.01
Grade 0–1 vs. grade 2, 3, 4		
0–1	1.00 (reference)	–
2	1.620 (0.982–2.671)	0.059
3	2.759 (1.703–4.469)	< 0.01
4	7.045 (3.634–13.659)	< 0.01

LN, lymph node; OR, odds ratios.
[a]*Compared with reference.*

axillary LN metastasis independently (28, 29). Possible reasons might be that the non-circumscribed margins reflect the invasiveness of tumors and aggressive of metastatic lesions to lymphocytes. Besides, tumor growth rate might influence the internal echo of the masses, and rapid tumor growth could lead to an increased likelihood of LN metastasis (29). These ultrasound features, such as irregular shape, non-circumscribed margin, and architectural distortion, were considered as determining characteristics of ultrasound BI-RADS, which may explain the higher incidence of LN metastasis in BI-RADS 5 tumors in this study.

A previous study revealed that tumor staging was an independent predictor of breast cancer LN metastasis (30). A tumor size of 20 mm or greater on ultrasound was an independent predictor of LN metastasis (31). In the present study, however, there was no significant difference in the number of tumors evaluated as T1 and T2/T3 in LN-positive and -negative groups (p = 0.340), which was similar to the results of the previous study; that is, tumor size on ultrasound was not an independent predictor of axillary LN metastasis (28). Breast tumor location has been acknowledged as an independent factor of tumor prognosis, and malignant tumors in the lower inner quadrant were related to a worse prognosis because of the internal mammary lymphatic pathway without sufficient studies (32). However, there was no significant difference in tumor location (p = 0.390) in LN-positive and -negative groups in our study.

The grade system has been extensively used in categorical variables of clinical and imaging data analysis (1, 33). In the prediction of LN metastasis, there was no statistical difference among the logistic regression equation, the weighting method, and the counting method (p > 0.05 for all). For ease of use, a 0–4 grade risk stratification system for LN metastasis was established based on the counting method in this study. The risk system was based on quantitative and qualitative SWE features as well as BI-RADS category. To the best of our knowledge, no studies have combined quantitative and qualitative SWE characteristics of tumors, especially peritumor tissues, and the BI-RADS category to predict axillary LN metastasis in breast invasive cancer. Our results showed that the risk of axillary LN metastasis of breast cancer increased as the risk grade rises from 0 to 4, with the increase of the four significant factors (BI-RADS, "stiff rim" sign,

E_{min} of internal tumor, and shell 3 mm). A patient with a BI-RADS 5 breast tumor exhibiting the "stiff rim" sign (value of E_{min} in internal tumor and shell 3 mm ≤3.31 and ≤3.52, respectively) was assessed as grade 4 in our risk system. The OR of each grade in the risk system gradually increased as the increasing levels were 3.095 ($p = 0.074$, 95% CI = 0.896–10.696) for grade 1, 4.211 ($p < 0.05$, 95% CI = 1.243–14.271) for grade 2, 7.172 ($p < 0.05$, 95% CI = 2.132–24.132) for grade 3, and 18.318 ($p < 0.05$, 95% CI = 5.016–66.892) for grade 4, respectively, compared with grade 0, indicating that patients with a higher risk grade were more likely to have LN metastasis. The diagnostic performance of this risk grade system is significant in assessing the risk grade of LN metastasis, although the AUC is not excellent (0.656, 95% CI, 0.617–0.693).

The merit of our research is that this is a multicenter study covering 17 hospitals, which effectively improves the reliability of the research data. However, the study still has some limitations. Firstly, the evaluation of the BI-RADS category and the SWE features of breast lesions were influenced by the experience of the radiologist. However, previous studies have confirmed that the reproducibility ranged from moderate to substantial for the BI-RADS category and ranged from substantial to almost perfect for both the quantitative SWE features and the "stiff rim" sign (16, 34, 35). Secondly, all SWE information were acquired from the same ultrasound system (Resona 7, Mindray), which would limit the generalizability of the study results. Further studies using different ultrasound SWE systems are warranted in the future. Third, the AUC of the risk system is not very high. In the future, better predictive models may be obtained if the information provided by ultrasound, MRI, mammography, and clinical data is utilized jointly.

CONCLUSION

A 0–4 grade risk stratification system, based on quantitative and qualitative SWE features as well as BI-RADS category, for predicting axillary LN metastasis in invasive breast cancer was established in this study. Although the diagnostic performance of this risk system was not excellent, there is no doubt that it has the potential to predict axillary LN metastasis by the combination of SWE and conventional ultrasound.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the generated datasets belong to department of ultrasound, Ruijin Hospital, Shanghai Jiaotong University School of Medicine. Requests to access the datasets should be directed to JQZ, zhousu30@126.com.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Ruijin Hospital, Shanghai Jiaotong University School of Medicine. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JWZ drafted and edited the manuscript. HTZ, YJD, XHJ, JWZ, ZYL, ZRC, YJX, BH, YXH, CC, JFX, FJD, XNX, CRW, WJH, GW, QYL, QC, WYD, QCJ, YLM, HNY, XJX, HJY, PZ, YS, LGC, PH, LXQ, JPL, LYS, YNZ, and YYX collected the clinical and ultrasound data. YJD and WWZ analyzed the images. XHJ collected the ultrasound images. XHJ and JWZ collected the autopsy and pathology data. HTZ prepared for the literature. YYS and HTZ completed the statistical analysis. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation of China (grant no. 82071928).

ACKNOWLEDGMENTS

We would like to express our sincere gratitude to all doctors from the 17 hospitals participating in the multicenter study who provided data for this study.

REFERENCES

1. T. I. A. f. R. o. Cancer. *Latest Global Cancer Data: Cancer Burden Rises to 19.3 Million New Cases and 10.0 Million Cancer Deaths in 2020*. Geneva: World Health Organization (2020).
2. Araujo T, Aresta G, Castro E, Rouco J, Aguiar P, Eloy C, et al. Classification of Breast Cancer Histology Images Using Convolutional Neural Networks. *PloS One* (2017) 12(6):e0177544. doi: 10.1371/journal.pone.0177544
3. Cao Z, Duan L, Yang G, Yue T, Chen Q. An Experimental Study on Breast Lesion Detection and Classification From Ultrasound Images Using Deep Learning Architectures. *BMC Med Imaging* (2019) 19(1):51. doi: 10.1186/s12880-019-0349-x
4. Zeng H, Zheng R, Zhang S, Zou X, Chen W. Female Breast Cancer Statistics of 2010 in China: Estimates Based on Data From 145 Population-Based Cancer Registries. *J Thorac Dis* (2014) 6(5):466–70. doi: 10.3978/j.issn.2072-1439.2014.03.03
5. Tao C, Chen K, Han L, Peng Y, Li C, Hua Z, et al. New One-Step Model of Breast Tumor Locating Based on Deep Learning. *J Xray Sci Technol* (2019) 27 (5):839–56. doi: 10.3233/XST-190548
6. Giuliano AE, Edge SB, Hortobagyi GN. Eighth Edition of the AJCC Cancer Staging Manual: Breast Cancer. *Ann Surg Oncol* (2018) 25(7):1783–5. doi: 10.1245/s10434-018-6486-6
7. Lyman GH, Giuliano AE, Somerfield MR, Benson AB, Bodurka 3DC, Burstein HJ, et al. American Society of Clinical Oncology Guideline Recommendations for Sentinel Lymph Node Biopsy in Early-Stage Breast Cancer. *J Clin Oncol* (2005) 23(30):7703–20. doi: 10.1200/JCO.2005.08.001
8. Mercado CL. BI-RADS Update. *Radiologic Clinics of North America* (2014) 52 (3):481-7. doi: 10.1016/j.rcl.2014.02.008

9. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, van Ginneken B, Karssemeijer N, Litjens G, et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA* (2017) 318(22):2199–210. doi: 10.1001/jama.2017.14585
10. Zheng X, Yao Z, Huang Y, Yu Y, Wang Y, Liu Y, et al. Deep Learning Radiomics can Predict Axillary Lymph Node Status in Early-Stage Breast Cancer. *Nat Commun* (2020) 11(1):1236. doi: 10.1038/s41467-020-15027-z
11. Valente SA, Levine GM, Silverstein MJ, Rayhanabad JA, Weng-Grumley JG, Ji L, et al. Accuracy of Predicting Axillary Lymph Node Positivity by Physical Examination, Mammography, Ultrasonography, and Magnetic Resonance Imaging. *Ann Surg Oncol* (2012) 19(6):1825–30. doi: 10.1245/s10434-011-2200-7
12. Chou Y-H, Tiu C-M, Lai Y-C, Lin Y-H. Elastography of Breast Tumors: Pitfalls. *Ultrasound Med Biol* (2019) 45:S22. doi: 10.1016/j.ultrasmedbio.2019.07.484
13. Nori J, Vanzi E, Bazzocchi M, Bufalini FN, Distante V, Branconi F, et al. Role of Axillary Ultrasound Examination in the Selection of Breast Cancer Patients for Sentinel Node Biopsy. *Am J Surg* (2007) 193(1):16–20. doi: 10.1016/j.amjsurg.2006.02.021
14. Lee MC, Eatrides J, Chau A, Han G, Kiluk JV, Khakpour N, et al. Consequences of Axillary Ultrasound in Patients With T2 or Greater Invasive Breast Cancers. *Ann Surg Oncol* (2011) 18(1):72–7. doi: 10.1245/s10434-010-1171-4
15. Mercado CL. BI-RADS Update. *Radiol Clin North Am* (2014) 52(3):481–7. doi: 10.1016/j.rcl.2014.02.008
16. Zhou J, Zhan W, Chang C, Zhang X, Jia Y, Dong Y, et al. Breast Lesions: Evaluation With Shear Wave Elastography, With Special Emphasis on the "Stiff Rim" Sign. *Radiology* (2014) 272(1):63–72. doi: 10.1148/radiol.14130818
17. Evans A, Rauchhaus P, Whelehan P, Thomson K, Purdie CA, Jordan LB, et al. Does Shear Wave Ultrasound Independently Predict Axillary Lymph Node Metastasis in Women With Invasive Breast Cancer? *Breast Cancer Res Treat* (2014) 143(1):153–7. doi: 10.1007/s10549-013-2747-z
18. Colleoni M, Rotmensz N, Maisonneuve P, Sonzogni A, Pruneri G, Casadio C, et al. Prognostic Role of the Extent of Peritumoral Vascular Invasion in Operable Breast Cancer. *Ann Oncol* (2007) 18(10):1632–40. doi: 10.1093/annonc/mdm268
19. de Mascarel I, Bonichon F, Durand M, Mauriac L, MacGrogan G, Soubeyran I, et al. Obvious Peritumoral Emboli: An Elusive Prognostic Factor Reappraised. Multivariate Analysis of 1320 Node-Negative Breast Cancers. *Eur J Cancer* (1998) 34(1):58–65. doi: 10.1016/s0959-8049(97)00344-4
20. Barr RG, Nakashima K, Amy D, Cosgrove D, Farrokh A, Schafer F, et al. WFUMB Guidelines and Recommendations for Clinical Use of Ultrasound Elastography: Part 2: Breast. *Ultrasound Med Biol* (2015) 41(5):1148–60. doi: 10.1016/j.ultrasmedbio.2015.03.008
21. Zhou J, Song Y, Zhan W, Wei X, Zhang S, Zhang R, et al. Thyroid Imaging Reporting and Data System (TIRADS) for Ultrasound Features of Nodules: Multicentric Retrospective Study in China. *Endocrine* (2021) 72(1):157–70. doi: 10.1007/s12020-020-02442-x
22. Evans A, Whelehan P, Thomson K, McLean D, Brauer K, Purdie C, et al. Invasive Breast Cancer: Relationship Between Shear-Wave Elastographic Findings and Histologic Prognostic Factors. *Radiology* (2012) 263(3):673–7. doi: 10.1148/radiol.12111317
23. Youk JH, Gweon HM, Son EJ, Kim JA, Jeong J. Shear-Wave Elastography of Invasive Breast Cancer: Correlation Between Quantitative Mean Elasticity Value and Immunohistochemical Profile. *Breast Cancer Res Treat* (2013) 138(1):119–26. doi: 10.1007/s10549-013-2407-3
24. Wen X, Yu X, Tian Y, Liu Z, Cheng W, Li H, et al. Quantitative Shear Wave Elastography in Primary Invasive Breast Cancers, Based on Collagen-S100A4 Pathology, Indicates Axillary Lymph Node Metastasis. *Quant Imaging Med Surg* (2020) 10(3):624–33. doi: 10.21037/qims.2020.02.18
25. Wernicke M, Roitman P, Manfre D, Stern R. Breast Cancer and the Stromal Factor. The "Prometastatic Healing Process" Hypothesis. *Med (B Aires)* (2011) 71(1):15–21.
26. Auvinen P, Tammi R, Parkkinen J, Tammi M, Agren U, Johansson R, et al. Hyaluronan in Peritumoral Stroma and Malignant Cells Associates With Breast Cancer Spreading and Predicts Survival. *Am J Pathol* (2000) 156(2):529–36. doi: 10.1016/S0002-9440(10)64757-8
27. Guo Q, Dong Z, Zhang L, Ning C, Li Z, Wang D, et al. Ultrasound Features of Breast Cancer for Predicting Axillary Lymph Node Metastasis. *J Ultrasound Med* (2018) 37(6):1354–3. doi: 10.1002/jum.14469
28. Bae MS, Shin SU, Song SE, Ryu HS, Han W, Moon WK. Association Between US Features of Primary Tumor and Axillary Lymph Node Metastasis in Patients With Clinical T1-T2N0 Breast Cancer. *Acta Radiol* (2018) 59(4):402–8. doi: 10.1177/0284185117723039
29. Zhang H, Sui X, Zhou S, Hu L, Huang X. Correlation of Conventional Ultrasound Characteristics of Breast Tumors With Axillary Lymph Node Metastasis and Ki-67 Expression in Patients With Breast Cancer. *J Ultrasound Med* (2019) 38(7):1833–40. doi: 10.1002/jum.14879
30. Stachs A, Thi AT, Dieterich M, Stubert J, Hartmann S, Glass A, et al. Assessment of Ultrasound Features Predicting Axillary Nodal Metastasis in Breast Cancer: The Impact of Cortical Thickness. *Ultrasound Int Open* (2015) 1(1):E19–24. doi: 10.1055/s-0035-1555872
31. Zhu AQ, Li XL, An LW, Guo LH, Fu HJ, Sun LP, et al. Predicting Axillary Lymph Node Metastasis in Patients With Breast Invasive Ductal Carcinoma With Negative Axillary Ultrasound Results Using Conventional Ultrasound and Contrast-Enhanced Ultrasound. *J Ultrasound Med* (2020) 39(10):2059–70. doi: 10.1002/jum.15314
32. Gou ZC, Liu XY, Xiao Y, Zhao S, Jiang YZ, Shao ZM. Decreased Survival in Patients With Carcinoma of Axillary Tail Versus Upper Outer Quadrant Breast Cancers: A SEER Population-Based Study. *Cancer Manag Res* (2018) 10:1133–41. doi: 10.2147/CMAR.S165291
33. Liu Z, Wang R, Zhou J, Zheng Y, Dong Y, Luo T, et al. Ultrasound Lymphatic Imaging for the Diagnosis of Metastatic Central Lymph Nodes in Papillary Thyroid Cancer. *Eur Radiol* (2021) 31(11):8458–67. doi: 10.1007/s00330-021-07958-y
34. Schwab F, Redling K, Siebert M, Schotzau A, Schoenenberger CA, Zanetti-Dallenbach R. Inter- and Intra-Observer Agreement in Ultrasound BI-RADS Classification and Real-Time Elastography Tsukuba Score Assessment of Breast Lesions. *Ultrasound Med Biol* (2016) 42(11):2622–9. doi: 10.1016/j.ultrasmedbio.2016.06.017
35. Cosgrove DO, Berg WA, Dore CJ, Skyba DM, Henry JP, Gay J, et al. Shear Wave Elastography for Breast Masses is Highly Reproducible. *Eur Radiol* (2012) 22(5):1023–32. doi: 10.1007/s00330-011-2340-y

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhang, Dong, Jia, Zhang, Li, Chuan, Xu, Hu, Huang, Chang, Xu, Dong, Xia, Wu, Hu, Wu, Li, Chen, Deng, Jiang, Mou, Yan, Xu, Yan, Zhou, Shao, Cui, He, Qian, Liu, Shi, Zhao, Xu, Song, Zhan and Zhou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 23 March 2022
doi: 10.3389/fonc.2022.794975

Establishment and Validation of a Novel MRI Radiomics Feature-Based Prognostic Model to Predict Distant Metastasis in Endemic Nasopharyngeal Carcinoma

Hao-Jiang Li[1†], Li-Zhi Liu[1†], Ying Huang[2†], Ya-Bin Jin[3], Xiang-Ping Chen[3], Wei Luo[3], Jian-Chun Su[4], Kai Chen[4], Jing Zhang[4] and Guo-Yi Zhang[4]*

[1] Department of Radiology, State Key Laboratory of Oncology in Southern China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Sun Yat-Sen University Cancer Center, Guangzhou, China, [2] Department of Radiation Oncology, State Key Laboratory of Oncology in Southern China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Sun Yat-Sen University Cancer Center, Guangzhou, China, [3] Clinical Research Institute, Foshan Academy of Medical Sciences, Sun Yat-Sen University Foshan Hospital and The First People's Hospital of Foshan, Foshan, China, [4] Department of Radiation Oncology, Foshan Academy of Medical Sciences, Sun Yat-Sen University Foshan Hospital and The First People's Hospital of Foshan, Foshan, China

Purpose: We aimed to establish a prognostic model based on magnetic resonance imaging (MRI) radiomics features for individual distant metastasis risk prediction in patients with nasopharyngeal carcinoma (NPC).

Methods: Regression analysis was applied to select radiomics features from T1-weighted (T1-w), contrast-enhanced T1-weighted (T1C-w), and T2-weighted (T2-w) MRI scans. All prognostic models were established using a primary cohort of 518 patients with NPC. The prognostic ability of the radiomics, clinical (based on clinical factors), and merged prognostic models (integrating clinical factors with radiomics) were identified using a concordance index (C-index). Models were tested using a validation cohort of 260 NPC patients. Distant metastasis-free survival (DMFS) were calculated by using the Kaplan-Meier method and compared by using the log-rank test.

Results: In the primary cohort, seven radiomics prognostic models showed similar discrimination ability for DMFS to the clinical prognostic model (P=0.070-0.708), while seven merged prognostic models displayed better discrimination ability than the clinical prognostic model or corresponding radiomics prognostic models (all P<0.001). In the validation cohort, the C-indices of seven radiomics prognostic models (0.645-0.722) for DMFS prediction were higher than in the clinical prognostic model (0.552) (P=0.016 or <0.001) or in corresponding merged prognostic models (0.605-0.678) (P=0.297 to 0.857), with T1+T1C prognostic model (based on Radscore combinations of T1 and T1C Radiomics models) showing the highest C-index (0.722). In the decision curve analysis of

OPEN ACCESS

Edited by:
Tingying Peng, Helmholtz zentrum München, Germany

Reviewed by:
Carlos M. Chiesa-Estomba, Donostia University Hospital, Spain
Yu Zhao, Technical University of Munich, Germany

***Correspondence:**
Guo-Yi Zhang
guoyizhff@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *14 October 2021*
Accepted: *23 February 2022*
Published: *23 March 2022*

Citation:
Li H-J, Liu L-Z, Huang Y, Jin Y-B, Chen X-P, Luo W, Su J-C, Chen K, Zhang J and Zhang G-Y (2022) Establishment and Validation of a Novel MRI Radiomics Feature-Based Prognostic Model to Predict Distant Metastasis in Endemic Nasopharyngeal Carcinoma. Front. Oncol. 12:794975. doi: 10.3389/fonc.2022.794975

the validation cohort for all prognostic models, the T1+T1C prognostic model displayed the best performance.

Conclusions: Radiomics models, especially the T1+T1C prognostic model, provided better prognostic ability for DMFS in patients with NPC.

Keywords: MRI, radiomics, nasopharyngeal carcinoma, prognosis, predictive model

INTRODUCTION

Nasopharyngeal carcinoma (NPC) is endemic to southern China, southeastern Asia, and northern Africa, where the peak incidence rate is 20–50 cases per 100 000 individuals (1, 2). Advances in NPC diagnosis and treatment using MRI, intensity-modulated radiotherapy, and combined chemotherapy, has significantly improved locoregional control, and distant metastasis is regarded as the main cause of treatment failure (3–5). In a meta-analysis that reviewed 13304 participants, intensity-modulated radiotherapy was associated with better 5-year locoregional control (odds ratio 2.08; 95% confidence interval [CI]=1.82–2.37), while there were no significant differences seen in distant metastasis-free survival (DMFS) compared with conventional radiotherapy (4). Although the current 5-year overall survival rate of NPC patients reach 80-88%, 15–25% of patients develop distant metastasis after treatment, especially in advanced NPC (5, 6), so to identify high-risk patients to guide optimal treatment decisions is vital.

Currently, the TNM staging system is widely used to predict prognosis and establish treatment strategies among patients with NPC. However, patients at the same clinical stage receiving similar treatments often have different outcomes (7). Among NPC patients with distant metastasis after treatment, nearly 80% are stage III or IV (8), when it is difficult to identify high-risk patients. The presence of plasma Epstein-Barr virus (EBV) DNA influences the disease process in NPC patients and is regarded as a potential biomarker with clinical applications (9). Other risk factors have been associated with prognosis in patients with NPC, including age, sex, smoking status, hepatitis B surface antigen, serum lactate dehydrogenase, high-sensitivity C-reactive protein, and microRNA or mRNA; however, results are inconsistent (10–15). Therefore, the prognostic value of these markers needs confirmation.

Radiomics focuses on extracting quantitative features from medical images to find a possible association with tumor phenotypic characteristics (16). Radiomic features were found to be associated with gene and/or protein expression profiles, treatment response, and clinical outcomes in different cancer types (16, 17). Pretreatment MRI radiomics or MRI-based radiomics nomograms were proposed to predict locoregional recurrence, distant metastasis, and early response to induction chemotherapy in endemic NPC (18–29). However, there is no effective radiomics-based model to distinguish different risks levels for distant metastasis among NPC patients.

We conducted a study of multiparametric MRI radiomics features to identify and validate an MRI-based radiomics model discriminating distant metastasis in patients with NPC. Moreover, we compared the accuracy of a radiomics prognostic model (based on selected radiomics features), a clinical prognostic model (based on clinical risk factors), and a prognostic model merging the two, for discriminating distant metastasis in NPC patients.

Abbreviations: CI confidence interval; C-index concordance index; DMFS distant metastasis-free survival; EBV Epstein-Barr virus; NPC nasopharyngeal carcinoma; T1C-w contrast-enhanced; T1-weighted; T1-w T1-weighted; T2-w T2-weighted.

METHODS

Participants

This retrospective study was approved by the Clinical Research Ethics Committee of the Sun Yat-Sen University Cancer Center (SYSUCC); pre-treatment written informed consent was obtained from all patients or their next of kin. The authenticity of the study was validated by uploading the fully raw data onto the Research Data Deposit (RDD) public platform (http://www.researchdata.org.cn), with the approval RDD number as RDDA 2021135093. There were 903 patients newly diagnosed with untreated and non-metastatic NPC initially enrolled between January 2010 and November 2012; 125 patients were excluded from analysis (see **Supplementary Materials**). The remaining 778 NPC patients were randomly assigned in a 2:1 proportion to the primary (n=518) or validation cohort (n=260).

All patients underwent pretreatment evaluation, including clinical examinations of the head and neck region, fiber optic nasopharyngoscopy, neck and nasopharynx MRI, chest radiography, abdominal sonography, and whole-body bone scan, and staged according to the eighth American Joint Committee on Cancer TNM staging manual (30). Pre-treatment blood samples were collected to determine blood type and the presence or levels of EBV DNA, hepatitis B surface antigen, lactate dehydrogenase, high-sensitivity C-reactive protein, platelets, and leucocytes. A quantitative polymerase chain reaction method was used for detecting plasma EBV DNA (31). Clinical patient characteristics are listed in **Table 1**.

All patients were treated according to the SYSUCC guidelines for NPC patients, which recommend that stage I patients have to receive radiotherapy alone, stage II radiotherapy alone, or concomitant chemoradiotherapy, and stage III–IVa concomitant chemoradiotherapy with or without induction chemotherapy. All patients received definitive intensity-modulated radiotherapy with 6-MV photons during the entire radiotherapy course.

TABLE 1 | Patient demographic characteristics in the primary and validation cohorts.

Characteristic	Primary cohort (*N*=518)	Validation cohort (*N*=260)	*P*
Age (years)			0.944
Median (IQR)	44 (38–53)	46 (38–52)	
Gender			0.389
Male	371 (71.6%)	194 (74.6%)	
Female	147 (28.4%)	66 (25.4%)	
WHO pathologic classification			0.253
Type I/II	27 (5.2%)	19 (7.3%)	
Type III	491 (94.8%)	241 (92.7%)	
T classification			0.389
T1	140 (27%)	61 (23.5%)	
T2	60 (11.6%)	37 (14.2%)	
T3	186 (35.9%)	103 (39.6%)	
T4	132 (25.5%)	59 (22.7%)	
N classification			0.425
N0	128 (24.7%)	53 (20.4%)	
N1	282 (54.4%)	151 (58.1%)	
N2	69 (13.3%)	40 (15.4%)	
N3	39 (7.5%)	16 (6.2%)	
AJCC clinical stage (2010)			0.181
I	53 (10.2%)	19 (7.3%)	
II	117 (22.6%)	57 (21.9%)	
III	185 (35.7%)	112 (43.1%)	
IVa	163 (31.5%)	72 (27.7%)	
Treatment regimen			0.490
RT alone	71 (13.7%)	34 (13.1%)	
CCRT	201 (38.8%)	91 (35%)	
IC + CCRT	246 (47.5%)	135 (51.9%)	
EBV DNA (10^3 copies/ml)*			0.262
<1	232 (44.8%)	125 (48.1%)	
<10	117 (22.6%)	67 (25.8%)	
<100	129 (24.9%)	49 (18.8%)	
≥100	40 (7.7%)	19 (7.3%)	
Blood type			0.145
A	139 (26.8%)	56 (21.5%)	
B	130 (25.1%)	67 (25.8%)	
AB	17 (3.3%)	16 (6.2%)	
o	232 (44.8%)	121 (46.5%)	
HBsAg			0.620
Negative	427 (82.4%)	210 (80.8%)	
Positive	91 (17.6%)	50 (19.2%)	
LDH (U/L)*			0.142
<245	492 (95%)	240 (92.3%)	
≥245	26 (5%)	20 (7.7%)	
hs-CRP (g/ml)*			0.186
<1	196 (37.8%)	101 (38.8%)	
1–3	166 (32%)	96 (36.9%)	
≥3	156 (30.1%)	63 (24.2%)	
Platelet counts (10^9/L)*			0.459
<100	4 (0.8%)	2 (0.8%)	
100–300	436 (84.2%)	227 (87.3%)	
≥300	78 (15.1%)	31 (11.9%)	
Leucocyte counts (10^9/L)*			0.347
<4	9 (1.7%)	9 (3.5%)	
4–10	457 (88.2%)	225 (86.5%)	
≥10	52 (10%)	26 (10%)	
Follow-up time (months)			0.886
Median (min, max)	84.6 (3.3–104.1)	84.4 (6.9–103.9)	

*IQR, interquartile range; WHO, World Health Organization; Type I, keratinizing; Type II, non-keratinizing differentiated; Type III, non-keratinizing undifferentiated; T, tumor; N, node; AJCC, American Joint Committee on Cancer; RT, radiotherapy; CCRT, concurrent chemoradiotherapy; IC, induction chemotherapy; EBV DNA, Plasma Epstein–Barr virus DNA; HBsAg, hepatitis B surface antigen; LDH, serum lactate dehydrogenase levels; hs-CRP, high-sensitivity C-reactive protein; min, minimum; max, maximum. *Results before treatment.*

Concurrent chemotherapy, induction chemotherapy, and intensity-modulated radiotherapy data are presented in the **Supplementary Material**.

Complete follow-up data at 3, 5, and 8 years were available for 97.8%, 96.4%, and 87.6% of patients, respectively. Follow-up visits occurred at least once every 3 months during the first 3 years and then once every 6 months after treatment. DMFS was calculated from the date of treatment initiation to that of distant metastasis at any site, death from any cause, or the date of the last follow-up. Distant metastases were assessed using two imaging methods, including chest radiography, abdominal sonography, bone scan, and CT or MRI; elevation of plasma EBV DNA was confirmed by pathological biopsy if necessary.

MRI Acquisition/Segmentation and Radiomics Feature Extraction

All patients underwent MRI with a 1.5-T system (Signa CV/i; General Electric Healthcare, Chalfont St. Giles, United Kingdom) or 3.0 T system (Siemens Magnetom Tim Trio, Erlangen, Germany), employing the fast spin-echo technique. The scan region imaged ranged from the suprasellar cistern to the inferior margin of the sternal end of the clavicle using a head-and-neck combined coil. T1-weighted images (T1-w) in the axial, coronal, and sagittal planes, T2-weighted (T2-w) images in the axial plane, and contrast-enhanced T1-weighted (T1C-w) in the axial, coronal, and sagittal planes were acquired in all patients. Detailed information on MRI scans is shown in Supplementary Materials.

Axial T1-w, T2-w, and T1C-w Digital Imaging and Communications in Medicine (DICOM) images were retrieved from the picture archiving and communication system (PACS) and loaded into AnalyzePro (https://analyzedirect.com/analyzepro/) for semi-manual segmentation with normalization. One radiologist (Z. Guoyi) with over 10 years of experience in head and neck cancers outlined the tumor contour regions of interest (ROI) on each MRI slice, and each segmentation was validated by a senior radiologist (L. Lizhi) with 20 years of experience in evaluating MRI scans of patients with NPC. Differences were resolved by consensus. Extractions of radiomics features were performed using the open-source PyRadiomics (http://www.radiomics.io/pyradiomics.html). We extracted 4527 radiomics features for each patient from axial T1-w, T2-w, and T1C-w axial images (1509 features from each unimodal MR image). Data from the radiomics features are shown in the **Supplementary Material**.

Statistical Analysis Methods

The interclass correlation coefficient was used to assess the effect of variations in manual segmentation on radiomics feature values from 50 patients randomly extracted from the primary cohort. Features with high reproducibility (median ± standard deviation: interclass correlation coefficient >0.80) were retained for subsequent analysis. After deleting features that had a Pearson correlation coefficient ≥0.75, univariate analysis using the Cox regression model was employed as the first step of feature selection to assess the possible risk factor of DMFS. The least absolute shrinkage and selection operator method was the second feature-selection process for model building.

Supplementary Figures S1, S2 show the processes of radiomics feature selection and radiomics feature selection using the least absolute shrinkage and selection operator regression model, respectively.

Kaplan-Meier survival curves and log-rank tests were used to evaluate the time-event data. Clinical characteristics between the primary and validation cohorts were compared using the chi-square test for categorical variables and the Mann-Whitney U test for continuous variables. Harrell's concordance indices (C-indices) were applied to evaluate the discriminating ability of each prognostic model. The establishment flowchart of the clinical prognostic model, radiomics prognostic models, and merged prognostic models is shown in **Supplementary Figure S3**.

Statistical analyses were performed using R software version 3.2.5. with the following R packages: the caret package for Pearson correlation analysis; the survival package and survminer for Kaplan-Meier survival curves; the glmnet package for least absolute shrinkage and selection operator Cox regression; the ggplot2 package for score plot; the rms package for calibration curves and nomograms; the Hmisc package for comparisons between models in terms of C-indices; the pheatmap and gplots packages for heatmaps. All statistical tests were two-tailed with $P<0.05$ indicating statistical significance.

RESULTS

Participants

Table 1 lists the clinical characteristics of patients with NPC. No significant differences were found in clinical characteristics between the primary and validation cohorts ($P=0.142$ to 0.944). During the follow-up period, 12.9% (67/518) and 13.1% (34/260) of patients developed distant metastases in the primary and validation cohorts, respectively. There was no significant difference between the two cohorts in the distant metastases rate ($P=0.961$). The median DMFS time was 84.3 months (range, 3.3–104.1) for the primary cohort, and 84.4 months (range, 6.6–103.9) for the validation cohort.

Clinical Prognostic Model Building

Univariate analysis indicated that pre-treatment T stage, N stage, clinical stage, and plasma EBV DNA were associated with DMFS in patients with NPC (**Supplementary Table S1**). Multivariable analyses further identified that pre-treatment T stage, N stage, and plasma EBV DNA were independent predictors for DMFS (**Supplementary Table S2**). The clinical prognostic model for DMFS discriminating was based on these three independent factors. The C-index of this model for the primary cohort was 0.736 (95% CI=0.680–0.791), significantly higher than the C-indices for the validation cohort, 0.552 (95% CI=0.457–0.647) (**Table 2**).

Radiomics Prognostic Model Building

Supplementary Table S3 summarizes the features most strongly associated with DMFS based on analysis of the primary cohort. The T1, T1c, and T2 prognostic models for DMFS prediction were based on selected radiomics features derived from T1-w, T1C-w, and T2-w MRI scans, respectively (**Table 3**). The T1+T1C, T1+T2, T2+T1C, and T1+T2+T1C prognostic models were based on Radscore combinations of the T1 and T1C, T1 and T2, T2 and T1C, and T1, T2 and T1C prognostic models, respectively (**Supplementary Figure S3** and **Table 3**).

As shown in **Table 2**, in the primary cohort, seven radiomics prognostic models showed similar DMFS discriminating ability as that of the clinical prognostic model ($P=0.070$ to 0.708). While in the validation cohort, the C-indices for DMFS prediction by the seven models (0.645-0.722) were significantly higher than the C-index calculated using the clinical prognostic model ($P=0.016$ or <0.001). Intriguingly, in the seven radiomics models of the validation cohort, the T1+1C prognostic model showed the highest C-index (0.722, 95% CI=0.632-0.811) for DMFS discriminating, significantly superior to the C-indices of the T1 (0.676), T2 (0.645), and T1C (0.711) prognostic models ($P=0.008$, 0.043, and 0.048, respectively) and higher than those of the other three radiomics models, although without significant differences (**Supplementary Table S4**).

Merged Prognostic Model Building

To identify whether integrating clinical factors with radiomics prognostic models could increase predictive accuracy, we built seven merged prognostic models (MT1, MT2, MT1C, MT1+T2, MT1+T1C, MT2+T1C, and MT1+T2+T1C) based on combinations of three clinical factors and the T1, T2, T1C, T1+T2, T1+T1C, T2+T1C, and T1+T2+T1C radiomics models, respectively (**Supplementary Figure S3**). In the primary cohort, merged prognostic models displayed better discriminating ability in DMFS than the corresponding radiomics prognostic models (all $P<0.001$). While in the validation cohort, the C-indices of seven merged prognostic models (0.605-0.678) for DMFS discriminating were all lower than those of the corresponding radiomics prognostic models; although, there were no significant differences (**Table 2**), indicating that adding the clinical factors to the radiomics models weakened their discriminating ability.

To confirm this finding, we removed the T stage from the merged prognostic models and constructed seven additional merged prognostic models (rMT1, rMT2, rMT1C, rMT1+T2, rMT1+T1C, rMT2+T1C, and rMT1+T2+T1C) based on the radiomics prognostic model, N stage, and EBV-DNA. The C-indices of these seven models (0.612-0.685) were slightly better than those of the corresponding merged prognostic models in the validation cohort (**Table 2**).

Performance of the Clinical and T1+1C Prognostic Model

Based on the results above, T1+1C radiomics model was considered to provide the best prognostic ability for DMFS (having the highest C-index and Least degree of freedom). We further compared the discriminating ability between the clinical prognostic model and the T1+T1C prognostic model for DMFS from their nomograms, calibration curves, risk score distributions, and decision curve analysis.

TABLE 2 | C-index values of different prognostic models for DMFS prediction in the primary and validation cohorts.

Prognostic model		Degree of freedom	Primary cohort			Validation cohort		
			C-index (95% CI)	*P*†	*P**	C-index (95% CI)	*P*†	*P**
Clinical prognostic model		9	0.736 (0.68 0.791)	Reference		0.552 (0.457, 0.647)	Reference	
Radiomics prognostic model	T1	7	0.723 (0.666, 0.780)	0.181	reference	0.676 (0.588, 0.764)	<0.001	reference
	T2	5	0.715 (0.656, 0.774)	0.232	reference	0.645 (0.546, 0.744)	0.016	reference
	T1C	9	0.733 (0.676, 0.791)	0.708	reference	0.711 (0.615, 0.807)	<0.001	reference
	T1+2	2	0.771 (0.720, 0.823)	0.175	reference	0.679 (0.594, 0.763)	<0.001	reference
	T1+1C	2	0.757 (0.703, 0.81)	0.468	reference	0.722 (0.632, 0.811)	<0.001	reference
	T2+1C	2	0.763 (0.712, 0.813)	0.430	reference	0.697 (0.599, 0.795)	<0.001	reference
	T1+2+1C	3	0.784 (0.737, 0.831)	0.070	reference	0.711 (0.622, 0.799)	<0.001	reference
Merged prognostic model	MT1	10	0.784 (0.731, 0.837)	<0.001	<0.001	0.640 (0.558, 0.723)	0.057	0.511
	MT2	10	0.773 (0.719, 0.826)	<0.001	<0.001	0.605 (0.500, 0.710)	0.055	0.647
	MT1C	10	0.791 (0.741, 0.841)	<0.001	<0.001	0.661 (0.563, 0.759)	0.003	0.857
	MT1+2	11	0.801 (0.752, 0.850)	<0.001	<0.001	0.648 (0.560, 0.735)	0.004	0.322
	MT1+1C	11	0.812 (0.763, 0.860)	<0.001	<0.001	0.678 (0.589, 0.767)	<0.001	0.391
	MT2+1C	11	0.806 (0.758, 0.854)	<0.001	<0.001	0.653 (0.554, 0.753)	<0.001	0.600
	MT1+2+1C	12	0.818 (0.771, 0.865)	<0.001	<0.001	0.677 (0.587, 0.767)	<0.001	0.297
Remerged prognostic model	rMT1	7	0.780 (0.726, 0.834)	<0.001	<0.001	0.653 (0.570, 0.736)	<0.001	0.843
	rMT2	7	0.773 (0.719, 0.827)	<0.001	<0.001	0.612 (0.508, 0.716)	0.046	0.964
	rMT1C	7	0.789 (0.739, 0.839)	<0.001	<0.001	0.671 (0.576, 0.766)	<0.001	0.893
	rMT1+2	8	0.799 (0.749, 0.848)	<0.001	<0.001	0.661(0.573, 0.748)	<0.001	0.480
	rMT1+1C	8	0.809 (0.761, 0.857)	<0.001	<0.001	0.685 (0.597, 0.774)	<0.001	0.663
	rMT2+1C	8	0.804 (0.757, 0.851)	<0.001	<0.001	0.665 (0.566, 0.763)	<0.001	0.874
	rMT1+2+1C	9	0.815 (0.769, 0.862)	<0.001	<0.001	0.683 (0.592, 0.775)	<0.001	0.385

Note.—C-index = concordance index; DMFS = distant metastasis-free survival; CI = confidence interval. MT1, MT2, MT1C, MT1+T2, MT1+T1C, MT2+T1C and MT1+T2+T1C prognostic models were built, and they integrated clinical risk factors (T stage, N stage, and plasma EBV DNA) with the T1, T2, T1C, T1+T2, T1+T1C, T2+T1C and T1+T2+T1C radiomics prognostic models, respectively. rMT1, rMT2, rMT1C, rMT1+T2, rMT1+T1C, rMT2+T1C and rMT1+T2+T1C prognostic models were built based on N stage, plasma EBV DNA with the T1, T2, T1C, T1+T2, T1+T1C, T2+T1C and T1+T2+T1C radiomics prognostic models, respectively.

†P-values were calculated compared with the clinical prognostic model.

**P-values were calculated by comparing with the corresponding radiomics prognostic model. For example, P=.0511 was the result of the comparison between the MT1 model and the T1 model in the validation cohort.*

Nomograms and calibration curves for the clinical and T1+T1C radiomics prognostic models are shown in **Figure 1**. The T1+T1C radiomics prognostic model predicted post-treatment metastasis risk for patients with a 3-year DMFS probability ≥40% or 5-year DMFS probability ≥30%, while the clinical prognostic model required a 3-year DMFS probability ≥70% or 5-year DMFS probability ≥60% to predicted metastasis risk (**Figures 1A1, A2**). Calibration curves in the probability of 5-year DMFS by a 1000-iteration bootstrap resampling experiment in the validation cohort also showed the T1+1C prognostic model had better discriminating ability than the clinical prognostic model. Observations showed better agreement with model predictions when the probability of 5-year DMFS was ≥68.9% in the T1+1C prognostic model, while it had to be ≥83.3% in the clinical prognostic model (**Figures 1B2, B4**).

We calculated the risk scores in the clinical and T1+T1C prognostic models for each patient and then classified patients into the low- and high-risk groups, with zero as risk score cutoff. The distributions of risk scores and 5-year DMFS in the low- and high-risk groups are shown in **Figure 2**. As shown, post-treatment metastatic patients were concentrated in the high score area, and the survival curve of the T1+T1C prognostic model showed good prognostic stratification for patients in the low- and high-risk groups in the validation cohort, while such trends were not observed in the validation cohort of the clinical prognostic model. When maximally selected rank statistics were used to generate the optimal risk score cutoff value for the clinical and T1+T1C prognostic models to divide patients into low- and high-risk groups, the results were identical to when zero was used as cutoff value (**Supplementary Figures S4, 5**).

The decision curve analysis of the validation cohort for the clinical prognostic model, radiomics models T1, T2, T1C, T1+T1C, T1+T2+T1C, MT1+T1C model, and rMT1C model is shown in **Figure 3** (other prognostic models are not listed). The T1+T1C prognostic model provided the best performance, while the clinical prognostic model showed no net benefit for patients with a 5-year DMFS probability ≥14.9%.

DISCUSSION

Our radiomics prognostic models, especially the T1+T1C prognostic model, had better discriminating ability than either the clinical prognostic model or corresponding merged prognostic models integrating clinical factors with selected radiomics features. All seven radiomics prognostic models could stratify patients into high- and low-risk groups based on significantly different 5-year DMFS rates. To our knowledge, this is the first study to clarify the role of multiple versus single MRI sequence radiomics, of radiomics features versus clinical factors, or the combination of clinical factors and radiomics features to

TABLE 3 | Multivariable analysis of radiomic features for the primary cohort.

Prognostic model	Variable	DMFS		
		coefficient	HR (95% CI)	*P*
T1 radiomics prognostic model	T1_shape_Sphericity	-4.67	0.01 (1.78E-04, 0.49)	0.021
	T1_W_{LHH}_GLSZM_LGLZE	-1.47	0.23 (0.04, 1.37)	0.107
	T1_W_{HHL}_GLCM_IMC2	-6.33	1.78E-03 (5.51E-05, 5.76E-02)	< 0.001
	T1_W_{HHH}_GLCM_IMC2	4.93	138.93 (2.39, 8.09E3)	0.017
	T1_W_{HLH}_GLCM_IMC2	-3.41	0.03 (6.41E-04, 1.71)	0.090
	T1_log.sigma.3.0.mm.3D_NGTDM_Strength	1.36	3.91 (1.19, 12.89)	0.025
	T1_W_{HHH}_NGTDM_Contrast	-14.64	4.37E-07(3.20E-16, 595.36)	0.172
T2 radiomics prognostic model	T2_W_{HLL}_GLDM_LDHGLE	8.04E-05	1.00008 (1.000006, 1.000115)	0.034
	T2_log.sigma.5.0.mm.3D_FOS_Skewness	-0.45	0.64 (0.38, 1.06)	0.084
	T2_W_{HHL}_GLSZM_SALGLE	-23.50	6.25E-11 (6.92E-18, 5.64E-04)	0.004
	T2_logarithm_NGTDM_Coarseness	17.62	4.48E+07 (1.22E+02, 1.65E+13)	0.007
	T2_W_{LLH}_GLCM_IDMN	25.21	8.86E+10 (1.22E-01, 6.45E+22)	0.070
T1C radiomics prognostic model	T1C_W_{HLL}_GLCM_Correlation	6.59	7.30E+02 (6.08, 8.77E+04)	0.007
	T1C_W_{LLH}_GLSZM_SAHGLE	0.03	1.03 (1.01, 1.05)	0.010
	T1C_Gradient_GLCM_IMC1	9.33	1.13E+04 (0.40, 3.23E+08)	0.075
	T1C_Square_GLCM_Correlation	3.22	25.02 (1.95, 3.21E+02)	0.013
	T1C_Gradient_GLSZM_ZE	0.60	1.83 (0.93, 3.58)	0.079
	T1C_square_GLRLM_RE	-1.66	0.19 (0.06, 0.61)	0.005
	T1C_log.sigma.3.0.mm.3D_GLSZM_ZE	0.94	2.55 (1.07, 6.07)	0.035
	T1C_W_{LHL}_GLSZM_ZE	-1.05	0.35 (0.11, 1.06)	0.063
	T1C_gradient_GLSZM_GLN	0.04	1.04 (0.99, 1.08)	0.104
T1+1C radiomics prognostic model	T1 radscore*	0.59	1.80 (1.25, 2.59)	0.002
	T1C radscore*	0.75	2.12 (1.53, 2.94)	<0.001
T1+2+1C radiomics prognostic model	T1 radscore*	0.42	1.52 (1.04, 2.23)	0.031
	T2 radscore*	0.48	1.61 (1.14, 2.28)	0.007
	T1C radscore*	0.58	1.79 (1.26, 2.54)	0.001

CI, confidence interval; DMFS, distant metastasis free survival; HR, hazard ratio. Textural features should be decomposed into three-dimensional wavelet transform (8 decompositions), and the wavelet decompositions are labeled as W_{LLL}, W_{LLH}, W_{LHL}, W_{LHH}, W_{HLH}, W_{HHL}, W_{HHL}, and W_{HHH}. T1-w = T1-weighted; T2-w = T2-weighted; T1C-w = contrast-enhanced T1-weighted. GLSZM, Gray Level Size Zone Matrix; GLCM, Gray Level Co-occurrence Matrix; NGTDM, Neighbouring Gray Tone Difference Matrix; GLDM, Gray Level Dependence Matrix; FOS, First order statistics; GLRLM, Gray Level Run Length Matrix; LGLZE, Low Gray Level Zone Emphasis; IMC, Informational measure of correlation; JA, Joint Average; LDHGLE, Large Dependence High Gray Level Emphasis; SALGLE, Small Area Low Gray Level Emphasis; IDMN, Inverse Difference Moment Normalized; SAHGLE, Small Area High Gray Level Emphasis; ZE, Zone Entropy; RE, Run Entropy; GLN, Gray Level Non-Uniformity.

**radscore from nomogram.*

predict prognosis with such a large sample size of nasopharyngeal carcinoma patients. Our findings could be useful for the evaluation of individual distant metastasis risk and the choice of a personalized therapeutic regimen in nasopharyngeal carcinoma patients.

The TNM staging system is routinely used in clinical practice to guide assessment of individual prognosis and treatment strategy decisions. However, an obvious shortcoming of the TNM staging system is that its basis on anatomical tumor extent does not reflect intra-tumor heterogeneity, which has pronounced effects on tumor diagnosis and prognosis. Plasma EBV DNA is a potential biomarker for NPC clinical management, but the lack of a global standardized testing methodology limits its widespread value (9). Since NPC is spatially and temporally heterogeneous (32), plasma EBV DNA does not completely characterize tumor heterogeneity. In this study, the C-index of the clinical prognostic model was 0.736 in the primary cohort, while it was significantly reduced to 0.552 in the validation cohort, indicating that clinical risk factors were inadequate for distinguishing between different metastasis risk groups and not a reliable method for predicting DMFS in NPC patients.

Radiomics transforms tumor images into detailed quantifications of tumor characteristics to find possible prognostic information for patients (16). Our study showed that radiomics features could be used as biomarkers for DMFS prediction in NPC patients without considering the TNM staging system or other clinical risk factors. There might be two reasons for this result. First, clinical risk factors could not distinguish between different metastasis risks; therefore, adding clinical factors into the radiomics prognostic models could not significantly improve the models' prediction efficiency. Second, the selected radiomics features might contain prognostic information hidden in clinical factors, and overfitting occurrence might weaken their prediction performance when integrating clinical factors with radiomics prognostic models. Zhang et al. (21) showed that the C-index of a radiomics prognostic model using the TNM staging system or clinical factors (based on age, sex, and hemoglobin) was reduced by 0.009 or 0.013 in the validation cohort, respectively.

T1-w, T1C-w, and T2-w are three common methods used in MRI tumor diagnosis. In single-MRI prognostic models for DMFS, the T1C prognostic model had the highest discriminating ability in our study, followed by the T1 and T2, in both primary and validation cohorts. Differences in discriminating ability may be due to imaging capabilities. T1-w can detect tumor anatomical details; T2-w images are sensitive for detecting effusion or edema; T1C-w images may reflect tumor

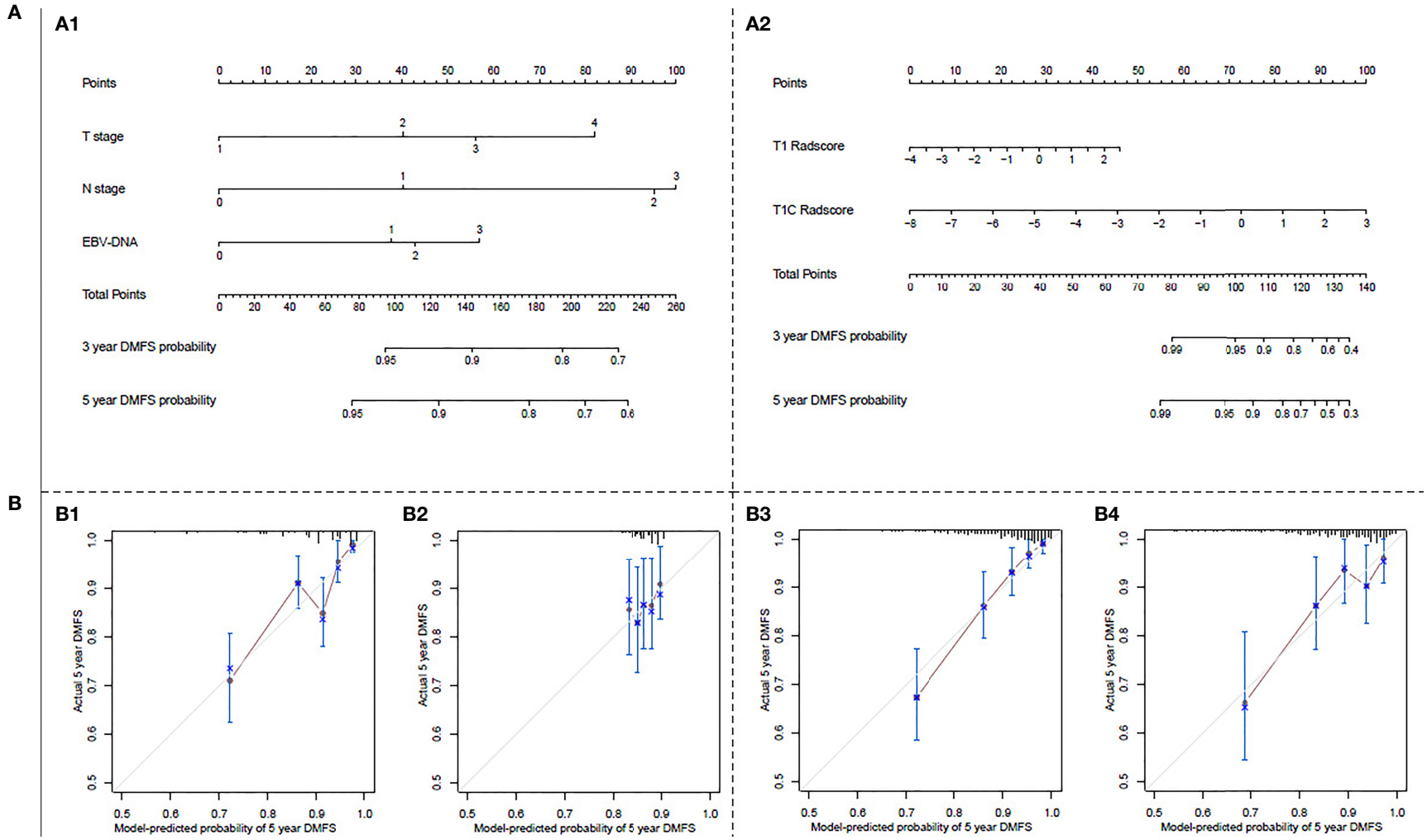


FIGURE 1 | **(A)** Nomograms for 3- and 5-year distant metastasis-free survival (DMFS): (A1) for the clinical prognostic model, (A2) for the T1+T1C prognostic model. The nomogram allows the user to obtain the probability of 3- and 5-year DMFS corresponding to a patient combination of covariates. As an example, locate the patient T stage and draw a line straight upward to the "Points" axis to determine the score associated with that T stage. Repeat the process for each variable and sum of the scores achieved for each covariate; after that, locate this sum on the "Total Points" axis. Draw a line straight down to determine the likelihood of 3- or 5-year DMFS. **(B)** Calibration curves for predicting 5-year DMFS: (B1) in the primary cohort of the clinical prognostic model, (B2) in the validation cohort of the clinical prognostic model, (B3) in the primary cohort of T1+T1C prognostic model, (B4) in the validation cohort of T1+T1C prognostic model. The Y-axis shows observed survival estimated by the Kaplan–Meier method, and the X-axis shows predicted survival calculated using the prognostic model. The solid lines represent the ideal reference line for which predicted survival corresponds with actual survival.

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-83250-751-3
DOI 10.3389/978-2-83250-751-3

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

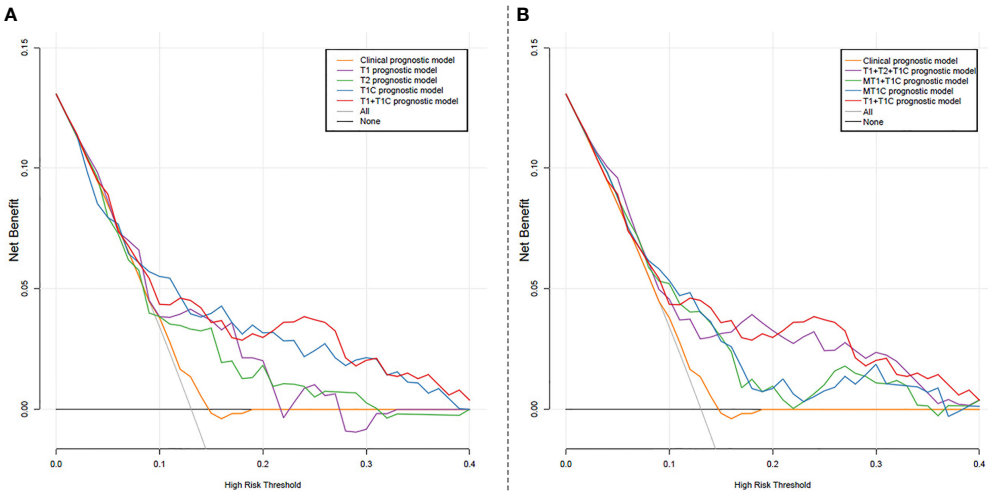


FIGURE 3 | Decision curve analysis for the prognostic models: **(A)** Clinical prognostic model, and radiomics models T1, T2, T1C, T1+T1C; **(B)** Clinical prognostic model, radiomics models T1+T2+T1C, T1+T1C, MT1+T1C model, and rMT1C mode. The X-axis represents the probability of the 5-year distant metastasis-free survival ranging from 0 to 100%. The Y-axis shows the net benefit. The black line indicates that no distant metastasis occurred in all patients. The gray line represents the assumption that all patients developed distant metastasis.

angiogenesis, closely related to tumor invasion and metastasis. For multiple-MRI prognostic models for DMFS, the T1+T1C prognostic model had the best performance in the validation cohort, with a higher C-index than the T1+T2, T2+T1C or T1+T2+T1C prognostic models. Among all prognostic models, the T1+T1C prognostic model also showed the best performance in the decision curve analysis of the validation cohort. Therefore, the radiomics features from T1-w and T1C-w MRI sequences better characterize metastatic biological behaviors than other combinations of the three MRI modalities. Except for one shape fature (sphericity), the other 15 features of the composition of T1+T1C radiomics model are textural features, which reflect tumor heterogeneity. This indicates a radiomic signature of some textural features can be reguarded as a potential prognostic biomark to discriminate different metastasis risk groups in NPC patients.

Our study had several limitations. First, the patients in the primary and validation cohorts were enrolled from the same institution, which may reduce the generalizability of our findings. Second, follow-up was incomplete for some patients, which may have led to distortions. Third, although PyRadiomics adheres for the most part to the Image Biomarker Standardization Initiative (IBSI), some small differences have been duly noted for it in terms of quantifying the grey level values, removing features associated with the ROI volum from the very beginning for example. Fourth, other MRI-based radiomics were not included, especially diffusion-weighted imaging and dynamic contrast-enhanced, significant for treatment monitoring and outcome prediction in different cancer types (33–35). Thus, the usefulness of diffusion-weighted imaging and dynamic contrast-enhanced MRI-based radiomics in NPC patients should be further explored.

In conclusion, a T1+T1C radiomics prognostic model could be a reliable approach for individual risk discriminating of distant metastasis in nasopharyngeal carcinoma patients. This model provided better prognostic performance than other radiomics prognostic models, clinical risk factors, and combination models, and may facilitate personalized risk stratification and treatment strategies for patients with nasopharyngeal carcinoma. To expand the generalizability of this prognostic model, validation with data from prospective, large sample, multicenter studies is required.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Medical Ethics Committee of Sun Yat-sen University Cancer center. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

H-JL, L-ZL, YH, and G-YZ designed the research. H-JL and G-YZ performed the data analysis and drafted the manuscript.

Y-BJ, X-PC, WL, J-CS, KC, and JZ provided scientific and technical support. H-JL, L-ZL and G-YZ collected clinical data. H-JL, YH, and G-YZ critically revised the manuscript. All authors critically reviewed the article and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 81572652), Science and Technology Planning Program of Guangdong Province (No. 2017A020215128), Science and Technology Planning Project of Guangzhou City (No. 201907010043), Foshan city climbing peak plan (No. 2019A009), and Science and Technology Innovation Platform of Foshan City (No. FS0AA-KJ218-1301-0007). The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

ACKNOWLEDGMENTS

We would like to thank Editage (www.editage.com) for English language editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.794975/full#supplementary-material

REFERENCES

1. Wee JT, Ha TC, Loong SL, Qian CN. Is Nasopharyngeal Cancer Really a "Cantonese Cancer"? *Chin J Cancer* (2010) 29(5):517–26. doi: 10.5732/cjc.009.10329
2. Jemal A, Bray F, Center MM, et al. Global Cancer Statistics. *CA Cancer J Clin* (2011) 61(2):69–90. doi: 10.3322/caac.20107
3. Lai SZ, Li WF, Chen L, et al. How Does Intensity-Modulated Radiotherapy Versus Conventional Two-Dimensional Radiotherapy Influence the Treatment Results in Nasopharyngeal Carcinoma Patients? *Int J Radiat Oncol Biol Phys* (2011) 80(3):661–8. doi: 10.1016/j.ijrobp.2010.03.024
4. Du TF, Xiao JF, Qiu ZL, Wu K. The Effectiveness of Intensity-Modulated Radiation Therapy Versus 2D-RT for the Treatment of Nasopharyngeal Carcinoma: A Systematic Review and Meta-Analysis. *PloS One* (2019) 14(7):e0219611. doi: 10.1371/journal.pone.0219611
5. Mao YP, Tang LL, Chen L, et al. Prognostic Factors and Failure Patterns in Non-Metastatic Nasopharyngeal Carcinoma After Intensity-Modulated Radiotherapy. *Chin J Cancer* (2016) 35(1):103. doi: 10.1186/s40880-016-0167-2
6. Chua DT, Ma J, Sham JS, et al. Long-Term Survival After Cisplatin Based Induction Chemotherapy and Radiotherapy for Nasopharyngeal Carcinoma: A Pooled Data Analysis of Two Phase III Trials. *J Clin Oncol* (2005) 23(6):1118–24. doi: 10.1200/JCO.2005.12.081
7. Mao YP, Xie FY, Liu LZ, et al. Re-Evaluation of 6th Edition of AJCC Staging System for Nasopharyngeal Carcinoma and Proposed Improvement Based on Magnetic Resonance Imaging. *Int J Radiat Oncol Biol Phys* (2009) 73(5):1326–34. doi: 10.1016/j.ijrobp.2008.07.062
8. Tang LL, Chen YP, Mao YP, et al. Validation of the 8th Edition of the UICC/AJCC Staging System for Nasopharyngeal Carcinoma From Endemic Areas in the Intensity-Modulated Radiotherapy Era. *J Natl Compr Canc Netw* (2017) 15(7):913–9. doi: 10.6004/jnccn.2017.0121
9. Lin JC, Wang WY, Chen KY, et al. Quantification of Plasma Epstein-Barr Virus DNA in Patients With Advanced Nasopharyngeal Carcinoma. *N Engl J Med* (2004) 350(24):2461–70. doi: 10.1056/NEJMoa032260
10. Zhou GQ, Tang LL, Mao YP, et al. Baseline Serum Lactate Dehydrogenase Levels for Patients Treated With Intensity-Modulated Radiotherapy for Nasopharyngeal Carcinoma: A Predictor of Poor Prognosis and Subsequent Liver Metastasis. *Int J Radiat Oncol Biol Phys* (2012) 82(3):e359–65. doi: 10.1016/j.ijrobp.2011.06.1967
11. Tang LQ, Hu DP, Chen QY, et al. Elevated High-Sensitivity C-Reactive Protein Levels Predict Decreased Survival for Nasopharyngeal Carcinoma Patients in the Intensity-Modulated Radiotherapy Era. *PloS One* (2015) 10(4):e0122965. doi: 10.1371/journal.pone.0122965
12. Weng JJ, Wei JZ, Li M, et al. Effects of Hepatitis B Virus Infection and Antiviral Therapy on the Clinical Prognosis of Nasopharyngeal Carcinoma. *Cancer Med* (2020) 9(2):541–51. doi: 10.1002/cam4.2715
13. Chen C, Shen LJ, Li BF, Gao J, Xia YF. Smoking is a Poor Prognostic Factor for Ale Nasopharyngeal Carcinoma Treated With Radiotherapy. *Radiother Oncol* (2014) 110(3):409–15. doi: 10.1016/j.radonc.2013.08.003
14. Tang XR, Li YQ, Liang SB, et al. Development and Validation of a Gene Expression-Based Signature to Predict Distant Metastasis in Locoregionally Advanced Nasopharyngeal Carcinoma: A Retrospective, Multicentre, Cohort Study. *Lancet Oncol* (2018) 19(3):382–93. doi: 10.1016/S1470-2045(18)30080-9
15. Liu N, Chen NY, Cui RX, et al. Prognostic Value of a microRNA Signature in Nasopharyngeal Carcinoma: A microRNA Expression Analysis. *Lancet Oncol* (2012) 13(6):633–41. doi: 10.1016/S1470-2045(12)70102-X
16. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169
17. Dong D, Fang MJ, Tang L, et al. Deep Learning Radiomic Nomogram can Predict the Number of Lymph Node Metastasis in Locally Advanced Gastric Cancer: An International Multicenter Study. *Ann Oncol* (2020) 31(7):912–20. doi: 10.1016/j.annonc.2020.04.003
18. Zhao LN, Gong J, Xi YB, et al. MRI-Based Radiomics Nomogram May Predict the Response to Induction Chemotherapy and Survival in Locally Advanced Nasopharyngeal Carcinoma. *Eur Radiol* (2020) 30(1):537–46. doi: 10.1007/s00330-019-06211-x
19. Wang GY, He L, Yuan C, Huang Y, Liu Z, Liang C. Pretreatment MR Imaging Radiomics Signatures for Response Prediction to Induction Chemotherapy in Patients With Nasopharyngeal Carcinoma. *Eur J Radiol* (2018) 98:100–6. doi: 10.1016/j.ejrad.2017.11.007
20. Zhang L, Zhou HY, Gu DS, et al. Radiomic Nomogram: Pretreatment Evaluation of Local Recurrence in Nasopharyngeal Carcinoma Based on MR Imaging. *J Cancer* (2019) 10:4217–25. doi: 10.7150/jca.33345
21. Zhang B, Tian J, Dong D, et al. Radiomics Features of Multiparametric MRI as Novel Prognostic Factors in Advanced Nasopharyngeal Carcinoma. *Clin Cancer Res* (2017) 23(15):4259–69. doi: 10.1158/1078-0432.CCR-16-2910
22. Zhang B, Ouyang FS, Gu DS, et al. Advanced Nasopharyngeal Carcinoma: Pre-Treatment Prediction of Progression Based on Multi-Parametric MRI Radiomics. *Oncotarget* (2017) 8(42):72457–65. doi: 10.18632/oncotarget.19799
23. Yang KX, Tian JF, Zhang B, et al. A Multidimensional Nomogram Combining Overall Stage, Dose Volume Histogram Parameters and Radiomics to Predict Progression-Free Survival in Patients With Locoregionally Advanced Nasopharyngeal Carcinoma. *Oral Oncol* (2019) 98:85–91. doi: 10.1016/j.oraloncology.2019.09.022
24. Ming X, Oei RW, Zhai R, et al. MRI-Based Radiomics Signature Is a Quantitative Prognostic Biomarker for Nasopharyngeal Carcinoma. *Sci Rep* (2019) 9(1):10412. doi: 10.1038/s41598-019-46985-0
25. Peng H, Dong D, Fang MJ, et al. Prognostic Value of Deep Learning PET/CT-Based Radiomics: Potential Role for Future Individual Induction Chemotherapy in Advanced Nasopharyngeal Carcinoma. *Clin Cancer Res* (2019) 25(14):4271–9. doi: 10.1158/1078-0432.CCR-18-3065

26. Zhuo EH, Zhang WJ, Li HJ, et al. Radiomics on Multi-Modalities MR Sequences Can Subtype Patients With Non-Metastatic Nasopharyngeal Carcinoma (NPC) Into Distinct Survival Subgroups. *Eur Radiol* (2019) 29 (10):5590–9. doi: 10.1007/s00330-019-06075-1
27. Lv WB, Yuan QY, Wang QS, et al. Radiomics Analysis of PET and CT Components of PET/CT Imaging Integrated With Clinical Parameters: Application to Prognosis for Nasopharyngeal Carcinoma. *Mol Imaging Biol* (2019) 21(5):954–64. doi: 10.1007/s11307-018-01304-3
28. Mao JJ, Fang J, Duan XH, et al. Predictive Value of Pretreatment MRI Texture Analysis in Patients With Primary Nasopharyngeal Carcinoma. *Eur Radiol* (2019) 29(8):4105–13. doi: 10.1007/s00330-018-5961-6
29. Zhang B, He X, Ouyang FS, et al. Radiomic Machine-Learning Classifiers for Prognostic Biomarkers of Advanced Nasopharyngeal Carcinoma. *Cancer Lett* (2017) 403:21–7. doi: 10.1016/j.canlet.2017.06.004
30. Lydiatt WM, Patel SG, O'Sullivan B, et al. Head and Neck Cancers-Major Changes in the American Joint Committee on Cancer Eighth Edition Cancer Staging Manual. *CA Cancer J Clin* (2017) 67(2):122–37. doi: 10.3322/caac.21389
31. Tang LQ, Li CF, Li J, et al. Establishment and Validation of Prognostic Nomograms for Endemic Nasopharyngeal Carcinoma. *J Natl Cancer Inst* (2015) 108(1):djv291. doi: 10.1093/jnci/djv291
32. Liang ZG, Tan HQ, Zhang F, et al. Comparison of Radiomics Tools for Image Analyses and Clinical Prediction in Nasopharyngeal Carcinoma. *Br J Radiol* (2019) 92(1102):20190271. doi: 10.1259/bjr.20190271
33. Kato H, Kanematsu M, Tanaka O, Mizuta K. Head and Neck Squamous Cell Carcinoma: Usefulness of Diffusion-Weighted MR Imaging in the Prediction of a Neoadjuvant Therapeutic Effect. *Eur Radiol* (2009) 19(1):103–9. doi: 10.1007/s00330-008-1108-5
34. Heethuis SE, Goense L, van Rossum PSN, et al. DW-MRI and DCE-MRI are of Complementary Value in Predicting Pathologic Response to Neoadjuvant Chemoradiotherapy for Esophageal Cancer. *Acta Oncol* (2018) 57(9):1201–8. doi: 10.1080/0284186X.2018.1473637
35. Lee SH, Hayano K, Zhu AX, Sahani DV, Yoshida H. Dynamic Contrast-Enhanced MRI Kinetic Parameters as Prognostic Biomarkers for Prediction of Survival of Patient With Advanced Hepatocellular Carcinoma: A Pilot Comparative Study. *Acad Radiol* (2015) 22(11):1344–60. doi: 10.1016/j.acra.2015.05.012

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Li, Liu, Huang, Jin, Chen, Luo, Su, Chen, Zhang and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Oncology

ORIGINAL RESEARCH
published: 31 March 2022
doi: 10.3389/fonc.2022.854253

Circulating MicroRNAs as Quantitative Biomarkers for Diagnosis and Prognosis of Uveal Melanoma

*Wen-Da Zhou†, Lei Shao†, Li Dong, Rui-Heng Zhang, Yi-Fan Li, He-Yan Li, Hao-Tian Wu, Xu-Han Shi and Wen-Bin Wei**

Beijing Tongren Eye Center, Beijing Key Laboratory of Intraocular Tumor Diagnosis and Treatment, Beijing Ophthalmology & Visual Sciences Key Lab, Medical Artificial Intelligence Research and Verification Key Laboratory of the Ministry of Industry and Information Technology, Beijing Tongren Hospital, Capital Medical University, Beijing, China

For uveal melanoma (UM) patients, it is significant to establish diagnosis and prognosis evaluation systems through imaging techniques. However, imaging examinations are short of quantitative biomarkers and it is difficult to finish early diagnosis of UM. In order to discover new molecular biomarkers for the diagnosis and prognostic evaluation of UM, six circulating miRNAs (mir-132-3p, mir-21-5p, mir-34a-5p, mir-126-3p, mir-199a-3p, mir-214-3p) were chosen as candidates for independent validation. Validation of these miRNAs was performed in a cohort of 20 patients, including 10 spindle-shaped melanoma and 10 epithelioid cell melanoma, and 10 healthy donors. Then 5 patients with metastatic UM were included to validate the performance of miRNAs in advanced UM. Serum levels of miRNAs were determined using quantitative real-time PCR. We confirmed significantly higher levels of three miRNAs in serum of UM patients in comparison to healthy controls, and miR-199a-3p had the best performance ($p < 0.0001$; AUC = 0.985). MiR-214-3p and miR-21-5p were significantly upregulated in serum of epithelioid cell melanoma patients compared to spindle-shaped melanoma patients and miR-132-3p and, conversely, were significantly downregulated in serum of epithelioid cell melanoma patients. MiR-21-5p shows their best performance ($p < 0.0001$; AUC = 0.980). Both miR-199a-3p and miR-21-5p showed great performance in advanced UM. Significantly higher levels of miR-21-5p ($p < 0.001$) were found in serum of metastatic UM patients compared to patients with localized spindle-shaped melanoma, and significantly higher levels of miR-199a-3p ($p < 0.001$) were detected in serum of metastatic UM patients compared to healthy controls. Our preliminary data indicate promising diagnostic utility of circulating miR-199a-3p and promising prognostic utility of circulating miR-21-5p in both early and advanced UM patients.

Keywords: prognostic, diagnostic, biomarker, microRNAs, uveal melanoma

OPEN ACCESS

Edited by:
Tingying Peng, Helmholtz zentrum München, Germany

Reviewed by:
Luca Falzone, G. Pascale National Cancer Institute Foundation (IRCCS), Italy
Dantong Wang, Helmholtz Association of German Research Centres (HZ), Germany

***Correspondence:**
Wen-Bin Wei
weiwenbintr@163.com

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *13 January 2022*
Accepted: *04 March 2022*
Published: *31 March 2022*

Citation:
Zhou W-D, Shao L, Dong L, Zhang R-H, Li Y-F, Li H-Y, Wu H-T, Shi X-H and Wei W-B (2022) Circulating MicroRNAs as Quantitative Biomarkers for Diagnosis and Prognosis of Uveal Melanoma. Front. Oncol. 12:854253. doi: 10.3389/fonc.2022.854253

INTRODUCTION

Uveal melanoma (UM) is the most common intraocular malignancy in adults and arises from melanocytes in the iris, ciliary body, or choroid. The clinical diagnosis of uveal melanoma is based on the slit lamp and indirect ophthalmoscope. Experienced retina specialists have no difficulty in diagnosing typical uveal melanoma based on the morphology of the tumor (1). However, in some cases, the diagnosis of UM will be challenging, especially when it is complicated with retinal detachment, refractive interstitial opacity, and small melanoma, which needs imaging techniques including ultrasonography and magnetic resonance imaging (MRI) to confirm the diagnosis. Metastatic melanoma was the leading single cause of death, and 40%–50% of patients with UM die from metastatic disease. Patients with metastatic disease frequently receive various treatments including surgical resection, chemotherapy, immunotherapy, and targeted therapy, but these treatments produce relatively modest health benefits (2).In spite that several cytogenetic features have been proved to have great value to estimate the prognosis of patients with UM, these methods are invasive and labor-intensive (2, 3). It is of great clinical significance to obtain *a priori* information of a poor prognosis through a non-invasive way. In summary, the current non-invasive examinations, mainly imaging techniques, are difficult to diagnose and evaluate the prognosis of UM at an early stage and it also lacks quantitative biomarkers. Therefore, novel non-invasive biomarkers for diagnosis and prognosis evaluation in very early stage are urgently needed.

MicroRNAs (miRNA) are small, non-coding, single-stranded RNA molecules that posttranscriptionally regulate gene expression by inhibiting or inactivating target messenger RNAs (4, 5). MiRNA expression patterns are frequently dysregulated in human tumors (6–10), including uveal melanoma (11, 12). Previous studies have shown that circulating miRNAs, detected in plasma or serum, could be an ideal class of blood-based biomarkers for cancer detection (13). The serum microRNA expression profile of tumor patients is significantly different from that of healthy people, and even tumor subtypes can be distinguished based on the serum microRNA expression profile (14). Unlike imaging examination, serum microRNAs change before tumors cause morphological changes, which makes circulating microRNAs an ideal non-invasive biomarker that can screen for tumors at an early stage (15, 16). In uveal melanoma, most studies on miRNA biomarkers have paid attention to the analysis of tumor tissue, while recently some studies have reported the diagnostic potential of circulating miRNAs (17–19). However, the prognostic utility of circulating miRNAs in uveal melanoma still needs to be proved and the diagnostic utility needs to be validated on larger cohorts of patients. Such studies are of great significance because they may provide quantitative biomarkers for the early diagnosis and prognostic evaluation of UM.

According to literature review, we finally selected six candidate miRNAs (miR-132-3p, miR-21-5p, miR-34a-5p, miR-126-3p, miR-199a-3p, miR-214-3p). In this study, we verified the level of circulating miRNAs in the plasm of UM patients with different histologic types to find a circulating biomarker for the early diagnosis and prognostic evaluation of UM.

METHODS AND MATERIALS

Selection of Candidate MicroRNAs

The selection of candidate microRNAs was based on the results of our previous experiments. The specimens of epithelioid cell melanoma (4 specimens) and spindle-shaped melanoma (4 specimens) of UM confirmed by histopathology and immunochemistry were collected in Beijing Tongren Hospital from March 2013 to October 2015. The expression profile of miRNA was assayed by miRNA array. Normal uveal specimens were obtained from 8 donors as controls. The differentially expressing miRNAs were screened by intergroup differential folds of ≥2. The microarray outcomes were validated by real-time quantitative PCR. The mutual upregulated miRNA in both spindle-shaped melanoma and epithelioid cell melanoma were miR-132-3p, miR-21-5p, miR-34a-5p, and miR-34b-5p, and mutual downregulated ones were miR-125b-2-3p, miR-126-3p, miR-199a-3p, and miR-214-3p (**Table 1**). Likely secreted miRNAs were selected from these miRNAs based on literature search.

Serum Sampling

We collected 20 blood samples from UM patients who underwent eyeball enucleation in Beijing Tongren Hospital, Capital Medical University, Beijing, China, between December 2017 and October 2019. All blood samples were collected before treatment. As of the last follow-up, 20 patients with localized UM were in good survival conditions, and no melanoma metastasis was found. The blood of 10 healthy subjects was collected as a control. All healthy subjects have no history of tumors or long-term medication history before collecting blood samples.

To further evaluate the applicability of the biomarkers in patients with advanced UM, the blood samples of 5 patients with

TABLE 1 | The results of real-time quantitative PCR performed on UM tissue samples.

miRNA	Spindle-shaped melanoma $2^{-\triangle\triangle ct}$(Mean ± SD)	*p* value	Epithelioid cell melanoma $2^{-\triangle\triangle ct}$(Mean ± SD)	*p* value
miR-132-3p	7.44 ± 1.98	<0.001	4.18 ± 0.99	<0.001
miR-21-5p	4.37 ± 1.02	0.049	3.41 ± 0.97	0.011
miR-34a-5p	8.93 ± 0.98	<0.001	9.2 ± 1.03	<0.001
miR-34b-5p	241.20 ± 7.02	<0.001	3.83 ± 0.54	<0.001
miR-125-2-3p	0.03 ± 0.01	<0.001	0.05 ± 0.01	<0.001
miR-126-3p	0.22 ± 0.01	0.002	0.15 ± 0.01	0.003
miR-199a-3p	0.25 ± 0.02	<0.001	0.12 ± 0.02	<0.001
miR-214-3p	0.12 ± 0.01	<0.001	0.02 ± 0.01	<0.001

metastatic UM undergoing primary ocular treatment in Beijing Tongren Hospital were collected from January 1, 2020, to September 1, 2020, as a verification set. All metastatic tumors were identified by imaging and pathological examination. Blood samples were collected before all primary ocular treatments.

The fasting venous blood was taken as a blood sample. The blood sample was placed in a vacuum tube at room temperature for half an hour, and then it was centrifuged at 1,200 rpm for 15 min at 4°C to obtain the serum sample. Serum was divided into aliquots and stored at -80°C until analysis. All patients and healthy donors were informed of the purpose of the study and asked to sign an informed consent form. This study was reviewed by the Ethics Committee of Capital Medical University.

Clinicopathological Data

All patients underwent thorough ophthalmologic evaluations before enucleation to confirm the diagnosis of UM and confirm that they did not meet the indications for plaque radiotherapy or local resection. Histological examination was performed at the Unit of Pathology, Beijing Tongren Hospital, Capital Medical University, Beijing, China. The clinicopathological data include the age at diagnosis, sex, location of the tumor, the maximum diameter of the tumor, the height of the tumor (the size of tumor was measured after enucleation), tumor cell type, and extrascleral extension. All patients that underwent imaging examination including MRI and ultrasonography before enucleation, the tumor size, ciliary body involvement, and extraocular extension were initially assessed by imaging techniques (**Figure 1**), and pathological examination was employed for further assessment after enucleation.

RNA Isolation

Total RNA enriched for small RNAs was isolated using miRcute Serum/Plasma Kit (Tiangen Biotech, Ltd., Beijing, China) from 200 µl of blood serum according to the manufacturer's protocol. The concentration and purity of RNA were determined using the NanoDrop ND-2000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The samples were either stored at –80°C or immediately processed.

qRT-PCR Quantification of miRNA Expression in Serum

Hifair® II 1st Strand cDNA Synthesis Kit (Yeasen Biotechnology Co., Ltd., Shanghai, China) was used for reverse transcription of 3 µl of RNA sample. Real-time PCR was performed using the Hieff® qPCR SYBR Green Master Mix (Yeasen Biotechnology Co., Ltd., Shanghai, China). The 20-µl PCR reaction mixture included 10 µl Hieff® qPCR SYBR Green Master Mix, 5 µl cDNA, and 1 µl of primer (0.5 µl forward primer and 0.5 µl reverse primer). Reactions were incubated in a 96-well optical plate at 95°C for 5 min, followed by 40 cycles at 95°C for 10 s, 60°C for 20 s, and 72°C for 20 s. All real-time PCR reactions were run in triplicates.

Statistical Analysis

The average levels of all measured miRNAs were normalized by the use of U6 and subsequently analyzed by the 2−ΔCt method. If the CT value is greater than 35, the data are unreliable and will not be adopted. Statistical analyses were performed on GraphPad Prism version 6 (GraphPad software, USA). Sensitivity, specificity, and area under curve (AUC) for miRNA levels were determined using receiver operator characteristic (ROC) analysis. Student's t test was adopted if the clinical-pathological data were normally distributed. We performed the Mann–Whitney-U-test and Fisher's exact test when the data were not normally distributed.

RESULTS

Clinicopathological Characteristics of Uveal Melanomas

A total of 20 patients with UM were included in the study, including 10 spindle-shaped melanomas and 10 epithelioid cell melanomas. In addition, the blood of 10 healthy donors was

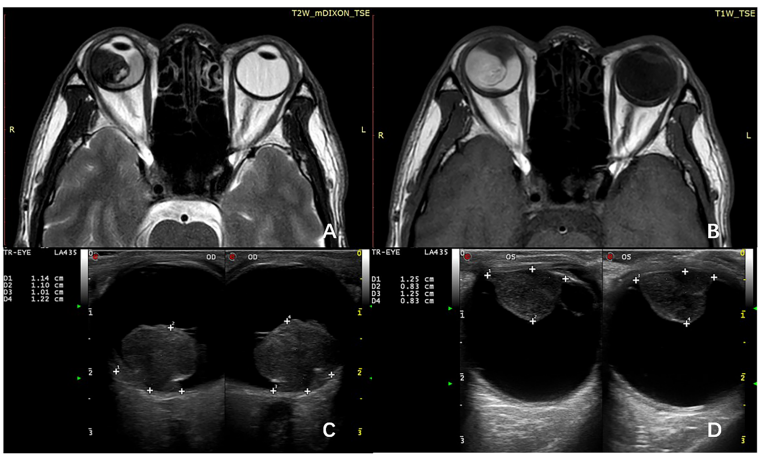


FIGURE 1 | MRI shows the primary tumor has spread to the ciliary body on T1-weighted image **(A)** and T2-weighted image **(B)**. B-scan ultrasonography can roundly measure the size of choroid melanoma **(C)** and ciliary body melanoma **(D)**.

collected as a normal control (**Table 2**). The average age of healthy donors was 41 ± 5.2 years, the average age of patients with spindle-shaped melanoma was 44.7 ± 4.5 years at the time of diagnosis, and the average age of patients with epithelioid cell melanoma was 40.9 ± 7.8 years at the time of diagnosis. The proportion of men in the three groups was 50.0%. There were no significant differences in clinicopathological parameters between patients with epithelioid cell melanoma and patients with spindle-shaped melanoma, including average tumor diameter (t = 0.075, $p = 0.941$), average tumor height (t = 0.561, $p = 0.582$), and tumor location ($p = 1.00$), extrascleral extension ($p = 1.00$), or systemic metastasis ($p = 1.00$). The clinicopathological data of participants with metastatic UM are given in **Table 3**. Most of the patients are male (80%), and the histological type is mainly epithelioid cell melanoma (80%). The average age, average tumor diameter, and average tumor height of metastatic patients are 52.8 ± 12.6 years, 15.3 ± 3.7 mm, and 11.7 ± 2.0 mm, respectively.

Imaging Features of Uveal Melanomas

Most of the UMs in this study show a typical pattern in B-scan ultrasonography, which appears as mushroom configuration, acoustic hollowness, and choroidal excavation (**Figure 2**). They also show a fairly typical pattern on T1-weighted images, demonstrating a bright signal relative to the vitreous compared with T2-weighted images, which provide a low tumor signal compared with the vitreous (**Figure 3**).

miRNA Expression in Serum of UM Patients

The normalized serum levels of miRNAs in UM patients and healthy controls are given in **Table 4**. The CT value of miR-34a-5p is greater than 35. The miR-21-5p (Z = 3.83, $p < 0.0001$), miR-199a-3p (Z = 4.27, $p < 0.0001$), and miR-132-3p (Z = 3.83, $p < 0.0001$) tested had significantly higher levels in serum of UM patients compared to healthy donors (**Figure 4**). The ROC curve was used to evaluate the diagnostic performance of the above three miRNAs (**Figure 5**), and the results showed that miR-199a-3p had the highest diagnostic performance, with an AUC of 0.985; when the normalized serum level of miR-199a-3p is 2.115, the sensitivity is 95% and the specificity is 100%.

The normalized serum levels of miRNAs in spindle-shaped melanoma and epithelioid cell melanoma are given in **Table 5**. The CT value of miR-34a-5p is greater than 35. The level of miR-132-3p (Z = 2.80, $p < 0.01$), miR-214-3p (Z = 3.25, $p < 0.001$), and miR-21-5p (Z = 3.63, $p < 0.0001$) in serum had a significant difference between spindle-shaped melanoma and epithelioid cell melanoma (**Figure 6**). The ROC curve was used to evaluate the ability to distinguish histologic types of the above three miRNAs (**Figure 7**), and the results showed that miR-21-5p had the best performance, with an AUC of 0.98. When the normalized serum level of miR-21-5p is 9.97, the sensitivity is 100% and the specificity is 90%.

As is depicted in **Figure 8**, significantly higher levels of miR-21-5p (Z = 2.51, $p < 0.001$) were found in serum of metastatic UM patients compared to patients with localized spindle-shaped melanoma. Moreover, beyond that, significantly higher levels of miR-199a-3p (Z = 3.07, $p < 0.001$) were detected in serum of metastatic UM patients compared to healthy donors.

DISCUSSION

In this study, we found that the serum levels of miR-199a-3p, miR-21-5p, and miR-132-3p are significantly increased in UM

TABLE 2 | Data of participants in this study.

Variable	Healthy controls	Spindle-shaped melanoma	Epithelioid melanoma		*p* value
Age (years)	41 ± 5.2	44.7 ± 4.5	40.9 ± 7.8		
Sex (male/female)	5/5	5/5	5/5		
Largest diameter (mm)		10.9 ± 6.2	10.7 ± 6.8	t = 0.075	0.941
Thickness (mm)		8.2 ± 3.3	9.4 ± 5.7	t = 0.561	0.582
Location (n, %)					
Choroid		9 (90.0)	8 (80.0)		1.000
Ciliary body		1 (10.0)	2 (20.0)		
Extrascleral extension (n, %)					
Yes		0 (0.0)	0 (0.0)		1.000
No		10 (100.0)	10 (100.0)		
Metastasis (n, %)					
Yes		0 (0.0)	0 (0.0)		
No		10 (100.0)	10 (100.0)		

TABLE 3 | Data of patients with metastatic UM.

Sex	Age	Location	Thickness (mm)	Largest diameter (mm)	Cell type	Extrascleral extension	Metastasis site
Female	49	Choriod	10.7	11.7	Spindle	No	Bone
Male	61	Choriod	9.8	19.1	Epithelioid	No	Liver
Male	56	Choriod	14	16.2	Epithelioid	No	Liver
Male	33	Choriod	10.4	11.2	Epithelioid	No	Liver
Male	65	Choriod	13.8	18.3	Epithelioid	No	Liver

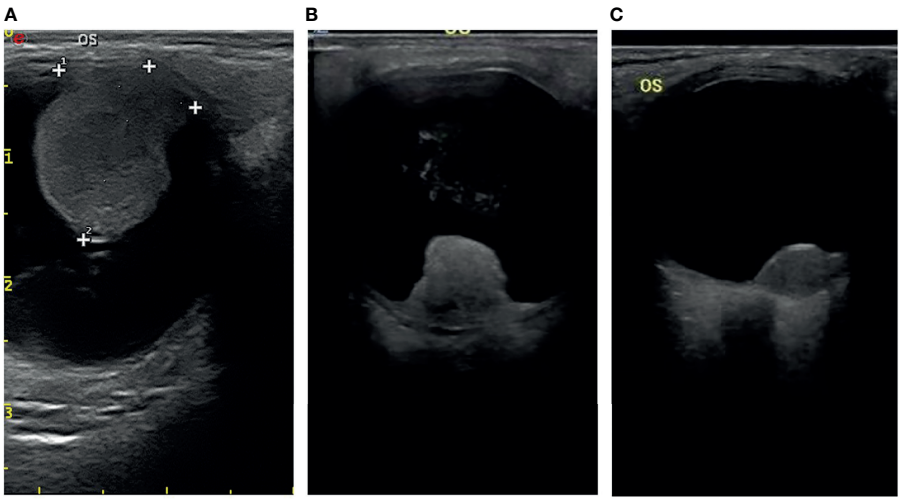


FIGURE 2 | The typical pattern of UM in B-scan ultrasonography **(A)**, mushroom configuration **(B)**, and acoustic hollowness **(C)** choroidal excavation.

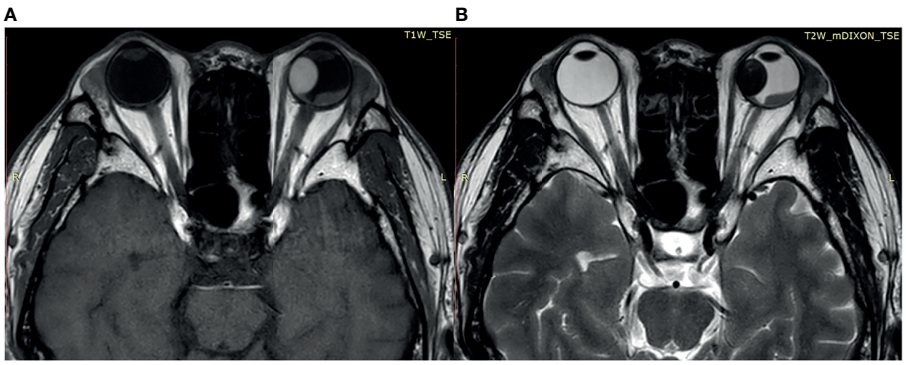


FIGURE 3 | **(A)** The tumor demonstrates a relatively bright signal on the T1-weighted image. **(B)** The tumor demonstrates a relatively low signal on the T2-weighted image.

TABLE 4 | The normalized serum levels of miRNAs in UM patients and healthy controls.

miRNAs	HC M (P25, P75)	UM M (P25, P75)	Wilcoxon Mann–Whitney	
			Z value	*p* value
miR-21-5p	0.78 (0.57, 2.00)	11.72 (3.64, 30.62)	3.83	<0.0001
miR-199a-3p	1.54 (0.48, 2.02)	6.28 (3.17, 7.98)	4.27	<0.0001
miR-132-3p	1.63 (0.25, 2.99)	6.59 (4.25, 9.81)	3.83	<0.0001
miR-126-3p	1.06 (0.47, 2.19)	1.44 (1.17, 2.19)	0.88	0.379
miR-214-3p	0.78 (0.68, 1.46)	0.55 (0.38, 0.76)	2.42	0.015

patients compared with healthy controls, and miR-199a-3p has the best diagnostic performance. We also found that the serum levels of miR-21-5p, miR-132-3p, and miR-214-3p differ significantly between epithelioid cell melanoma patients and spindle-shaped melanoma patients, and miR-21-5p has the best performance to distinguish histologic types. The diagnosis of UM is mainly primary diagnosis with a relatively high accuracy in the specialized ophthalmic center. In order to confirm the diagnosis of UM with non-invasive methods, a variety of imaging techniques can be employed (20, 21). Ultrasonography and MRI play an important role in the diagnosis of UM, especially in eyes with opaque media (22). Fundus fluorescein angiography (FFA) evaluates the intrinsic tumor vasculature which is useful in the differential diagnosis of choroidal lesions demonstrating different vascular patterns. Optical coherence tomography (OCT) is particularly useful to

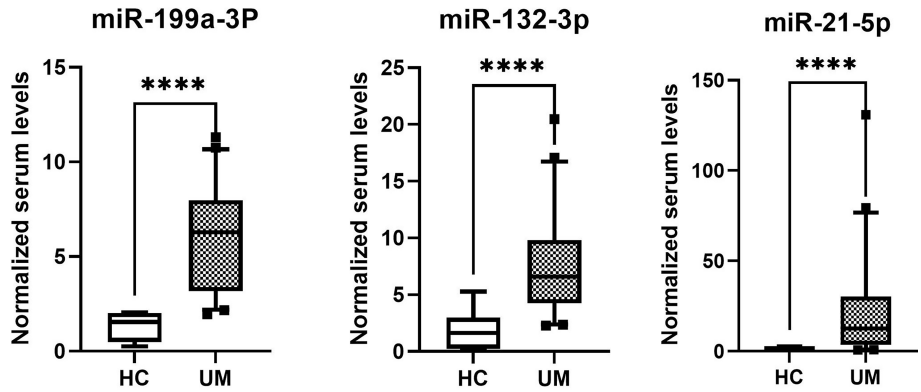


FIGURE 4 | MicroRNAs with significant different levels in serum of uveal melanoma (UM) patients and healthy controls (HC), ****$p < 0.0001$.

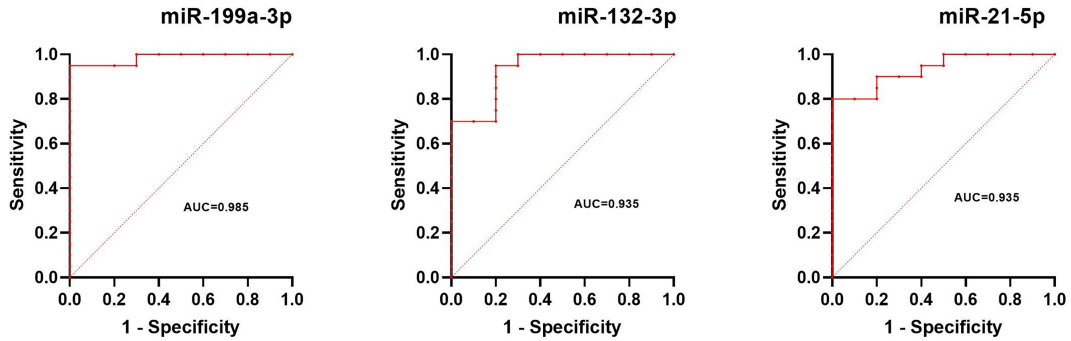


FIGURE 5 | Results of ROC analysis of miR-199a-3p, miR-132-3p, and miR-21-5p between UM patients and healthy controls.

TABLE 5 | The normalized serum levels of miRNAs in spindle-shaped melanoma and epithelioid cell melanoma.

miRNAs	Spindle-shaped melanoma M (P25, P75)	Epithelioid cell melanoma M (P25, P75)	Wilcoxon Mann–Whitney	
			Z value	*p* value
miR-21-5p	3.97 (1.76, 9.45)	29.98 (18.84, 59.17)	3.63	<0.0001
miR-199a-3p	6.80 (4.81, 9.88)	3.65 (2.34, 7.30)	4.27	0.07
miR-132-3p	9.1 (6.77, 14.45)	5.02 (2.55, 6.57)	2.80	<0.01
miR-126-3p	1.30 (0.96, 1.82)	1.57 (1.14, 2.65)	0.8	0.236
miR-214-3p	0.42 (0.30, 0.50)	0.75 (0.64, 1.40)	3.25	<0.001

image small melanomas which are 3 mm thick or less, and it is valuable for the early detection of UM, especially when considering factors for distinguishing a choroidal naevus from a UM (23). However, the diagnosis with imaging techniques is mainly based on the morphological features of UM, which is not obvious in the early stage (24, 25). Our study found that circulating miR-199a-3p can be used to effectively diagnose UM. Compared with healthy controls, miR-199a-3p, miR-21-5p, and miR-132-3p are significantly increased in the serum of UM patients, and miR-199a-3p has the best diagnostic performance. 20 patients with localized UM in this study were in good survival conditions after enucleation; the serum level of miR-199a-3p can increase in the early stage of UM. With the significantly higher levels of miR-199a-3p in serum of metastatic UM patients, the serum level of miR-199a-3p can also increase in the advanced UM. Compared with tumor biopsy, blood collection carries a low risk of procedure-related complications. MiR-199a-3p is a new, non-invasive, and quantitative biomarker that can be used for the early diagnosis of UM. Many studies have demonstrated that miR-199a-3p acts as antioncogene in a variety of tumors (26–29). Worely et al. (30) analyzed the miRNA expression patterns in 24 specimens of UM and miR-199a was significantly upregulated. However, in our previous study, the level of miR-199a-3p was significantly downregulated on formalin-fixed, paraffin-embedded UM samples compared with choroid tissue. On the one hand, the sample size was relatively small in our previous study, which may

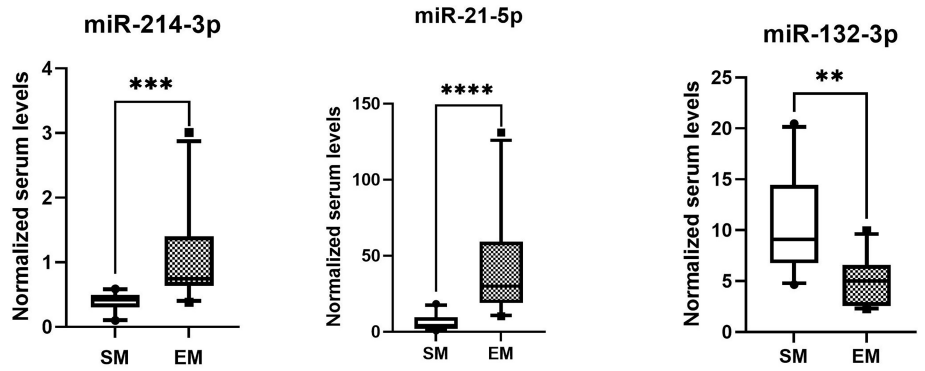


FIGURE 6 | MicroRNAs with significant difference in serum of spindle melanoma (SM) and epithelioid cell melanoma (EM) ****$p < 0.0001$, ***$p < 0.001$; **$p < 0.01$.

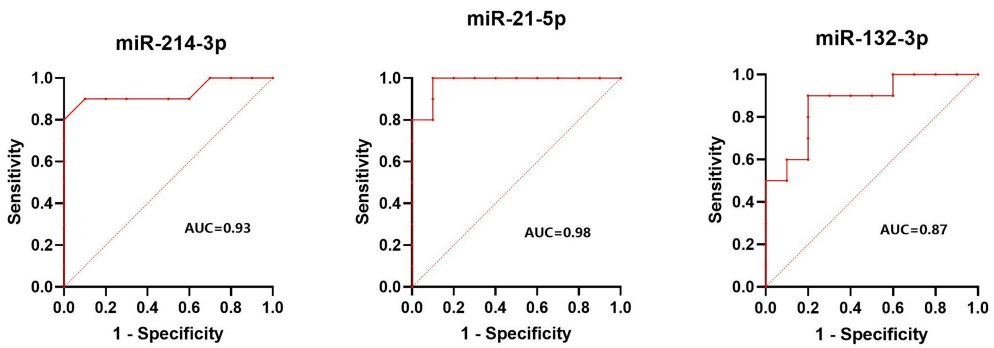


FIGURE 7 | Results of ROC analysis of miR-214-3p, miR-21-5p, and miR-132-3p between spindle melanoma and epithelioid cell melanoma.

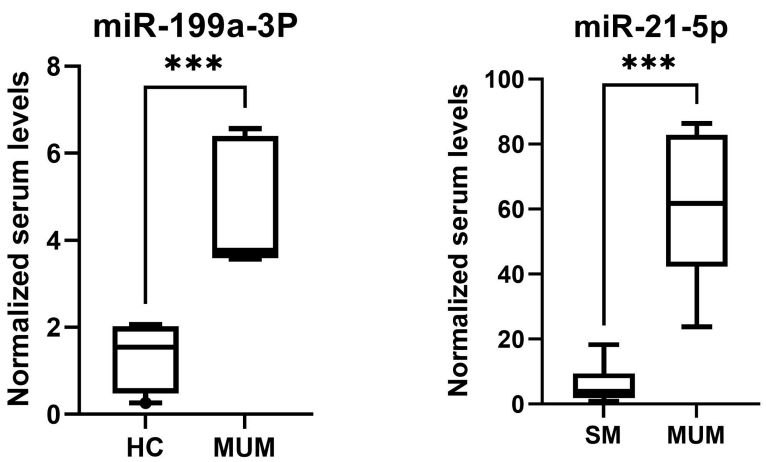


FIGURE 8 | The higher levels of miR-199a-3p of metastatic UM patients in comparison to healthy controls (HC) and higher levels of miR-21-5p of metastatic UM patients (MUM) in comparison to patients with localized spindle shaped melanoma (SM). ***$p < 0.001$.

lead to sampling errors. On the other hand, the circulating miRNAs do not completely reflect the tumor cellular profile (31). The exact molecular mechanisms about how tumor cells sort miRNAs to the blood remain quite elusive (32), and the role of miR-199a-3p in the UM and its underlying mechanism still need further study.

The tumor cells have disseminated in the very early stage, as primary ocular treatment of the eye with a UM does not prevent the development of these metastases (1). As a result, it is of great significance to assess the prognosis followed by the development of the best treatment. The prognosis of UM is related to histopathologic features, tumor size and location, extraocular extension, and genetic changes (33). Chromosomal alterations like monosomy 3 and gain of chromosome 8q and mutations like *BAP1*, *SF3B1*, and *EIF1AX* are strongly associated with high metastatic risk and poor survival (34). The abovementioned genetic changes have a clinical value as prognostic marker; however, the genetic testing is time-consuming and labor-intensive and the sample needs to be obtained by invasive biopsy, which makes it difficult to be applied in clinical practice. Tumor size, ciliary body involvement and extraocular extension are also important negative prognostic factors. These negative prognostic factors enable imaging examinations to evaluate the prognosis of UM at a limited level. Although it may cause radiation damage to patients, positron emission tomography/computed tomography (PET/CT) may have a role in predicting the prognosis of UM, since the standardized uptake value (SUV) may be correlated with risk factors for metastasis, like the proportion of epithelioid cell and tumor size (35). Unfortunately, imaging examinations cannot diagnose UM in an early stage and quantitatively evaluation of the prognosis of UM is either not possible. Circulating microRNA is an ideal blood-based biomarker. Studies have shown that the microRNA expression profiling of UM patients is significantly different from that of healthy controls and also has significant differences in UM with different prognosis (19, 30). Histologic types are the most commonly used method to assess the prognosis of patients after enucleation in routine clinic practice. Three histologic types of uveal melanoma include spindle-shaped melanoma, mixed melanoma, and epithelioid cell melanoma. Epithelioid melanoma is more malignant than the other two types, and the higher the proportion of epithelial cells, the more malignant the UM (36, 37). We found that the level of miR-21-5p in the blood of epithelioid cell melanoma patients was significantly increased compared with spindle-shaped melanoma. In our previous study, miR-21-5p was significantly upregulated in epithelioid cell melanoma tissue and spindle-shaped melanoma tissue, which was consistent with the significant upregulation of miR-21-5p in the serum of UM patients. MiR-21-5p acts as an oncogene in many tumors, and it has been shown to be involved in the epithelial–mesenchymal transition necessary for metastasis (38–40). Other research showed that the level of miR-21-5p is significantly increased in high-risk UM tissues, which implies that it can be responsible for UM metastasis (41). Since one miRNA can target a large number of genes and most studies use target genes identified by online prediction algorithms, it is important to perform validation experiments in UM cell lines (12, 31, 41). The validation experiment of miR-21 has already been performed in UM cell lines. It may inhibit the expression of wild-type p53 genes and increase the expression of *LASP1* at the mRNA level and protein level, which promotes the proliferation and metastasis of UM (42). Our findings may provide a quantitative prognostic biomarker for UM patients.

Several limitations of this study should be noted. All our patients were from the same ophthalmic center, which did not cover persons from multiple ethnic and regional populations. Another limitation of this study is the small sample size of participants with UM. Prospective studies with a large sample size and comparing varying ethnic groups are required to further evaluate the performance of the biomarkers.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article Supplementary Material. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Capital Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

W-DZ, LS, and W-BW contributed to study conception. W-DZ, H-TW, R-HZ, H-YL, X-HS, and Y-FL contributed to collect the data of healthy controls and UM patients and get blood samples. W-DZ, LS, and LD performed the laboratory operation. W-DZ and LS wrote the initial draft of the manuscript. LS and LD contributed to editing of the present version. W-BW commented on the manuscript and gave final approval of the version to be published. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Capital Health Research and Development of Special (2020-1-2052); Science & Technology Project of Beijing Municipal Science & Technology Commission (Z201100005520045, Z181100001818003); the Priming Scientific Research Foundation for the Junior Researcher in Beijing Tongren Hospital, Capital Medical University (No.2018-YJJ-ZZL-045); and Beijing Outstanding Talent Project Youth Backbone Foundation (2018000021469G207).

REFERENCES

1. Jager MJ, Shields CL, Cebulla CM, Abdel-Rahman MH, Grossniklaus HE, Stern MH, et al. Uveal Melanoma. *Nat Rev Dis Primers* (2020) 6(1):24. doi: 10.1038/s41572-020-0158-0
2. Chattopadhyay C, Kim DW, Gombos DS, Oba J, Qin Y, Williams MD, et al. Uveal Melanoma: From Diagnosis to Treatment and the Science in Between. *Cancer* (2016) 122(15):2299–312. doi: 10.1002/cncr.29727
3. Kaliki S, Shields CL, Shields JA. Uveal Melanoma: Estimating Prognosis. *Indian J Ophthalmol* (2015) 63(2):93–102. doi: 10.4103/0301-4738.154367
4. Stefani G, Slack FJ. Small non-Coding RNAs in Animal Development. *Nat Rev Mol Cell Biol* (2008) 9(3):219–30. doi: 10.1038/nrm2347
5. Lin S, Gregory RI. MicroRNA Biogenesis Pathways in Cancer. *Nat Rev Cancer* (2015) 15(6):321–33. doi: 10.1038/nrc3932
6. Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs With a Role in Cancer. *Nat Rev Cancer* (2006) 6(4):259–69. doi: 10.1038/nrc1840
7. Calin GA, Croce CM. MicroRNA Signatures in Human Cancers. *Nat Rev Cancer* (2006) 6(11):857–66. doi: 10.1038/nrc1997
8. Falzone L, Grimaldi M, Celentano E, Augustin LSA, Libra M. Identification of Modulated MicroRNAs Associated With Breast Cancer, Diet, and Physical Activity. *Cancers (Basel)* (2020) 12(9):2555. doi: 10.3390/cancers12092555
9. Jiang D, Xie X, Lu Z, Liu L, Qu Y, Wu S, et al. Establishment of a Colorectal Cancer-Related MicroRNA-mRNA Regulatory Network by Microarray and Bioinformatics. *Front Genet* (2020) 11:560186. doi: 10.3389/fgene.2020.560186
10. Vannini I, Fanini F, Fabbri M. Emerging Roles of microRNAs in Cancer. *Curr Opin Genet Dev* (2018) 48:128–33. doi: 10.1016/j.gde.2018.01.001
11. Yang C, Wei W. The miRNA Expression Profile of the Uveal Melanoma. *Sci China Life Sci* (2011) 54(4):351–8. doi: 10.1007/s11427-011-4149-y
12. Falzone L, Romano GL, Salemi R, Bucolo C, Tomasello B, Lupo G, et al. Prognostic Significance of Deregulated microRNAs in Uveal Melanomas. *Mol Med Rep* (2019) 19(4):2599–610. doi: 10.3892/mmr.2019.9949
13. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, et al. Circulating microRNAs as Stable Blood-Based Markers for Cancer Detection. *Proc Natl Acad Sci USA* (2008) 105 (30):10513–8. doi: 10.1073/pnas.0804549105
14. Cui M, Wang H, Yao X, Zhang D, Xie Y, Cui R, et al. Circulating MicroRNAs in Cancer: Potential and Challenge. *Front Genet* (2019) 10:626. doi: 10.3389/fgene.2019.00626
15. Crimi S, Falzone L, Gattuso G, Grillo CM, Candido S, Bianchi A, et al. Droplet Digital PCR Analysis of Liquid Biopsy Samples Unveils the Diagnostic Role of hsa-miR-133a-3p and hsa-miR-375-3p in Oral Cancer. *Biology* (2020) 9 (11):379. doi: 10.3390/biology9110379
16. Herreros-Villanueva M, Duran-Sanchon S, Martín AC, Pérez-Palacios R, Vila-Navarro E, Marcuello M, et al. Plasma MicroRNA Signature Validation for Early Detection of Colorectal Cancer. *Clin Trans Gastroenterol* (2019) 10(1): e00003. doi: 10.14309/ctg.0000000000000003
17. Ragusa M, Barbagallo C, Statello L, Caltabiano R, Russo A, Puzzo L, et al. miRNA Profiling in Vitreous Humor, Vitreal Exosomes and Serum From Uveal Melanoma Patients: Pathological and Diagnostic Implications. *Cancer Biol Ther* (2015) 16(9):1387–96. doi: 10.1080/15384047.2015.1046021
18. Russo A, Caltabiano R, Longo A, Avitabile T, Franco LM, Bonfiglio V, et al. Increased Levels of miRNA-146a in Serum and Histologic Samples of Patients With Uveal Melanoma. *Front Pharmacol* (2016) 7:424. doi: 10.3389/fphar.2016.00424
19. Stark MS, Gray ES, Isaacs T, Chen FK, Millward M, McEvoy A, et al. A Panel of Circulating MicroRNAs Detects Uveal Melanoma With High Precision. *Transl Vis Sci Technol* (2019) 8(6):12. doi: 10.1167/tvst.8.6.12
20. Foti PV, Travali M, Farina R, Palmucci S, Spatola C, Raffaele L, et al. Diagnostic Methods and Therapeutic Options of Uveal Melanoma With Emphasis on MR Imaging-Part I: MR Imaging With Pathologic Correlation and Technical Considerations. *Insights Into Imaging* (2021) 12(1):66. doi: 10.1186/s13244-021-01000-x
21. Kivelä T. Diagnosis of Uveal Melanoma. *Developments Ophthalmol* (2012) 49:1–15. doi: 10.1159/000330613
22. Fallico M, Raciti G, Longo A, Reibaldi M, Bonfiglio V, Russo A, et al. Current Molecular and Clinical Insights Into Uveal Melanoma (Review). *Int J Oncol* (2021) 58(4):10. doi: 10.3892/ijo.2021.5190
23. Kapoor A, Beniwal V, Beniwal S, Mathur H, Kumar HS. Management of Uveal Tract Melanoma: A Comprehensive Review. *J Egyptian Natl Cancer Institute* (2016) 28(2):65–72. doi: 10.1016/j.jnci.2016.02.003
24. Branisteanu DC, Bogdanici CM, Branisteanu DE, Maranduca MA, Zemba M, Balta F, et al. Uveal Melanoma Diagnosis and Current Treatment Options (Review). *Exp Ther Med* (2021) 22(6):1428. doi: 10.3892/etm.2021.10863
25. Bustamante P, Piquet L, Landreville S, Burnier JV. Uveal Melanoma Pathobiology: Metastasis to the Liver. *Semin Cancer Biol* (2021) 71:65–85. doi: 10.1016/j.semcancer.2020.05.003
26. Wu MM, Shen WD, Zou CW, Chen HJ, Guo HM. LncRNA-HEIH Suppresses Hepatocellular Carcinoma Cell Growth and Metastasis by Up-Regulating miR-199a-3p. *Eur Rev Med Pharmacol Sci* (2020) 24(11):6031–8. doi: 10.26355/eurrev_202006_21497
27. Callegari E, D'Abundo L, Guerriero P, Simioni C, Elamin BK, Russo M, et al. miR-199a-3p Modulates MTOR and PAK4 Pathways and Inhibits Tumor Growth in a Hepatocellular Carcinoma Transgenic Mouse Model. *Mol Ther Nucleic Acids* (2018) 11:485–93. doi: 10.1016/j.omtn.2018.04.002
28. Qu F, Zheng J, Gan W, Lian H, He H, Li W, et al. MiR-199a-3p Suppresses Proliferation and Invasion of Prostate Cancer Cells by Targeting Smad1. *Oncotarget* (2017) 8(32):52465–73. doi: 10.18632/oncotarget.17191
29. Phatak P, Burrows WM, Chesnick IE, Tulapurkar ME, Rao JN, Turner DJ, et al. MiR-199a-3p Decreases Esophageal Cancer Cell Proliferation by Targeting P21 Activated Kinase 4. *Oncotarget* (2018) 9(47):28391–407. doi: 10.18632/oncotarget.25375
30. Worley LA, Long MD, Onken MD, Harbour JW. Micro-RNAs Associated With Metastasis in Uveal Melanoma Identified by Multiplexed Microarray Profiling. *Melanoma Res* (2008) 18(3):184–90. doi: 10.1097/CMR.0b013e3282feeac6
31. Barbagallo C, Platania CBM, Drago F, Barbagallo D, Di Pietro C, Purrello M, et al. Do Extracellular RNAs Provide Insight Into Uveal Melanoma Biology? *Cancers (Basel)* (2021) 13(23):5919. doi: 10.3390/cancers13235919
32. Ragusa M, Barbagallo C, Cirnigliaro M, Battaglia R, Brex D, Caponnetto A, et al. Asymmetric RNA Distribution Among Cells and Their Secreted Exosomes: Biomedical Meaning and Considerations on Diagnostic Applications. *Front Mol Biosci* (2017) 4:66. doi: 10.3389/fmolb.2017.00066
33. Rand Simpson E, Gallie BL, Saakyan S, Amiryan A, Finger PT, Seregard S, et al. International Validation of the American Joint Committee on Cancer's 7th Edition Classification of Uveal Melanoma. *JAMA Ophthalmol* (2015) 133 (4):376–83. doi: 10.1001/jamaophthalmol.2014.5395
34. Decatur CL, Ong E, Garg N, Anbunathan H, Bowcock AM, Field MG, et al. Driver Mutations in Uveal Melanoma: Associations With Gene Expression Profile and Patient Outcomes. *JAMA Ophthalmol* (2016) 134(7):728–33. doi: 10.1001/jamaophthalmol.2016.0903
35. Finger PT, Chin K, Iacob CE. 18-Fluorine-Labelled 2-Deoxy-2-Fluoro-D-Glucose Positron Emission Tomography/Computed Tomography Standardised Uptake Values: A Non-Invasive Biomarker for the Risk of Metastasis From Choroidal Melanoma. *Br J Ophthalmol* (2006) 90 (10):1263–6. doi: 10.1136/bjo.2006.097949
36. Guo L, Wu Z, Zhang S, Chen J, Ai S, Zheng H. [Research on the Relationship Between Pathological Features of the Uveal Melanoma and Prognosis]. *Zhonghua bing li xue za zhi = Chin J Pathol* (2002) 31(6):518–21. doi: 10.3760/j.issn:0529-5807.2002.06.008
37. McLean IW, Foster WD, Zimmerman LE. Uveal Melanoma: Location, Size, Cell Type, and Enucleation as Risk Factors in Metastasis. *Hum Pathol* (1982) 13(2):123–32. doi: 10.1016/s0046-8177(82)80116-0
38. Li Q, Li B, Li Q, Wei S, He Z, Huang X, et al. Exosomal miR-21-5p Derived From Gastric Cancer Promotes Peritoneal Metastasis *via* Mesothelial-to-Mesenchymal Transition. *Cell Death Dis* (2018) 9(9):854. doi: 10.1038/s41419-018-0928-8
39. Ren W, Hou J, Yang C, Wang H, Wu S, Wu Y, et al. Extracellular Vesicles Secreted by Hypoxia Pre-Challenged Mesenchymal Stem Cells Promote Non-Small Cell Lung Cancer Cell Growth and Mobility as Well as Macrophage M2 Polarization *via* miR-21-5p Delivery. *J Exp Clin Cancer Res CR* (2019) 38 (1):62. doi: 10.1186/s13046-019-1027-0
40. Jin XH, Lu S, Wang AF. Expression and Clinical Significance of miR-4516 and miR-21-5p in Serum of Patients With Colorectal Cancer. *BMC Cancer* (2020) 20(1):241. doi: 10.1186/s12885-020-06715-6
41. Smit KN, Chang J, Derks K, Vaarwater J, Brands T, Verdijk RM, et al. Aberrant MicroRNA Expression and Its Implications for Uveal Melanoma Metastasis. *Cancers (Basel)* (2019) 11(6):815. doi: 10.3390/cancers11060815
42. Wang YC, Yang X, Wei WB, Xu XL. Role of microRNA-21 in Uveal Melanoma Cell Invasion and Metastasis by Regulating P53 and Its

Downstream Protein. *Int J Ophthalmol* (2018) 11(8):1258–68. doi: 10.18240/ijo.2018.08.03

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhou, Shao, Dong, Zhang, Li, Li, Wu, Shi and Wei. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 11 May 2022
doi: 10.3389/fonc.2022.833978

Automatic Tumor Grading on Colorectal Cancer Whole-Slide Images: Semi-Quantitative Gland Formation Percentage and New Indicator Exploration

Shenlun Chen[1,2†], Meng Zhang[3†], Jiazhou Wang[1†], Midie Xu[3], Weigang Hu[3], Leonard Wee[2], Andre Dekker[2], Weiqi Sheng[3‡] and Zhen Zhang[1*‡]*

[1] Department of Radiation Oncology, Fudan University Shanghai Cancer Center, Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China, [2] MAASTRO (Department of Radiotherapy), GROW School of Oncology and Developmental Biology, Maastricht University and Maastricht University Medical Centre+, Maastricht, Netherlands, [3] Department of Pathology, Fudan University Shanghai Cancer Center, Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China

OPEN ACCESS

Edited by:
Tingying Peng,
Helmholtz zentrum München,
Germany

Reviewed by:
Eldad Klaiman,
Roche, Germany
Hyun-Jong Jang,
Catholic University of Korea,
South Korea

***Correspondence:**
Zhen Zhang
zhen_zhang@fudan.edu.cn
Weiqi Sheng
shengweiqi2006@163.com270

[†]These authors have contributed equally to this work and share first authorship

[‡]These authors have contributed equally to this work and share senior authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 12 December 2021
Accepted: 21 March 2022
Published: 11 May 2022

Citation:
Chen S, Zhang M, Wang J, Xu M, Hu W, Wee L, Dekker A, Sheng W and Zhang Z (2022) Automatic Tumor Grading on Colorectal Cancer Whole-Slide Images: Semi-Quantitative Gland Formation Percentage and New Indicator Exploration. Front. Oncol. 12:833978. doi: 10.3389/fonc.2022.833978

Tumor grading is an essential factor for cancer staging and survival prognostication. The widely used the WHO grading system defines the histological grade of CRC adenocarcinoma based on the density of glandular formation on whole-slide images (WSIs). We developed a fully automated approach for stratifying colorectal cancer (CRC) patients' risk of mortality directly from histology WSI relating to gland formation. A tissue classifier was trained to categorize regions on WSI as glands, stroma, immune cells, background, and other tissues. A gland formation classifier was trained on expert annotations to categorize regions as different degrees of tumor gland formation versus normal tissues. The glandular formation density can thus be estimated using the aforementioned tissue categorization and gland formation information. This estimation was called a semi-quantitative gland formation ratio (SGFR), which was used as a prognostic factor in survival analysis. We evaluated gland formation percentage and validated it by comparing it against the WHO cutoff point. Survival data and gland formation maps were then used to train a spatial pyramid pooling survival network (SPPSN) as a deep survival model. We compared the survival prediction performance of estimated gland formation percentage and the SPPSN deep survival grade and found that the deep survival grade had improved discrimination. A univariable Cox model for survival yielded moderate discrimination with SGFR (c-index 0.62) and deep survival grade (c-index 0.64) in an independent institutional test set. Deep survival grade also showed better discrimination performance in multivariable Cox regression. The deep survival grade significantly increased the c-index of the baseline Cox model in both validation set and external test set, but the inclusion of SGFR can only improve the Cox model less in external test and is unable to improve the Cox model in the validation set.

Keywords: tumor grading, whole-slide histopathology image, colorectal cancer, deep learning, gland formation, Pathology and clinical outcomes

1 INTRODUCTION

Though there has been intense research interest in molecular prognostic and predictive factors, visual histology examination continues to be the most dependable predictor of survival in patients with colorectal cancer (CRC). Pathologists define the "reference standard" of tumor grading by evaluating visual characteristics, such as nuclear atypia, mitotic activity, and morphological features. New features have been proposed as histopathological prognostic factors, such as poorly differentiated clusters (PDCs) (1), tumor budding (2), and tumor-infiltrating lymphocytes (TILs) (3). In routine clinical practice, prognostication by pathologist-reported overall tumor grade remains the basis for "standard care" in CRC (4–6).

The objective of this study was to evaluate the potential performance of "deep learning" models for individualized survival time predictions directly from histopathology whole-slide images (WSIs) of CRC.

Over 90% of all CRC are adenocarcinoma arising in epithelial cells of colorectal mucosae. The widely used WHO visual grading system defines histological tumor grade based on the density of gland-forming regions (7) on WSI. Tumors are classified as well differentiated, moderately differentiated, poorly differentiated, or undifferentiated, depending on the percentage of the tumor that is gland forming, at >95%, 50%–95%, 5%–50%, and <5%, respectively; however, visual grading has some obvious limitations. First, visual estimation of gland formation (GF) density might be highly subjective (8), and grade assignment thus implies significant inter-observer variability (9–11). Second, quantitative calculation of GF is not presently performed during manual tumor grading due to the lack of tools that work on WSI. Third, the WHO tumor grade alone does not appear to be a sufficiently strong biomarker for personalized therapy decision making.

Numerous publications to date have demonstrated the potential application of deep learning convolutional neural networks (CNNs) for image-based tissue classification and survival prognostication tasks, including on CRC WSI (12, 13). Clinical parameters, such as TIL density and tumor–stroma ratio, have been quantified automatically and accurately using CNNs (12, 14) but fall short of predicting outcome. Other modeling works do predict outcomes of CRC patients, either from WSI or by first computing surrogate indicators (15, 16) such as the tumor–stroma ratio (12) and the proportion of different tissues types present in the WSI (13). These approaches revealed the deep connection between WSIs and outcomes, and we hypothesize that a direct approach applying deep learning on the entire WSIs may provide new prognostic markers beyond the traditional clinical parameters.

To date, no study has shown a survival prediction model on the WHO's grading system by either a direct or indirect approach. As mentioned above, the most widely accepted method of tumor grading is GF percentage. Therefore, we aim to predict survival outcomes on GF by both the clinical parameter (GF) and the deep learning-based biomarker. It is our hypothesis that the GF percentage can be estimated by an artificial intelligence (AI) model, and subsequently, an AI model can derive a better prognostic marker from spatial heatmaps of GF compared to the estimated GF percentage.

The organization of our work is as follows. First, we developed separate CNN models i) for gland-forming region classification (we hereafter refer to these as tumor differentiation spatial heatmaps) and ii) for automatic segmentation of tissue type regions on the entire WSI. Next, we trained a deep learning Cox proportional hazards model using the aforementioned spatial heatmaps and tissue type segmentation masks as input. We evaluated the prognostic performance among these alternative stratification options for the primary outcome of overall survival (OS). This work combines two well-known The Cancer Genome Atlas (TCGA) public datasets as training and internal validation, and then we added a new private institutional dataset of Chinese subjects as an independent external test set.

2 MATERIALS AND METHODS

2.1 Patients and Datasets

Training and validation datasets consisted of TCGA Colon Adenocarcinoma (TCGA-COAD) and Rectum Adenocarcinoma (TCGA-READ) that have been described in detail elsewhere. Briefly, the former contains data from 454 colon and 7 rectosigmoid junction cancer patients, with 983 WSIs (H&E-stained) and an optical magnification factor of 40 (0.25 μm/pixel) or 20 (0.5 μm/pixel). We chose the highest magnification possible to provide detailed data for AI classifiers and sampled tiles that contain enough information for classifying tissues types and GFs, which were confirmed by our pathologists. The latter comprises 364 H&E-stained WSIs derived from 91 rectal, 71 rectosigmoid junctions, and 6 colon cancer patients also with a magnification factor of 40. For the OS model construction, we excluded only WSIs that had no matching survival records. Thus, 857 WSI (out of 983) and 297 WSI (out of 364) from TCGA-COAD and TCGA-READ, respectively, were used for training. Our local institutional dataset consisted of 108 H&E-stained WSIs with a magnification factor of 40 from 108 eligible patients treated from 2008 to 2009. The WSIs in the local dataset were all formalin-fixed paraffin-embedded (FFPE) slides rather than frozen slides in TCGA dataset for validating survival outcomes better. The inclusion criteria of the institutional dataset were as follows: i) follow-up data were available, ii) epithelial tissues covered an area of at least 20% on WSI, and iii) the WSI was clear enough for human grading and free from large contaminated regions.

Written informed consent from all institutional patients (050432-4-1911D) and ethics board approval had been obtained. The overall training procedure is as described sequentially in the following subsections and is supported by the illustration in **Figures 1A, B**.

2.2 Acquisition of Whole-Slide Image Classification Heatmaps

Spatial heatmaps of tissue classification and GF classification were necessary to obtain the estimated tumor grading and develop the survival prediction models. Two independent supervised deep learning classifiers were trained to generate these spatial heatmaps. To segment new WSI, we implemented

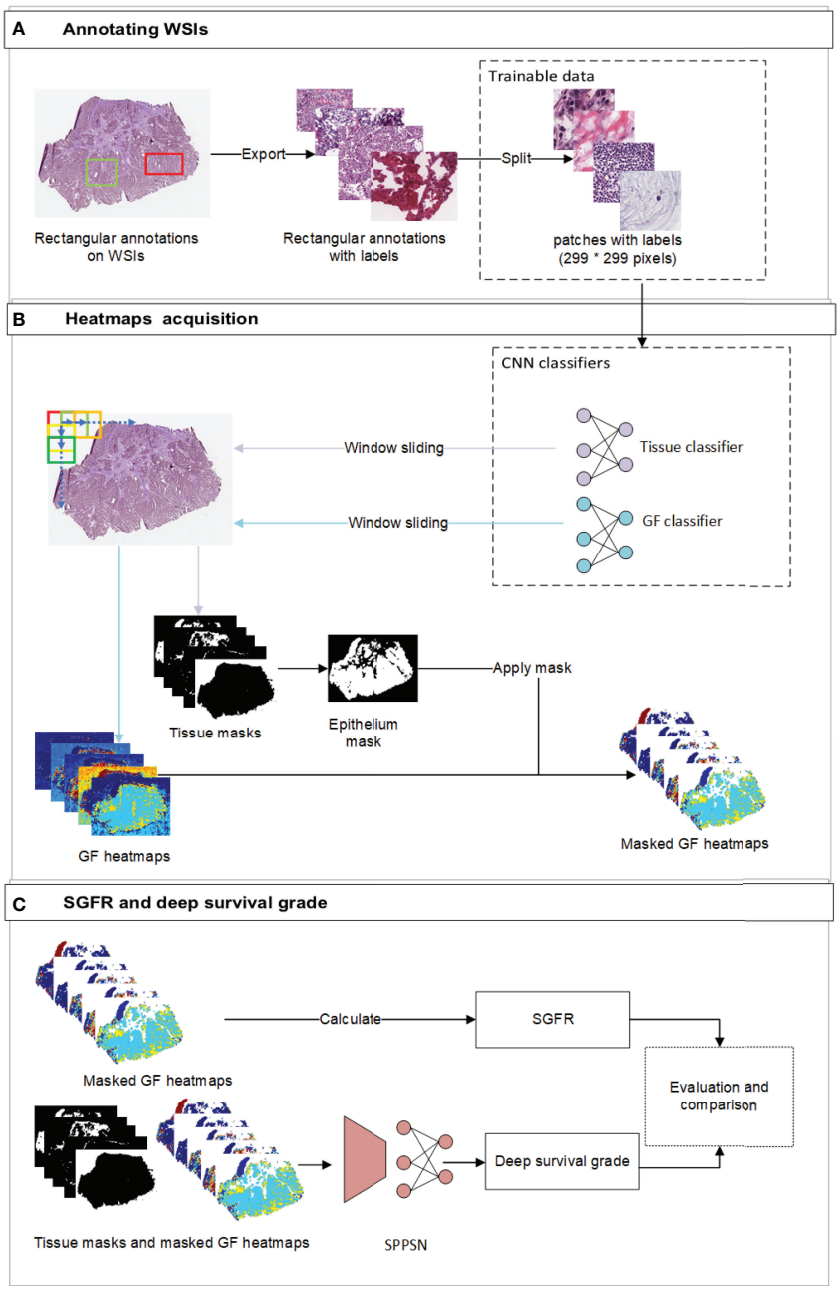


FIGURE 1 | The overview of the workflow. **(A)** The workflow of annotating labels and generating trainable patches. Our pathologists annotated rectangular annotations and related labels from original WSI. We then randomly sampled trainable patches from annotations and balanced the ratio of labels before training. **(B)** the workflow of developing gland formation and tissue category classifiers and generating spatial heatmaps. We used window sliding technique to generate masks of tissue categories and probability maps of tumor gland formation and other tissue types from classifiers. **(C)** The workflow of calculating SGFR and deep survival grade. The epithelium masks were specifically applied to gland formation heatmaps, which were further used in both calculating SGFR and developing SPPSN. WSI, whole-slide image; SGFR, semi-quantitative gland formation ratio; SPPSN, spatial pyramid pooling survival network.

a sliding window to span the entire WSI (**Figure 1B**). The reference labels for classification were manually annotated by institutional experts, and internal measures had been taken to reduce inter-observer error. For the tissue classifier, each region of interest (ROI) was classified as epithelium, stroma, immune cells, other tissues, and background. To train the GF classifier, each ROI was then classified by visual inspection as "GF3" (paucity of gland-forming cells), "GF2" (complex or irregular tubules with cribriform morphology glands), "GF1" (simple tubules only), or "X" (normal epithelial cells). The details of

developing the above two classifiers, including the reference label annotation method, are reported in **Supplementary Materials S1.1–S1.4**.

2.3 Semi-Quantitative Assessment for Gland Formation Ratio

The percentage of the tumor that is gland forming should be calculated to be able to use the WHO stratification system, which can be indirectly assessed by using the abovementioned differentiation maps and tissue category masks. We first focus on the gland regions, which were indicated by category masks, and then estimated the percentage of a gland-forming tumor using the spatial differentiation heatmaps with masks of the gland (**Figure 1C**).

In clinical practice, the estimation of the degree of GF is largely interpretive and lacks standardization for pathologists (8); therefore, the actual GF percentage could not be calculated exactly according to the current WHO definition. So we had to estimate the GF percentage by assigning a semi-quantitative GF ratio (SGFR) in the following manner. We assumed that GF2 regions (complex or irregular tubules with cribriform morphology glands and nuclear polarity partially lost, which sit between gland-forming and non-gland-forming regions) have an arbitrarily assigned weight of 0.5. To spread the weights uniformly, we assigned a weight of 1 to GF1 regions and correspondingly weight 0 to GF3 regions and then excluded all normal tissues. We further assumed the SGFR would be a weighted average of all GF1, GF2, and GF3 in tumor-containing regions as described above will be a reasonably close semi-quantitative approximation of the percentage of gland-forming tumors according to the WHO. The detailed formula of SGFR is provided in **Supplementary Material S2**.

2.4 Spatial Pyramid Pooling Survival Network

Spatial pyramid pooling survival network (SPPSN) was trained to predict the hazard of death using information derived from WSI. A "DeepSurv" neural network has been proposed (17) for predicting hazards on time and event data. The input to the network is designed to be a patient's baseline data. The hidden layers of the network consist of a fully connected layer of nodes, followed by a dropout layer. Therefore, the DeepSurv networks require one-dimensional input data with uniform size. This cannot be applied for our GF spatial heatmaps and category masks, since their shapes obviously change from WSI to WSI. The SPP layer was initially proposed in the field of object detection (18) and has the ability to use arbitrary shape input data. A patient's data can be transformed into one-dimensional uniform data and thus can be applied to the DeepSurv networks. We thus combined the SPP layer and the DeepSurv networks as an SPPSN and applied it to developing our survival prognostication (**Figure 1C**).

For the input data of the SPPSN, we first applied the epithelium tissue type masks for the GF heatmaps since linear predictor maps due to the GF classifier can only be used on gland-forming regions (**Figure 1C**). The input of the SPPSN consists of i) masked GF maps; ii) masks of stroma, immune cells, other tissues, and background; and iii) thumbnails of the original WSI. These were concatenated as the input layer of the SPPSN.

To develop the OS model, input data derived from both COAD and READ sets combined were randomly split into training and validation in an 80:20 ratio. The SPPSN consists of 1 SPP layer, 3 linear layers, and 1 activation layer. We performed the log-rank test to examine the OS distribution between the validation set and training sets. The p-value was 0.70, which indicated that the OS distribution between the validation set and training set had no statistical difference. The detailed architecture and training parameters of the SPPSN are provided in **Supplementary Material S1.5**.

2.5 Statistical Analysis

For the GF classifier and tissue type classifier, we used an area under the curve of the receiver operating curve (AUC) and the confusion matrix to evaluate the model discrimination performance.

For the evaluation of SGFR, we used a Harrell concordance (c-index) and the Kaplan–Meier curve log-rank test on TCGA dataset. To construct the Kaplan–Meier curves, we used three subgroup definitions based on the WHO's definitions (cutoff points of 0.05, 0.5, and 0.95) and an optimized cutoff (19). Optimized cutoff and WHO were compared for validating SGFR. The optimized cutoff point was obtained by the R package of "maxstat."

We used a c-index for evaluating the performance of SGFR and deep survival grade from the SPPSN. A Kaplan–Meier curve log-rank test was adopted for performance evaluation as well. A Spearman's correlation test was applied to determine any correlation between SGFR and deep survival grade. Next, we analyzed common potential OS influencing factors (sex, age, vascular invasion, and American Joint Committee on Cancer (AJCC) stage) in TCGA set and local set in a univariable Cox regression. A multivariable Cox regression of significant factors in the validation set from the last step was developed as the baseline model for comparison. Finally, the SGFR and deep survival grade were separately added into a reference Cox regression model for investigating these grades' significance and prediction-enhancing power. The performance of the two grades was evaluated and compared using the above multivariable Cox regression models. The calibration curves of Cox regression models were also performed to show the detail.

3 RESULTS

The clinical case-mix comparison of included patients with ROI annotations and available follow-up from the three datasets used in this study are given in **Table 1**. Note that the p-value is for the difference between the combined TCGA dataset and the local institutional dataset since the former was used as the discovery set and only the latter was the independent test set.

TABLE 1 | Patient characteristics for the datasets of TCGA-COAD, TCGA-READ, and the local institution.

	TCGA-COAD	TCGA-READ	Local dataset	p-Value
Number of subjects	451	155	106	-
Primary cancer				p < 0.0001
Colon	444 (98%)	5 (3%)	104 (98%)	(Fisher's exact test)
Rectosigmoid	7 (2%)	65 (42%)	0 (0%)	
Rectal	0 (0%)	83 (54%)	2 (2%)	
Other	0 (0%)	2 (1%)	0 (0%)	
Biological sex				0.113
Male	240 (53%)	87 (56%)	66 (62%)	(Pearson's test)
Female	211 (47%)	68 (3%)	40 (38%)	
Years of diagnosis	1998–2013	Not reported	2008–2009	–
Mean age (range)	70 (31–90)	67 (31–90)	57 (27–78)	p < 0.0001 (U test)
Overall staging				0.136
I	71 (16%)	33 (21%)	10 (9%)	(Pearson's test)
II	170 (38%)	49 (32%)	39 (37%)	
III	123 (27%)	50 (32%)	37 (35%)	
IV	64 (14%)	23 (15%)	20 (19%)	
Not known	23 (5%)	0 (0%)	0 (0%)	
Median survival time (months)	101	52	Not reached	-
Median follow-up time (months)	30	26	87	

TCGA-COAD, The Cancer Genome Atlas Colon Adenocarcinoma; TCGA-READ, The Cancer Genome Atlas Rectum Adenocarcinoma.

3.1 Performance of the Tissue Classifier and the Gland Formation Classifier

The AUC indices of each tissue type of segmentation were calculated, corresponding to discrimination of epithelium, stroma, immune cells, other, and background. The tissue classifier showed strong discriminating ability (AUC = 0.981, 0.977, 0.972, 0.979, and 0.999 for epithelium, stroma, immune cells, other, and background, respectively; micro-average AUC = 0.987, macro-average AUC = 0.982) (**Figure 2A**). Likewise, the AUC indices for gland-forming tumors and normal glandular tissue classification were calculated, corresponding to discrimination between GF1, GF2, GF3, and X. The GF classifier also showed strong discrimination performance (AUC = 0.983, 0.963, 0.963, and 0.981 for G1, G2, G3, and X, respectively; micro-average AUC = 0.973, macro-average AUC = 0.973) (**Figure 2B**). To illustrate the results of tissue classification and gland-forming grades, we show a representative example of a WSI and its maps in **Figure 3**. The original WSI is shown (**Figure 3A**) together with an overlay of the tissue type mask (**Figure 3B**), only the gland-forming regions overlaid on the WSI (**Figure 3C**), and finally, the GF differentiation spatial heatmap within the gland-forming regions overlaid on the WSI (**Figure 3D**). The color on the differentiation heatmap corresponds to the linear predictor of the GF model as shown on the color bar adjacent to **Figure 3D**.

3.2 Evaluation and Validation of Semi-Quantitative Gland Formation Ratio on Dataset of The Cancer Genome Atlas

As stated in method section 2.3, we calculated SGFR using the strata of WHO compared with optimized cutoffs. The c-index for SGFR was 0.552, which indicated that the GF percentage possessed poor prediction power. The Kaplan–Meier curves are shown in **Figure 4**. For the WHO's cutoff method, only a few WSIs belong to the high-differentiation (>0.95) or un-differentiation (<0.05) class. In detail, the number of un-

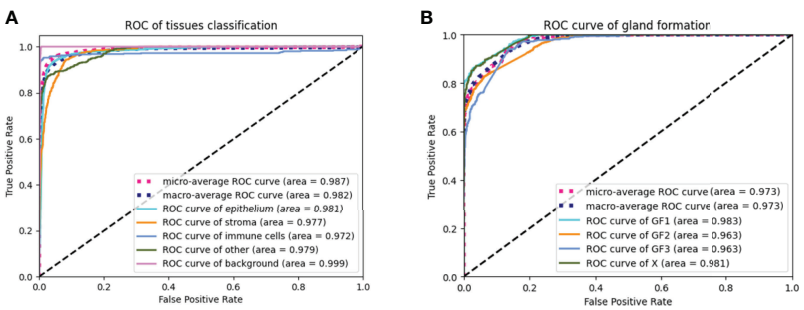


FIGURE 2 | ROC curves of the tissue classifier and the gland formation classifier **(A)** showed ROC curves of the tissue classifier. The AUC were 0.987, 0.982, 0.981, 0.977, 0.972, 0.979 and 0.999 for micro-average ROC, macro-average ROC, epithelium, stroma, immune cells, other and background, respectively. **(B)** showed ROC curves of the gland formation classifier. The AUC were o.973, 0.973, 0.983, 0.963, 0.963 and 0.981 for micro-average ROC, macro-average ROC, GF1, GF2, GF3 and X, respectively.

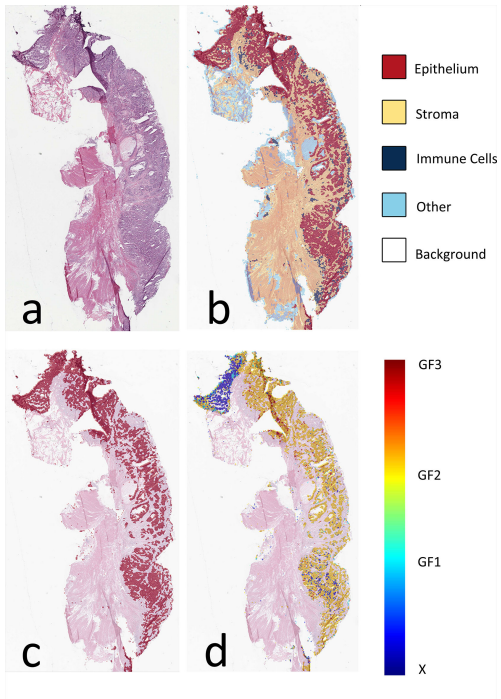


FIGURE 3 | Visual presentation of tissue category and gland formation. **(A)** Thumbnail of original WSI. **(B)** The fusion map of the thumbnail and tissue category map. Different colors represent corresponding tissue categories, and the details can be seen in the right side legend. **(C)** The fusion map of the thumbnail and the epithelium mask selected from panel **(B, D)** The fusion map of the thumbnail and gland formation from the linear predictor on panel **(C)** epithelium mask. WSI, whole-slide image.

differentiation WSIs and high-differentiation WSIs were 5 and 18, respectively, compared to a total of 1,157 valid WSIs. Therefore, we only focus on medium- and low-differentiation groups whose log-rank test p-values were 0.02. The median cutoff method showed no statistically significant differences (p-value 0.26). The optimized cutoff method indicated only one best point for SGFR (p = 0.01). Optimized cutoff also indicated the optimal threshold of 0.49, which ends up being very close to the WHO's proposed cutoff point of 0.5, and the WHO's proposed cutoff point could stratify WSIs well by itself (p = 0.01 *vs.* p = 0.02).

3.3 Comparison of Percentage of Gland Formation and Deep Survival Grade

3.3.1 Univariable Analysis

We applied a Cox model and derived Kaplan–Meier curves using the SGFR and deep survival grade on both validation set and test set (**Table 2**). We need to note that the SGFR value is inversely correlated to hazard (i.e., higher GF percentage means higher differentiation status and consequently leads to lower hazard), and thus the hazard ratios of SGFR came out as negative. In the validation set, SGFR showed no significant discriminatory performance, while deep survival grade performed better with a higher c-index and a p-value below 0.05. In the test set, both SGFR and deep survival grade showed significant discriminatory performance, and deep survival grade had a higher c-index of 0.64.

For Kaplan–Meier curves, we chose a median cutoff method because the WHO's cutoff point did not work in deep survival grade and an optimized cutoff point was not suitable for comparison. In the validation set, the Kaplan–Meier curve log-rank p-value of SGFR was 0.69, while that of deep survival grade was 0.02 (**Figure 5**). In the test set, the p-value of SGFR was slightly higher than 0.05 (0.07) and that of deep survival grade was 0.02 (**Figure 5**). In both the validation set and test set, the deep survival grade showed better prediction ability than the GF percentage according to univariable analysis.

3.3.2 Baseline Multivariable Cox Regression Model

We applied a Cox regression model on single factors of age, sex, vascular invasion, and AJCC pathology stage in the validation set (**Table 2**). Then we selected significant factors of age, vascular invasion, and AJCC pathology stage III and IV to build

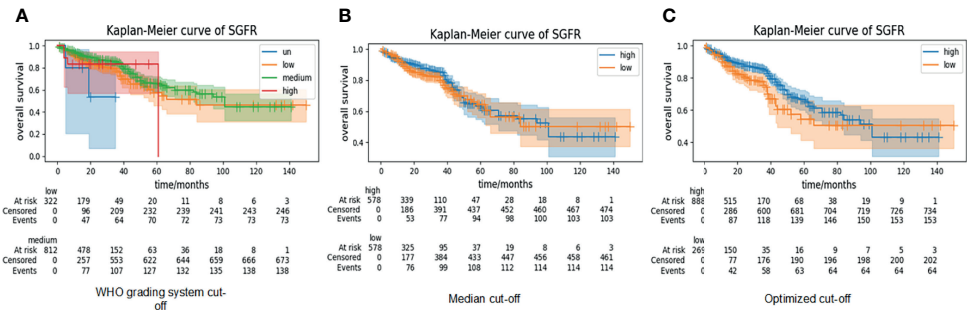


FIGURE 4 | Kaplan–Meier curves of semi-quantitative gland formation ratio (SGFR) with cutoff method of WHO, median, and optimized. According to WHO's grading system, the WSIs can be grouped into high differentiation (high, SGFR > 0.95), medium differentiation (medium, 0.5 < SGFR < 0.95), low differentiation (low, 0.05 < SGFR < 0.50), and un-differentiation (un, SGFR < 0.05), and the related Kaplan–Meier curves are shown in panel **(A, B)** KAPLAN–Meier curves of WSIs with SGFR higher than median (high) and SGFR lower than median (low). **(C)** Kaplan–Meier curves of WSIs with SGFR higher than optimized cutoff point (high) and SGFR lower than optimized cutoff point (low). WSIs, whole-slide images.

TABLE 2 | Univariable Cox models for each metric.

Metrics	Validation set			Test set		
	c-Index	log HR	p-Value	c-Index	log HR	p-Value
Age	0.70	0.065	<0.001	0.56	−0.01	0.38
Sex	0.51	−0.07	0.94	–	–	–
Vascular invasion	0.64	1.12	<0.001	0.57	0.58	0.11
AJCC stage III and IV	0.65	0.91	0.004	0.69	2.43	<0.001
SGFR	0.52	−0.44	0.66	0.62	−7.15	0.07
Deep survival grade	0.64	1.73	0.008	0.64	3.53	0.02

AJCC, American Joint Committee on Cancer; SGFR, semi-quantitative gland formation ratio.

multivariate Cox regression models and apply them on the validation set and test set as baseline models for further reference (**Figure 5**).

3.3.3 Performance Comparison in Multivariable Cox Regression Model

We added the GF percentage or deep survival grade separately into the baseline Cox model (Section 2.5) as extra information. The forest plot of Cox regression models is shown in **Figure 6**. When the GF percentage was added to the baseline model (GF enhanced model), it showed no significant performance as an independent hazard factor, and the c-index of the model could not be increased in the validation set, while SGFR enhanced the c-index from 0.74 to 0.76 in the test set. When deep survival grade was added (deep survival grade enhanced model), it showed significant performance as an independent hazard factor, and the c-index was increased from 0.75 to 0.77 in the validation set and from 0.74 to 0.77 in the test set. Hence, deep survival grade possessed an advantage over the GF percentage and enhanced the prediction power of a Cox model based solely on traditional clinical parameters. We also investigated calibration curves of Cox models, and the results are provided in **Supplementary Material S3**.

4 DISCUSSION

This study has been a Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) (20) Type 3 study combining model development and independent validation as one scope of work. We developed and externally validated a CNN model for assigning the GF percentage and a deep survival grade for CRC, based on automated quantitative analysis of WSI. Overall classification accuracy was over 90% for tissue classification and over 85% for GF classification; however, the dominant mode of failure appeared to be the incorrect assignment of GF2 to either normal tissue or high GF tumor slides.

Interpretability of the tissue and GF models was enhanced by providing pathologists with a color-coded composite map as output, as an overlay onto the original WSI. To examine SGFR, we compared the optimized cutoff point of SGFR and the WHO's

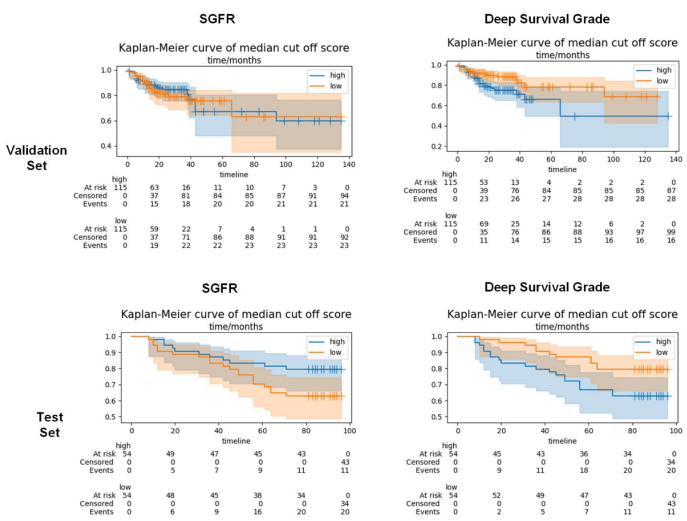


FIGURE 5 | Kaplan–Meier curves of semi-quantitative gland formation ratio (SGFR) and deep survival grade in validation set and test set. SGFR and deep survival grade were grouped into high group (higher than median) and low group (lower than median) and applied Kaplan–Meier curves in both validation set and test set.

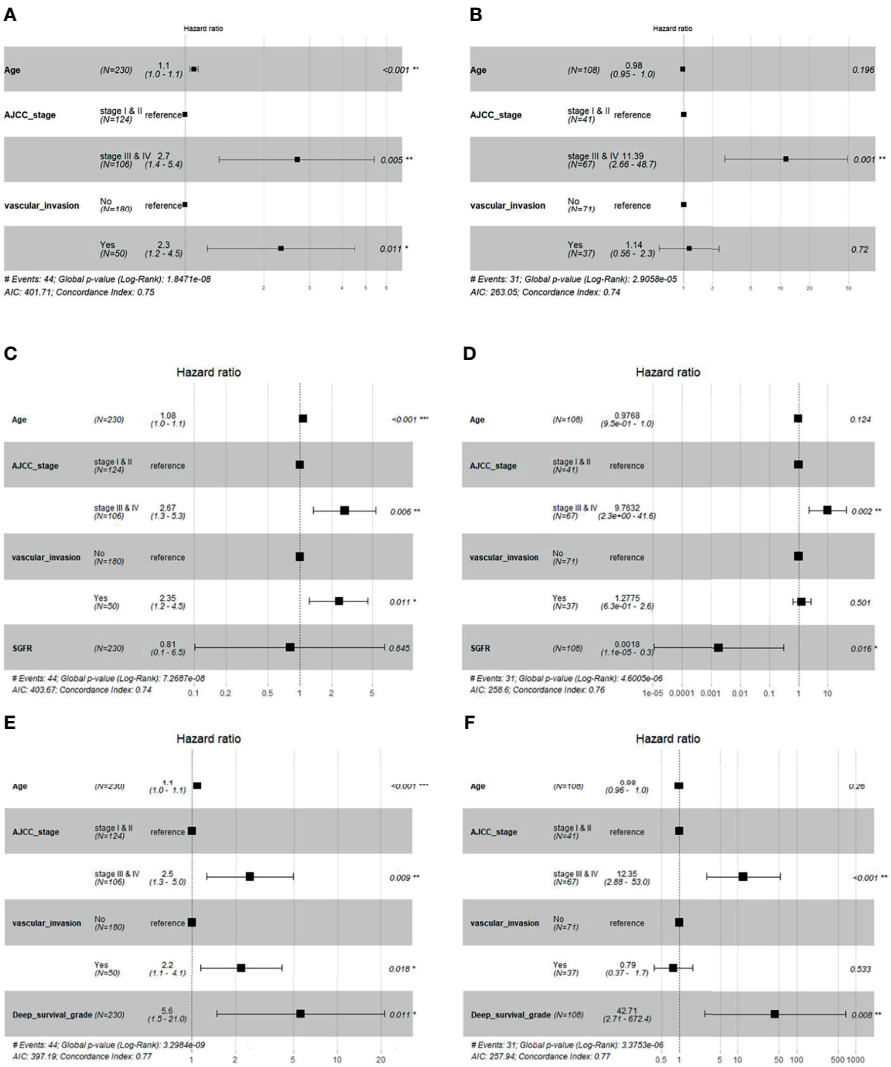


FIGURE 6 | Cox regression forest plot of baseline model, baseline + Semi-quantitative gland formation ratio (SGFR) model and baseline + deep survival grade model. **(A, B)** are the baseline multivariable Cox models of age, AJCC and vascular invasion for validation set and test set, respectively. **(C, D)** are multivariable Cox models of age, AJCC, vascular invasion and SGFR for validation set and test set, respectively. **(E, F)** are multivariable Cox models of age, AJCC, vascular invasion and deep survival grade for validation set and test set, respectively.

proposed cutoff point. Our results indicate that SGFR has moderate prediction ability on stratifying CRC patients whose optimized cutoff point was close to the WHO's proposed cutoff point of 50%. However, the WHO's proposed cutoff points of 5% and 95% were unable to be compared in TCGA dataset due to the low sample size. Our optimized cutoff algorithm also indicated that there exists only one cutoff point in TCGA dataset, which supports a two-tiered grading system (21). In external validation datasets, we found that the GF percentage still possessed moderate prediction ability but cannot increase the c-index of the multivariable baseline Cox regression model, which may allude to the limitation of the GF percentage in clinical usage. To discover a new indicator on GF beyond the GF percentage, we developed the SPPSN model, which provided a deep survival grade. Compared to the GF percentage, deep survival grade possessed a higher c-index on itself and could increase the c-index of a baseline clinical Cox model. Therefore, the deep survival grade shows a more promising potential on outcome prediction and even on clinical therapy than SGFR.

The quality of WSIs may illustrate the difference of GF percentage discrimination between TCGA dataset and the local dataset. The overall quality of TCGA's WSIs was lower than that of a local dataset. The low-quality WSIs such as stained WSIs, WSIs with little regions of epithelium tissue, and highly folded WSIs were common in TCGA set, while WSIs in the local dataset all contained adequate and clear epithelium tissues for tumor grading. Consequently, the GF percentage of the local dataset's WSIs was more stable and reliable due to better quality.

Our approach has a few advantages over a traditional pathology visual grading workflow. First, a trained machine algorithm is able to exhaustively slide a window of attention over an entire WSI and generate a composite picture of the most probable grade in each window. This helps the pathologist by providing an intuitive visual overview for assigning a tumor grade, along with the heterogeneity within the sample. In effect, the GF model has auto-segmented tumor subregions in the WSI according to their most likely differentiation grade. All of the original WSI information has been retained for closer scrutiny and more precise human review, if so desired.

Calculating the GF percentage was not done within the scope of this present study. This additional feature would be easily performed once the composite maps have been generated. Additionally, this abovementioned workflow will be efficient, objective, and reproducible, which might not always be assured among human observers. There is less chance that a small but significant WSI region might be accidentally overlooked by a human rater. Furthermore, the speed of grading WSIs can be revolutionized by the automatic calculation of SGFR. We also compared the performance of SGFR with human-grading differentiation grade in the test set. The human-grading differentiation grade also shows moderate discriminable ability for stratifying survival outcomes, with a c-index that was close to our SGFR. However, similar to SGFR, the human-grading differentiation grade increases less the c-index of baseline Cox regression model than that of deep survival grade and is also less significant than deep survival grade as an independent factor. Although both SGFR and human-grading differentiation grades were imprecise estimations of the WHO's differentiation grading system, they still show some value in stratifying survival outcomes. The performance of human-grading differentiation grade can be seen in **Supplementary Material S4**.

Beyond the traditional GF percentage, we also provide a stronger indicator of deep survival grade. Exploring a new indicator from CRC WSIs was previously conducted by several research groups (12, 13, 15, 16). There are two major differences between our work and previous works. First, our SPPSN is the first model that analyzed GF, which can bridge the deep learning output and clinical explanation. Deriving an indicator from a WSI can face the problem of balancing global and local information. The global information is derived from the whole WSI, while the local information uses local regions to represent the whole WSI. Most previous works adopted global information for outcome prediction such as ensemble score of tiles from WSIs (16), weighted classification results (13), or tumor–stroma ratio (12). However, pathologists could assess overall tumor grade either based on a qualitative impression of the sum of global histologic features, or the highest grade of local features observed anywhere in the neoplasm, or the relative proportion of the undifferentiated tumor, or the degree of differentiation along the advancing edge (8). Therefore, both global information and local information were adopted in pathologists' grading, and simply using one of the two is incomplete. Bychkov et al. included global and local information (15) by long short-term memory method, but the process was still different from pathologists' grading. Our SPPSN, however, included global information and local information with several levels by SPP layers, and this process was similar to the pathologists' grading procedure.

A number of limitations remain to be addressed in future work. We examined the errors of misclassification and found that uncertain findings tended to be assigned as GF2. This could be an undesirable side effect of the dataset being highly unbalanced with a much greater proportion of GF2 annotations. We attempted to ameliorate this with majority under-sampling, but it is likely that some vestigial influence persists in the model. Upon closer inspection, these disagreements were also present in pathologists' manual grade assignments. In real-world clinical practice, the accurate distinction between GF1 and GF2 can also be difficult for human experts. Our present results largely reflect the known difficulty and inherent uncertainty in clinical practice.

Our ground-truth annotations were only provided on so-called "perfect" regions on WSIs, while in real clinical grading, there are common defects such as unintended stains and folded tissues. Our classifiers performed abnormally when there were defects on WSIs, which reduced the classification accuracy of models and consequently influenced GF percentage calculation and deep survival grade prediction.

Of current concern is the misclassification of normal tissue as GF2 tumors. Such false-positive incidents could lead to unnecessary intervention and/or psychological distress to patients. Although the rate of false positives (normal tissue being misclassified as GF1, GF2, or GF3) appears low in this study, it is important to note that such errors can happen. Therefore, cautious clinical commissioning and routine quality assurance of such classification models are needed if they are included in clinical practice (22–24). Although this model has not yet been deployed into routine clinical practice, it has been designed and developed in close collaboration between data scientists and highly experienced clinicians; therefore, it is hoped that a clinical translation gap will be more easily bridged through future work.

In this study, we adopted the WHO's cutoff point as the reference to validate the estimation of SGFR. However, the cutoff points for the WHO four-tiered schema have not been deeply investigated for their relevance in survival prediction. For example, there exists an alternative two-tiered grading system that has been recommended by a multidisciplinary CRC working group (12), who proposed a 30% cutoff point of gland-forming regions (13). Although our SGFR's optimized cutoff point seems to agree with that of WHO, the SGFR may be not consistent with real GF percentage on the aspect of distribution.

The grading of CRC WSIs remains to be a complex work due to the existence in the literature of several different grading schemas without widespread acceptance and uniform use of any single system by practicing pathologists (8). However, our research was conducted only on the WHO's traditional definition of the GF percentage regardless of cytologic features. Tumor budding and TILs, which have been developed as interesting additive prognostic factors in CRC, were not included in our model. Thus, more factors considered in CRC tumor grading can be included in our future works.

There are other clinical factors with potential influence on survival that have not been included in the above model. For instance, different types of surgical and/or drug interventions have not been taken into account. Second, since the WSIs contain the resected tumor, there is little information in the image about the mucosa. Thus, tumor invasion depth and peritumoral morphology will not be available to the CNN to train on as potentially important survival features. No tumor invasion data have been included in the above model.

In conclusion, our study proposed a semi-quantitative estimation of the GF percentage with moderate discriminatory ability on stratifying WSIs' survival outcomes and an optimized cutoff point similar to the WHO's cutoff point. This SGFR can be regarded as a good estimation of the GF percentage and has a potential application on automatic CRC tumor grading. We also proposed a novel network architecture (SPPSN) to extract risk information from WSI. Through SPPSN, a new WSI-based predictor (deep survival grade) was created with better performance on stratifying WSIs than that of SGFR. Deep survival grade also showed better performance in a multivariable Cox model compared to SGFR.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the external dataset of whole-slide images and related patients' characteristics are not allowed to be shared by the local institution. The model weights, source code, and the annotations on TCGA datasets by pathologists of the local institution are available to share. Requests to access the datasets should be directed to shenlun.chen@maastro.nl.

AUTHOR CONTRIBUTIONS

SC designed and performed the experiments, analyzed the data, and wrote the original manuscript. MZ helped collect, annotate, and analyze the data. JW helped design and perform the experiments, analyze the data, and revise the manuscript. MX helped annotate the data and provided pathological knowledge. WH helped in the discussion. LW helped design the experiments, revise the manuscript, and analyze the data and in the discussion. AD helped revise the manuscript and in the discussion. WS helped annotate the data, provide pathological information, and design the experiments. ZZ helped design the experiments, analyze the data, and guide the research presented.

ACKNOWLEDGMENTS

The results of this study are partly based upon data generated by TCGA Research Network: https://www.cancer.gov/tcga.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.833978/full#supplementary-material

REFERENCES

1. Ryan E, Khaw YL, Creavin B, Geraghty R, Ryan EJ, Gibbons D, et al. Tumor Budding and PDC Grade Are Stage Independent Predictors of Clinical Outcome in Mismatch Repair Deficient Colorectal Cancer. *Am J Surg Pathol* (2018) 42(1):60–8. doi: 10.1097/PAS.0000000000000931
2. Koelzer VH, Zlobec I, Lugli A. Tumor Budding in Colorectal Cancer–Ready for Diagnostic Practice? *Hum Pathol* (2016) 47(1):4–19. doi: 10.1016/j.humpath.2015.08.007
3. Rozek LS, Schmit SL, Greenson JK, Tomsho LP, Rennert HS, Rennert G, et al. Tumor-Infiltrating Lymphocytes, Crohn's-Like Lymphoid Reaction, and Survival From Colorectal Cancer. *J Natl Cancer Inst* (2016) 108(8):1–15. doi: 10.1093/jnci/djw027
4. Chapuis PH, Dent OF, Fisher R, Newland RC, Pheils MT, Smyth E, et al. A Multivariate Analysis of Clinical and Pathological Variables in Prognosis After Resection of Large Bowel Cancer. *Br J Surg* (1985) 72(9):698–702. doi: 10.1002/bjs.1800720909
5. Mulcahy HE, Skelly MM, Husain A, O'Donoghue HE. Long-Term Outcome Following Curative Surgery for Malignant Large Bowel Obstruction. *Br J Surg* (1996) 83(1):46–50. doi: 10.1002/bjs.1800830114
6. Newland RC, Dent OF, Lyttle MN, Chapuis PH, Bokey EL. Pathologic Determinants of Survival Associated With Colorectal Cancer With Lymph Node Metastases. A Multivariate Analysis of 579 Patients. *Cancer* (1994) 73(8):2076–82. doi: 10.1002/1097-0142(19940415)73:8<2076::AID-CNCR2820730811>3.0.CO;2-6
7. Nagtegaal ID, Odze RD, Klimstra D, Paradis V, Rugge M, Schirmacher P, et al. The 2019 WHO Classification of Tumours of the Digestive System. *Histopathology* (2020) 76(2):182–8. doi: 10.1111/his.13975
8. Compton CC. Optimal Pathologic Staging: Defining Stage II Disease. *Clin Cancer Res* (2007) 13(22 Pt 2):6862s–70s. doi: 10.1158/1078-0432.CCR-07-1398
9. Blenkinsopp WK, Stewart-Brown S, Blesovsky L, Kearney G, Fielding LP. Histopathology Reporting in Large Bowel Cancer. *J Clin Pathol* (1981) 34 (5):509–13. doi: 10.1136/jcp.34.5.509
10. Jass JR, Atkin WS, Cuzick J, Bussey HJ, Morson BC, Northover JM, et al. The Grading of Rectal Cancer: Historical Perspectives and a Multivariate Analysis of 447 Cases. *Histopathology* (1986) 10(5):437–59. doi: 10.1111/j.1365-2559.1986.tb02497.x
11. Thomas GD, Dixon MF, Smeeton NC, Williams NS. Observer Variation in the Histological Grading of Rectal Carcinoma. *J Clin Pathol* (1983) 36(4):385–91. doi: 10.1136/jcp.36.4.385
12. Zhao K, Li Z, Yao S, Wang Y, Wu X, Xu Z, et al. Artificial Intelligence Quantified Tumour-Stroma Ratio Is an Independent Predictor for Overall Survival in Resectable Colorectal Cancer. *EBioMedicine* (2020) 61:103054. doi: 10.1016/j.ebiom.2020.103054
13. Kather JN, Krisam J, Charoentong P, Luedde T, Herpel E, Weis CA, et al. Predicting Survival From Colorectal Cancer Histology Slides Using Deep Learning: A Retrospective Multicenter Study. *PloS Med* (2019) 16(1): e1002730. doi: 10.1371/journal.pmed.1002730
14. Kather JN, Poleszczuk J, Suarez-Carmona M, Krisam J, Charoentong P, Valous NA, et al. *In Silico* Modeling of Immunotherapy and Stroma-Targeting Therapies in Human Colorectal Cancer. *Cancer Res* (2017) 77 (22):6442–52. doi: 10.1158/0008-5472.CAN-17-2006
15. Bychkov D, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, et al. Deep Learning Based Tissue Analysis Predicts Outcome in Colorectal Cancer. *Sci Rep* (2018) 8(1):3395. doi: 10.1038/s41598-018-21758-3
16. Skrede OJ, De Raedt S, Kleppe A, Hveem TS, Liestol K, Maddison J, et al. Deep Learning for Prediction of Colorectal Cancer Outcome: A Discovery and Validation Study. *Lancet* (2020) 395(10221):350–60. doi: 10.1016/S0140-6736(19)32998-8

17. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang TT, Kluger Y. DeepSurv: Personalized Treatment Recommender System Using a Cox Proportional Hazards Deep Neural Network. *BMC Med Res Method* (2018) 18: 1–12. doi: 10.1186/s12874-018-0482-1
18. He KM, Zhang XY, Ren SQ, Sun J. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Trans Pattern Anal Mach Intell* (2015) 37(9):1904–16. doi: 10.1109/TPAMI.2015.2389824
19. Hothorn T, Lausen B. On the Exact Distribution of Maximally Selected Rank Statistics. *Comput Stat Data Anal* (2003) 43(2):121–37. doi: 10.1016/S0167-9473(02)00225-6
20. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement. *Ann Intern Med* (2015) 162(1):55–63. doi: 10.7326/M14-0697
21. Compton CC, Fielding LP, Burgart LJ, Conley B, Cooper HS, Hamilton SR, et al. Prognostic Factors in Colorectal Cancer. College of American Pathologists Consensus Statement 1999. *Arch Pathol Lab Med* (2000) 124 (7):979–94. doi: 10.5858/2000-124-0979-PFICC
22. Diez FJ, Yebra M, Bermejo I, Palacios-Alonso MA, Calleja MA, Luque M, et al. Markov Influence Diagrams. *Med Decis Making* (2017) 37(2):183–95. doi: 10.1177/0272989X16685088
23. Yang L, Wang Q, Cui T, Huang J, Shi N, Jin H. Reporting of Coronavirus Disease 2019 Prognostic Models: The Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis Statement. *Ann Transl Med* (2021) 9(5):421. doi: 10.21037/atm-20-6933
24. Shi Z, Foley KG, Pablo de Mey J, Spezi E, Whybra P, Crosby T, et al. External Validation of Radiation-Induced Dyspnea Models on Esophageal Cancer Radiotherapy Patients. *Front Oncol* (2019) 9:1411. doi: 10.3389/fonc.2019.01411

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Chen, Zhang, Wang, Xu, Hu, Wee, Dekker, Sheng and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 20 May 2022
doi: 10.3389/fonc.2022.841801

Prostate Surface Distension and Tumor Texture Descriptors From Pre-Treatment MRI Are Associated With Biochemical Recurrence Following Radical Prostatectomy: Preliminary Findings

Rakesh Shiradkar[1], Soumya Ghose[2], Amr Mahran[3], Lin Li[1], Isaac Hubbard[1], Pingfu Fu[4], Sree Harsha Tirumani[5], Lee Ponsky[3], Andrei Purysko[6] and Anant Madabhushi[1]*

[1] Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH, United States, [2] GE Global Research, Niskayuna, NY, United States, [3] Department of Urology, University Hospitals Cleveland Medical Center, Cleveland, OH, United States, [4] Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, United States, [5] Department of Radiology, University Hospitals Cleveland Medical Center, Cleveland, OH, United States, [6] Department of Abdominal Imaging and Nuclear Radiology, Imaging Institute, Cleveland Clinic, Cleveland, OH, United States

OPEN ACCESS

Edited by:
Guang Yang, Imperial College London, United Kingdom

Reviewed by:
Ugo Giovanni Falagario, University of Foggia, Italy
Stefano Puliatti, University of Modena and Reggio Emilia, Italy

***Correspondence:**
Rakesh Shiradkar
rxs558@case.edu

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *22 December 2021*
Accepted: *13 April 2022*
Published: *20 May 2022*

Citation:
Shiradkar R, Ghose S, Mahran A, Li L, Hubbard I, Fu P, Tirumani SH, Ponsky L, Purysko A and Madabhushi A (2022) Prostate Surface Distension and Tumor Texture Descriptors From Pre-Treatment MRI Are Associated With Biochemical Recurrence Following Radical Prostatectomy: Preliminary Findings. Front. Oncol. 12:841801. doi: 10.3389/fonc.2022.841801

Objective: To derive and evaluate the association of prostate shape distension descriptors from T2-weighted MRI (T2WI) with prostate cancer (PCa) biochemical recurrence (BCR) post-radical prostatectomy (RP) independently and in conjunction with texture radiomics of PCa.

Methods: This retrospective study comprised 133 PCa patients from two institutions who underwent 3T-MRI prior to RP and were followed up with PSA measurements for ≥3 years. A 3D shape atlas-based approach was adopted to derive prostate shape distension descriptors from T2WI, and these descriptors were used to train a random forest classifier (C_S) to predict BCR. Texture radiomics was derived within PCa regions of interest from T2WI and ADC maps, and another machine learning classifier (C_R) was trained for BCR. An integrated classifier C_{S+R} was then trained using predictions from C_S and C_R. These models were trained on D_1 (N = 71, 27 BCR+) and evaluated on independent hold-out set D_2 (N = 62, 12 BCR+). C_{S+R} was compared against pre-RP, post-RP clinical variables, and extant nomograms for BCR-free survival (bFS) at 3 years.

Results: C_{S+R} resulted in a higher AUC (0.75) compared to C_R (0.70, p = 0.04) and C_S (0.69, p = 0.01) on D_2 in predicting BCR. On univariable analysis, C_{S+R} achieved a higher hazard ratio (2.89, 95% CI 0.35–12.81, $p < 0.01$) compared to other pre-RP clinical variables for bFS. C_{S+R}, pathologic Gleason grade, extraprostatic extension, and positive surgical margins were associated with bFS ($p < 0.05$). C_{S+R} resulted in a higher C-index (0.76 ± 0.06) compared to CAPRA (0.69 ± 0.09, $p < 0.01$) and Decipher risk (0.59 ± 0.06, $p < 0.01$); however, it was comparable to post-RP CAPRA-S (0.75 ± 0.02, p = 0.07).

Conclusions: Radiomic shape descriptors quantifying prostate surface distension complement texture radiomics of prostate cancer on MRI and result in an improved association with biochemical recurrence post-radical prostatectomy.

Keywords: magnetic resonance imaging, prostate cancer, retrospective studies, prostatectomy, artificial intelligence, machine learning

1 INTRODUCTION

An estimated 30%–35% of prostate cancer (PCa) patients experience biochemical recurrence (BCR) within 10 years post-radical prostatectomy (RP) (1). The occurrence of BCR is often found to be associated with metastasis (2) and PCa-specific mortality (3). Several predictors of BCR have been presented including pre-treatment CAPRA (4), post-surgical CAPRA-S (5), and Decipher risk (6). However, these models use invasive or post-treatment factors, are site dependent, and do not exclusively capture tumor heterogeneity and morphology. Non-invasive pre-treatment image-based prediction of BCR-free survival (bFS) may allow for treatment intensification or closer surveillance (7).

Multiparametric magnetic resonance imaging (mpMRI) is now increasingly used for PCa detection, staging, and prediction of the risk of BCR (8–17). Radiomic texture features provide an alternative representation for characterizing tumor heterogeneity and have been shown to improve PCa risk characterization (18–21) and also prognosticate BCR (22–24). However, they are susceptible to scanner variations, acquisition protocols, image artifacts, and non-standardized image intensities (25, 26).

There is evidence to suggest that cancerous lesions tend to induce mechanical stress in the surrounding tissue (27, 28). In the prostate, stresses induced by the tumors impact neighboring benign tissue (29) and cause deformation which may in turn impact the shape of the prostate capsule. Previous studies (30, 31) explored the idea of quantifying shape distension of the prostate between more and less aggressive diseases. Rusu et al. (30) have shown that prostate shape on T2-weighted MRI (T2WI) was significantly different between patients with and without cancerous lesions. Ghose et al. (31, 32) have shown that statistically significant differences in the shape of the prostate were observed between BCR+ and BCR− patients on T2WI. This now leads us to the hypothesis that radiomic descriptors that quantify differential distension of prostate shape may be associated with BCR outcome post-RP.

In this work, we present a new approach to quantify prostate distension in terms of radiomic shape descriptors from pre-treatment T2WI using a 3D shape atlas-based method. We seek to evaluate the association of these shape descriptors with BCR post-RP at 3 years. Since radiomic texture features of PCa from pre-treatment T2WI and apparent diffusion coefficient (ADC) maps have already been shown to be associated with BCR (17, 22–24), we evaluate the combination of prostate shape and tumor texture radiomics for their association with bFS. Additionally, we also sought to evaluate whether radiomic shape descriptors, which are less dependent on MRI intensities, are possibly more robust and resilient to scanner- and site-specific variations that could more substantially impact texture-based descriptors (25, 26). This was done by evaluating our approach using data from multiple cohorts acquired from two different sites. We also compared our integrated shape and texture radiomics approach against routine clinical variables and extant nomograms for predicting bFS.

2 MATERIALS AND METHODS

2.1 Patient Selection and Data Characteristics

This retrospective study included patients from two different institutions (I_1 and I_2), compliant with HIPAA regulations and approved by the local IRB, with a waiver of informed consent. A total of 263 patients from three cohorts (C_1–C_3) were identified with biopsy-confirmed PCa who underwent 3T mpMRI prior to RP. Patient records between 2009 and 2017 were reviewed and included in the study if they 1) underwent RP without additional therapies, 2) followed up for at least 3 years and had documented BCR (BCR+) defined as two consecutive readings of prostate-specific antigen (PSA) ≥0.2 ng/ml post-RP [18] or no BCR (BCR−), and 3) had axial T2WI and ADC maps without acquisition artifacts. $N = 133$ studies identified through this process were partitioned into a training set D_1 and independent hold-out validation set D_2 as illustrated in **Figure 1**. After RP, all patients underwent periodic follow-up according to established clinical protocol (3–6 months in the first year and 6–12 months in the following years). BCR− patients were censored at their last follow-up date for survival analysis.

Biparametric MRI protocol was used in this study that included T2WI (in-plane: 0.3 mm, slice thickness: 3 mm, TR/TE: 3,802–5,151/105–115 ms) and DWI (1.4 mm, 3 mm, 3,751–4,880/50–74 ms, *b*-values: 0–2,000 s/mm^2) along with ADC maps acquired *via* an endo-rectal coil (Achieva, Philips, Best, Netherlands) in C_1, T2WI (0.5 mm, 3 mm, 3,730/121 ms) and DWI (1.6 mm, 3 mm, 4,700/86 ms, 0–1,500 s/mm^2) acquired *via* a pelvic-phased array coil (PPAC) (Skyra, Siemens, Erlangen, Germany) in C_2, and T2WI (0.6 mm, 3 mm, 7,200/96 ms) and DWI (1.2 mm, 3 mm, 7,900/88 ms, 0–1,400 s/mm^2) *via* PPAC (Skyra, Siemens, Erlangen, Germany). Additional sequences, such as T1-weighted, dynamic contrast-enhanced (DCE), were acquired but not analyzed in the current study. Detailed dataset characteristics are provided in **Table 1** and imaging parameters in **Table 2**.

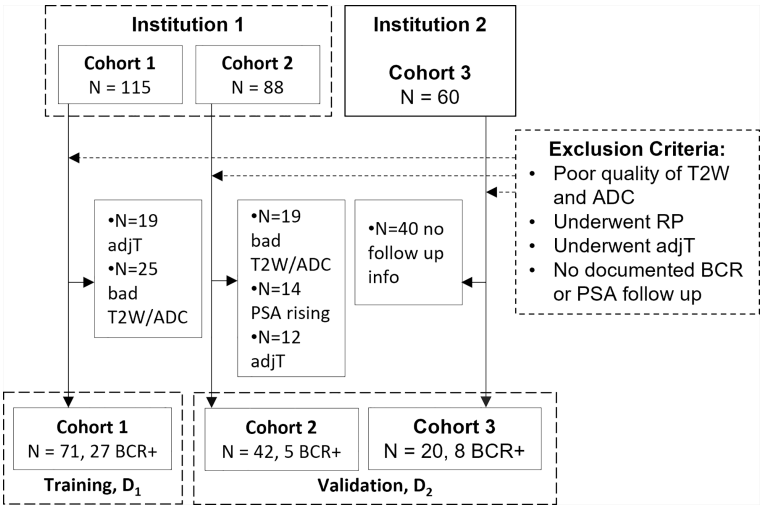


FIGURE 1 | Flowchart of patient selection from two different institutions. adjT, adjuvant or neoadjuvant therapy; RP, radical prostatectomy; BCR, biochemical recurrence; PSA, prostate-specific antigen.

2.2 Image Pre-Processing and Tumor Segmentation

A board-certified radiologist (10 years of experience in prostate imaging) reviewed the MRI scans in D_1. They used the 3D Slicer software (33) to delineate prostate capsule and dominant PCa lesions on T2WI using histopathology template reports from RP as reference. Prostate and lesion delineation in D_2 was performed by two board-certified radiologists (8 and 10 years of experience in prostate imaging) in a similar manner and assigned PI-RADS-v2.1 scores (34).

2.3 Methodology

The pipeline for radiomic prostate shape and lesion texture, along with their combination to evaluate the association with BCR, is illustrated in **Figure 2**. The term surface of interest (SOI) is used in the context of shape descriptors to describe a region on the 3D surface, while the region of interest (ROI) is used in the context of PCa lesions on 2D MRI slices.

2.3.1 Computation of a Consensus Surface of Interest

Prostate segmentations on T2WI within D_1 were co-registered via rigid registration followed by a deformable registration to create a 3D shape atlas for the BCR+ and BCR− cohorts. A spatially contextual SOI of the prostate capsule was uniquely identified from statistically significant shape differences between BCR+ and BCR− atlases as described earlier by Ghose et al. (31) (details in the **Supplementary Material**). To minimize the effect of the choice of template in establishing the SOI, the entire process was repeated N times, each time with a different set of BCR+ and BCR− atlases generating an SOI_i (where $I = 1, 2, 3 \ldots N$). Here, N was chosen as 27, equal to the number of patients in BCR+ class of D_1. The individual SOI_i was again co-registered to a common frame of reference using a rigid transformation. Next, a consensus SOI (SOI_C) was computed as a mean of all the SOI_i and the mean volume was binarized at a threshold of 0.5.

2.3.2 Prostate Shape Distension Features

Differential distension of the prostate capsule for each patient is quantified using the magnitudes of Gaussian curvature (κ) of the surface and orientation of the surface normal (θ, ϕ) at vertices of the prostate mesh. The Gaussian curvature measures the intrinsic degree of curvedness of a surface, and positive values indicate a greater differential expansion at the center, while negative values indicate expansion at the edges (35). Surface normal quantifies the local orientation of the surface. These features (κ, θ, and ϕ) are meaningful when extracted within the SOI_C that comprises those vertices where statistically significant shape differences between BCR+ and BCR− prostates were observed. Further, the SOI_C was cropped to include the mid-gland region alone after mapping it to individual prostate meshes to minimize the effect of inter-reader variations in prostate segmentation at the apex and base. The SOI_C overlaid onto a BCR+ and a BCR− prostate is illustrated in **Figure 3**. Surface normal orientation (θ, ϕ) provides direction of the prostate distension and surface mean curvature (κ) quantifies local shape deformation, and these were extracted for every vertex on the mapped SOI_C. For each patient, at each vertex, a vector of κ, θ, and ϕ was derived and a set of four statistical measures, namely, mean, standard deviation, skewness, and kurtosis, were computed resulting in $4 \times 3 = 12$ radiomic shape descriptors per patient.

2.3.3 Radiomic Texture Features of Prostate Cancer

A set of 75 radiomic features were extracted on a per-voxel basis from each of the standardized T2WI and ADC maps within the

TABLE 1 | Summary of the patient characteristics in different cohorts in terms of clinical variables.

Parameter	Training D_1	Validation D_2	
	Cohort 1	Cohort 2	Cohort 3
N (patients)	71	42	20
N (BCR+)	27	5	8
N (BCR–)	44	37	12
Median age (range), years	59 (47–79)	64 (42–76)	61 (47–86)
Mean PSA (range), ng/ml	10 (1–58)	10.6 (1.8–88.3)	9 (1.2–69.4)
Lesion volume (cm^3)	3.93 ± 6.47	3.34 ± 5.13	2.20 ± 5.05
Prostate volume (cm^3)	35.65 ± 12.37	38.81 ± 19.04	40.83 ± 16.99
Follow-up (months)	43 ± 28	30 ± 24	33 ± 18
biopsy Gleason Grade Group and RP Gleason Grade Group (*N*)			
1	15	7	3
2	20	27	6
3	8	13	5
4	10	3	4
5	18	6	2
RP pGG			
1	6	4	3
2	22	24	8
3	13	15	4
4	7	3	2
5	18	10	3
PI-RADS-v2.0			
1	0	1	0
2	10	8	2
3	6	3	2
4	21	10	6
5	34	34	10
EPE			
Yes	33	21	8
No	30	21	12
N/A	8	0	0
SVI			
Yes	24	5	4
No	42	37	16
N/A	5	0	0
PSM			
Yes	0	20	6
No	0	22	14
N/A	71	0	0
LNI			
Yes	12	15	1
No	41	22	18
N/A	18	5	1
Decipher risk			
Low	N/A	20	N/A
Intermediate	N/A	5	N/A
High	N/A	17	N/A

BCR, biochemical recurrence; RP, radical prostatectomy ;PI-RADS, Prostate Imaging Reporting and Data System; EPE, Extra Prostatic Extension; SVI, Seminal Vesicle Invasion; PSM, Positive Surgical Margin; LNI , Lymph Node Invasion; N/A, Not Available.

radiologist-delineated PCa ROIs. These include first- and second-order statistics, Gabor (36), Laws (37), Haralick (38), and CoLlAGe (39) features. These features characterize the underlying tissue heterogeneity and have previously been shown to be prognostic of BCR (17, 22–24). Four statistics, namely, mean, standard deviation, skewness, and kurtosis, were calculated for voxel-wise radiomic feature within each PCa ROI resulting in $75 \times 2 \times 4 = 600$ radiomic texture descriptors per patient. Strongly correlated features and unstable features were eliminated, and mRMR feature selection (40) was employed to identify a subset of features associated with BCR (details in the **Supplementary Material**).

2.4 Statistical Analysis and Comparison

A random forest machine learning classifier (C_S) with 50 trees and 5 as the maximum tree depth was trained using prostate distension features from D_1 to predict the binary outcome of BCR. The random forest classifier was chosen to ensure that parametric assumptions are not made for the distribution of shape distension features. A texture radiomics classifier (C_R) was trained in a similar manner as C_S.

A linear regression model (C_{S+R}) was trained on predictions from C_S and C_R using patients from D_1. The trained models (C_S, C_R, C_{S+R}) were evaluated using patients from the independent validation test set D_2. Kruskal–Wallis test was employed to assess

TABLE 2 | Imaging parameters of scans used in this study.

Parameter	Institution 1		Institution 2
	Scanner 1	Scanner 2	
Manufacturer	Philips Medical Systems, Best, Netherlands	Siemens Healthcare, Erlangen, Germany	Siemens Healthcare, Erlangen, Germany
Model	3T Achieva	3T Skyra	3T Skyra
Coils	ERC	PPAC	PPAC
T2-weighted sequence (T2WI)			
TR/TE, ms	3,802–5,151/105–115	3,730/121	7,200/96
Resolution, mm^3	0.3 × 0.3 × 3	0.5 × 0.5 × 3	0.6 × 0.6 × 3
Diffusion-weighted imaging (DWI)			
TR/TE, ms	3,751–4,880/50–74	4,700/86	7,900/88
Resolution, mm^3	1.4 × 1.4 × 3	1.6 × 1.6 × 3	1.2 × 1.2 × 3
b-values	0, 500, 1,000, 1,500, 2,000	0, 400, 900, 1,500	50, 600, 1,000, 1,400

TR, reconstruction time; TE, echo time.

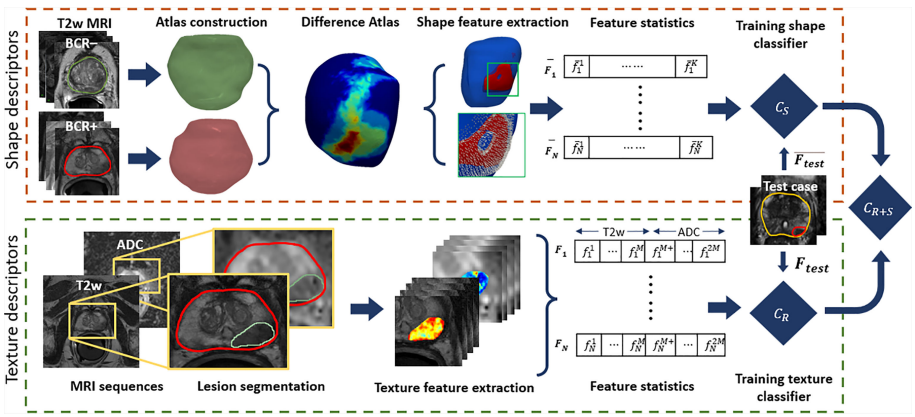


FIGURE 2 | Illustration of the radiomic shape and lesion texture descriptors pipeline used in this study. The shape descriptors are computed from a surface of interest determined from a 3D differential shape atlas between BCR+ and BCR– cohorts on T2WI. These are used to train a model for predicting BCR-free survival and integrated with another model trained using lesion texture descriptors derived from T2WI and ADC maps.

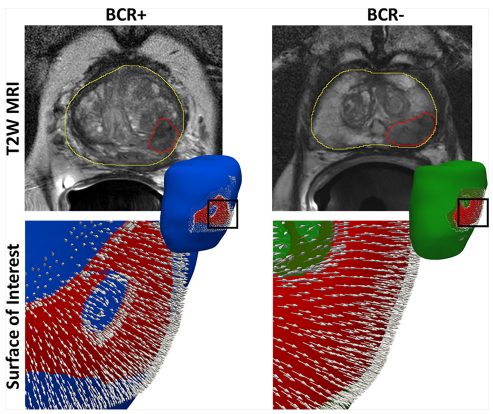


FIGURE 3 | The surface of interest (SOI_C) in red overlaid onto the prostate mesh of a BCR+ patient and a BCR– patient. We observe stronger surface distension in terms of more dense and divergent normal features in the BCR+ patient compared to the BCR– patient.

differences in radiomic shape and texture features between the cohorts without a parametric assumption. Youden's index was used to determine the optimal operating point within receiver-operating characteristics (ROC) analysis to convert the estimated posterior class probabilities to dichotomous labels (BCR+ and BCR−). DeLong's test was used to compare the AUCs between the classification models.

Univariable and multivariable analyses of C_{S+R} were conducted along with pre-RP clinical parameters including age, PI-RADS-v2, PSA, and biopsy Gleason grade (bGG) on D_1. Post-RP pathological Gleason grade (pGG), extraprostatic extension (EPE), seminal vesical invasion (SVI), and positive surgical margins (PSM) were compared on patients from D_2. A penalized Cox proportional hazards (CPH) regression model was used for this purpose given the class imbalance in D_2. The model C_{S+R} was also compared with other nomograms including CAPRA and CAPRA-S and the genomic assay Decipher on subsets of patients for whom sufficient clinical variables were available to estimate BCR-free survival on D_2. Wald test was used to determine statistical significance with p-values under 0.05, and the concordance index (C-index) and hazard ratios (HRs) were computed. Kaplan–Meier survival curves were estimated to determine the differences in BCR-free survival based on predictions of C_{S+R} and comparative nomograms. The log-rank test was used to determine statistical significance ($p < 0.05$).

3 RESULTS

3.1 Evaluation of Shape and Texture Radiomics for Their Association With BCR

The top radiomic shape and tumor texture descriptors based on their random forest Gini feature importance in the training set D_1 are listed in **Table 3**. Shape descriptors were consistent between the sites ($p > 0.05$), whereas all radiomic texture features except Haralick had significant variations ($p < 0.05$). Shape descriptors from the entire prostate mesh resulted in a lower AUC of 0.58 ± 0.08, $p < 0.01$, compared to using those from the SOI_C. Also, shape features from the SOI_C within the mid-gland region improved AUCs compared to those from the entire SOI_C (AUC = 0.75 vs. 0.78, $p = 0.04$), and the mid-gland SOI_C was used in all subsequent experiments. C_S trained using individual SOIs resulted in AUC = 0.64 ± 0.08 on D_1 in distinguishing BCR+ and BCR−, while the consensus SOI_C resulted in improved performance (AUC = 0.78, $p < 0.01$) (**Figure 4**). On the hold-out validation set D_2, the AUC using models trained from individual SOIs was 0.67 ± 0.12, while that from the SOI_C was 0.69 ($p = 0.02$). Inter-reader variations in shape descriptors were evaluated on D_2 and no significant differences were observed ($p > 0.05$) (**Supplementary Figure 2**).

The BCR prediction model (C_R) using radiomic texture descriptors of the PCa lesion resulted in a cross-validation AUC = 0.76 ± 0.09 on D_1 and AUC = 0.70 on D_2. The integrated model (C_{R+S}) resulted in a significantly higher ($p < 0.05$) AUC both on D_1 (0.85 ± 0.08) and D_2 (0.75) compared to both C_S and C_R, respectively. The AUC, sensitivity, and specificity of the models C_S, C_R, and C_{R+S} for predicting BCR status are summarized in **Table 4**.

3.2 Comparison With Clinical Parameters and Extant Nomograms for Predicting bFS at 3 Years

On univariable analysis for predicting bFS, C_{S+R} predictions resulted in the highest HR of 2.91 (95% CI 1.45–11.51, $p = 0.02$) on D_1 compared to other pre-treatment clinical variables including age, PSA, bGG, and PI-RADS-v2. On multivariable analysis, C_{S+R} and bGG showed independent predictive value ($p < 0.05$) for bFS (**Table 5**). Post-RP Gleason grade (pGG), EPE, and PSM had higher HRs of 2.63 (95% CI 1.16–5.93, $p = 0.01$), 2.51 (95% CI 1.06–11.26, $p = 0.04$), and 2.86 (95% CI 1.32–30.26, $p = 0.03$) on D_2. The C-index followed the same trend with C_{S+R} achieving the highest (0.76, $p = 0.03$) among pre-treatment clinical variables; however, it was lower than pGG (0.82, $p = 0.01$) (**Table 5**).

On univariable comparison with extant assays for predicting bFS, C_{S+R} resulted in higher HR (2.11, 95% CI 0.35–11.33, $p = 0.03$) compared to pre-treatment CAPRA (1.8, 95% CI 1.1–3.56, $p = 0.02$) and post-RP Decipher risk (1.41, 95% CI 0.61–2.45, $p = 0.18$); however, post-RP CAPRA-S resulted in comparable HR (2.12, 95% CI 1.2–5.72, $p = 0.03$). Integrating preoperative

TABLE 3 | Variations in the top 5-ranked radiomic shape and texture descriptors according to Gini importance between the cohorts.

	Feature	Gini importance (cohort 1)	Statistics across cohorts			
			Cohort 1 (mean ± std)	Cohort 2 (mean ± std)	Cohort 3 (mean ± std)	p-value (Kruskal–Wallis)
Radiomic shape	Normal_th_kt	1.26	0.68 ± 0.21	0.67 ± 0.16	0.64 ± 0.13	**0.25**
	Normal_phi_sk	1.14	0.52 ± 0.14	0.54 ± 0.09	0.52 ± 0.13	**0.37**
	Normal_phi_kt	0.83	0.63 ± 0.18	0.64 ± 0.12	0.66 ± 0.12	**0.23**
	Curvature_mn	0.87	0.26 ± 0.13	0.26 ± 0.13	0.29 ± 0.17	**0.72**
	Curvature_kt	0.81	0.27 ± 0.12	0.25 ± 0.12	0.28 ± 0.16	**0.7**
Radiomic texture	Haralick_IM_T2	1.41	0.34 ± 0.18	0.31 ± 0.14	0.34 ± 0.18	**0.99**
	Laws_edge_T2	1.07	0.44 ± 0.24	0.43 ± 0.19	0.56 ± 0.15	<0.01
	CoLIAGe_ent_ADC	0.9	0.71 ± 0.16	0.54 ± 0.24	0.71 ± 0.16	0.03
	Haralick_en_ADC	0.88	0.26 ± 0.14	0.51 ± 0.19	0.75 ± 0.13	<0.01
	Gabor_T2	0.84	0.69 ± 0.15	0.42 ± 0.21	0.71 ± 0.12	<0.01

The values in bold indicate no statistical significance p < *0.05. This implies that the feature descriptors were less affected by inter-site variations.*

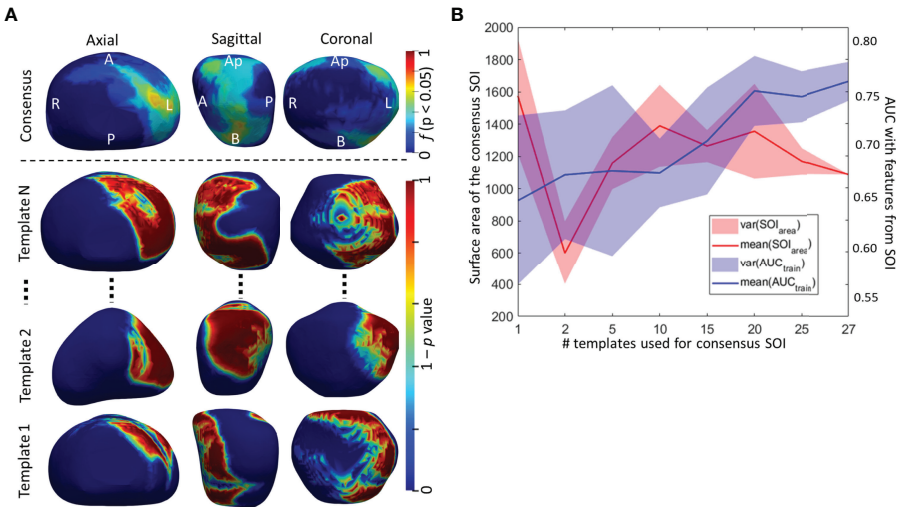


FIGURE 4 | **(A)** Surface of interest (SOI) determined from individual templates and a consensus SOI_C derived by averaging individual SOIs. **(B)** The surface area and AUC from the predictive model trained radiomic shape descriptors as a function of the number of SOIs used in building the consensus SOI_C.

clinical variables with C_{S+R} resulted in improved ($p < 0.05$) C-index (0.82, 95% CI 0.70–0.90) compared to CAPRA-S (0.75, 95% CI 0.69–0.78).

The C-indices also follow the same trend with C_{S+R} and CAPRA-S showing comparable results, and CAPRA and Decipher risk have lower values (**Table 6**). C_{S+R} and CAPRA-S predictions resulted in significant separation in bFS ($p < 0.05$), while CAPRA and Decipher risk showed no significant separation on D_2 (**Figure 5**).

4 DISCUSSION

In this study, we presented prostate shape distension descriptors derived from a 3D shape atlas of the prostate on T2WI and explored their association with PCa BCR post-RP. We observed that the shape descriptors were prognostic of BCR and, in conjunction with radiomic texture features of PCa on T2WI and ADC maps, resulted in an improved BCR prediction model. Since shape descriptors are based on prostate segmentations, they are relatively robust to scanner variations and acquisition parameters compared to texture features. We also validated our approach on hold-out datasets that were acquired from multiple institutions validating robustness and generalizability.

Previously presented radiomic methods (17, 23, 41, 42) for predicting BCR post-RP exclusively focused on MRI texture. Our study is different from these approaches wherein we explored the shape distension of the prostate capsule within a surface of interest as a biomarker for predicting BCR. Patients experiencing BCR tend to have a relatively more aggressive phenotype of PCa (43), and hence, it appears to deform the prostate surface more substantially compared to more indolent cancers that do not result in BCR. Differential distension of the prostate capsule was observed within an SOI located toward the left posterior region, similar to the results reported by Rusu et al. (30) and Ghose et al. (31). We observed that BCR+ patients had more variations in the surface normal orientation arising out of a higher degree of prostate distension compared to BCR− patients (**Figure 3**). Prostate capsular bulge on MRI was found to be a predictor of pathologic EPE after RP by Martini et al. (44). Pathologic EPE has been shown to be a predictor of BCR in several studies (45, 46) and was observed in our study as well. Radiomic texture features including Haralick, Laws, and CoLlAGe features from T2WI and ADC sequences were the top-ranking features that were used in building C_R. These features were also observed to be associated with BCR in previous works by Gnep et al. (22), Shiradkar et al. (23), and Li et al. (17).

TABLE 4 | Classification performance of prostate distension, lesion texture, and integrated classifiers for predicting biochemical recurrence-free survival.

Classifier	D_1			D_2		
	AUC (std)	Sensitivity	Specificity	AUC (std)	Sensitivity	Specificity
C_{Si} (templates)	0.64 (0.58–0.72)	0.63 (0.56–0.79)	0.71 (0.56–0.88)	0.67 (0.52–0.71)	0.49 (0.44–0.65)	0.61 (0.49–0.73)
C_S (consensus)	0.78 (0.69–0.82)	0.67 (0.59–0.71)	0.75 (0.69–0.88)	0.69	0.59	0.65
C_R	0.76 (0.73–0.88)	0.72 (0.63–0.78)	0.86 (0.73–0.88)	0.7	0.6	0.66
C_{S+R}	0.85 (0.76–0.93)	0.65 (0.61–0.75)	0.82 (0.77–0.91)	0.75	0.65	0.58

TABLE 5 | Univariable and multivariable analyses for predicting BCR-free survival with pre-surgical variables (N = 71, N_{BCR+} = 27, N_{BCR-} = 44).

	Parameter	Age	PSA	Biopsy GG	PI-RADS	cT stage	C_{S+R}
Univariable	HR	1.03	1.05	2.12	1.37	2.21	2.91
	Lower 0.95 CI	0.98	1.02	1.55	0.56	1.03	1.45
	Upper 0.95 CI	1.08	1.07	2.9	3.33	5.02	11.51
	C-index	0.53	0.69	0.72	0.64	0.67	0.78
	p-value	0.3	**0.03**	**0.01**	0.16	**0.02**	**0.02**
Multivariable	HR	0.92	1.01	1.21	1.16	2.17	4.51
	Lower 0.95 CI	0.69	0.78	0.34	0.42	1.12	1.87
	Upper 0.95 CI	1.31	1.17	2.65	3.75	4.32	14.55
	C-index				0.85 (95% CI 0.80–0.90)		
	p-value	0.18	**0.07**	**<0.01**	**0.06**	**0.01**	**<0.01**

HR, hazard ratio; PSA, prostate-specific antigen; GG, Gleason grade; CI, confidence interval; PI-RADS, Prostate Imaging Reporting and Data System v2.1. The values in bold indicate statistical significance.

We observed that the shape descriptors were largely consistent between the sites ($p > 0.05$). In terms of radiomic texture descriptors, other than the Haralick feature from PCa on T2W, all the texture features had significant variations between the cohorts ($p < 0.05$) (**Figure 6**; **Table 3**). Several previous studies (26, 47, 48) exploring intersite variations in radiomic texture features have reported texture features to be sensitive to site and scanner variations. Chirra et al. (26) have shown in the context of distinguishing PCa from normal regions that Haralick features remained relatively stable while Law's features showed significant variations which is consistent with our observations. Since radiomic shape descriptors used in our study are based on the shape of the prostate, they remain largely insensitive to site- and scanner-specific variations. This was also observed in a previous study by Merisaari et al. (49) where morphology- and shape-based radiomics were stable in terms of test–retest repeatability compared to texture features.

We observed that when compared to pre-treatment clinical variables including age, PSA, biopsy Gleason grade group (bGG), clinical T stage, or PI-RADS, C_{S+R} predictions resulted in the highest HR and C-index in predicting bFS. We observed that the integrated classifier C_{S+R} was independently prognostic of bFS compared to the other pre-treatment clinical variables and resulted in a high C-index in combination with those clinical variables. Post-RP Gleason grade group and EPE were however predictive of bFS on both univariable and multivariable analyses, also reported in previous studies (50, 51). The performance of the BCR predictive nomograms CAPRA and CAPRA-S in our study was in line with the results reported in previous large-scale validation studies (4). CAPRA score is derived from pre-treatment clinical variables and our radiomic shape and texture classifier C_{S+R} was superior to CAPRA in predicting bFS. CAPRA-S score which includes the post-RP Gleason score resulted in a comparable performance to C_{S+R} which includes pre-treatment parameters alone. On a subset of studies with the availability of Decipher in our study, we found no significant association between Decipher risk score and BCR. This was also observed in a previous study where Decipher down-classified a significant number of patients who experienced BCR (52).

We acknowledge that the study did have its limitations. Firstly, we did not explicitly control for the location of prostate cancer lesions and benign lesions and the presence of extraprostatic extension to explore their effect on prostate distension. Future studies need to be performed, involving controlling for the location of lesions in peripheral transition zones and for the size of lesions. Next, we explored inter-reader variations in prostate segmentation using only two readers and limited the SOI to the mid-gland region. Nevertheless, we observed that the shape features were robust to inter-reader and site-specific variations. We will continue to work on building a statistical model based on multireader segmentations to identify regions, where a high degree of confidence in prostate segmentation can be achieved, more precisely for subsequent shape feature analysis. We also had a smaller sample size in

TABLE 6 | Comparison of the integrated radiomic model C_{S+R} with post-surgical variables, nomograms CAPRA and CAPRA-S, and Decipher risk scores on the validation set (N = 62, N_{BCR+} = 13, N_{BCR-} = 49).

Parameter	HR	Lower 0.95 CI	Upper 0.95 CI	C-index	*p*-value
C_{S+R}	2.1	0.35	11.33	0.76	**0.03**
Pathologic GG	2.6	1.16	5.93	0.82	**0.01**
EPE	2.5	1.06	11.26	0.66	**0.04**
SVI	0.8	0.24	2.57	0.49	0.7
PSM	2.9	1.32	30.2	0.71	**0.03**
CAPRA	1.8	1.1	3.56	0.69	**0.02**
CAPRA-S	2.1	1.2	5.72	0.75	**0.03**
Decipher	1.4	0.61	2.45	0.59	0.18

GG, Gleason grade; EPE, extraprostatic extension; SVI, seminal vesicle invasion; PSM, positive surgical margins; CAPRA, Cancer of the Prostate Risk Assessment (UCSF nomogram); CAPRA-S, post-surgical CAPRA. The p-values in bold indicate statistical significance.

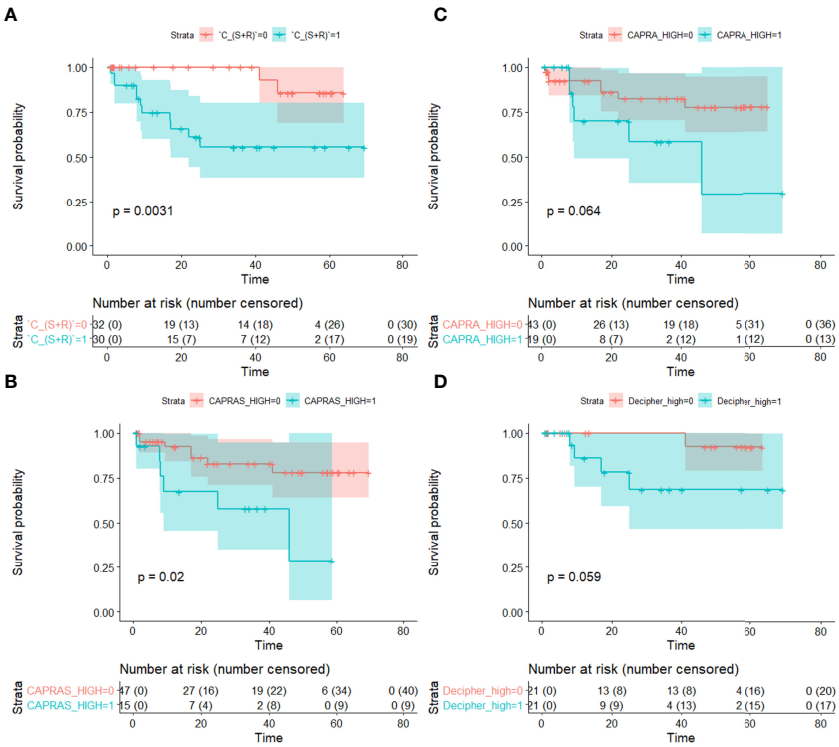


FIGURE 5 | Kaplan–Meier survival curves showing differences in biochemical recurrence-free survival between BCR+ and BCR– patients based on predictions from **(A)** the integrated texture and shape classifier (C_{S+R}), **(B)** CAPRA-S, **(C)** CAPRA and **(D)** Decipher risk. Statistically significant separation in BCR-free survival was observed with C_{S+R} and CAPRA-S ($p < 0.05$). Statistical significance ($p < 0.05$) was established using the log-rank test.

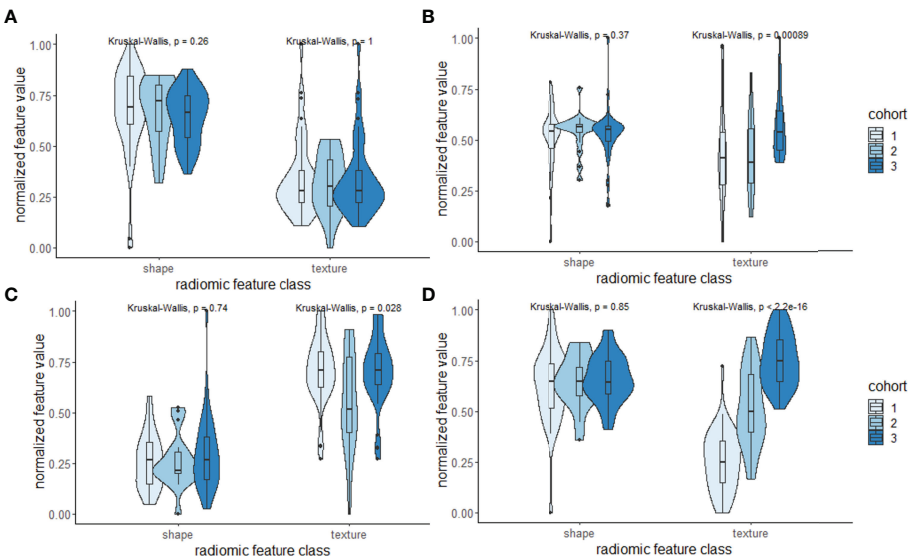


FIGURE 6 | **(A–D)** Violin and box plots of the top-ranked radiomic shape and texture descriptors according to the Gini importance score within the three cohorts used in this study. Shape descriptors tend to be largely consistent and less sensitive to variations across sites compared to texture descriptors.

training and validating our approach, and our validation set had a significant imbalance between the BCR+ and BCR− classes. However, our results still generalized well over the hold-out validation set that was acquired from two different institutions. There were differences in the resolution of T2WI and ADC maps and differences in *b*-values of DWI sequences for generating ADC between the cohorts which were not explicitly accounted for in our study. We will explicitly account for the sensitivity of shape and texture radiomics to variations in resolution in our future study. We also aim to control for positive surgical margins and explore their effect on shape radiomic features in our future work.

In conclusion, radiomic shape descriptors of the prostate capsule derived from T2WI were found to be associated with BCR in our study. In combination with radiomic texture features of prostate cancer lesion from T2WI and ADC sequences, radiomic shape distension features resulted in a better predictor of BCR. Following large-scale validation studies, this approach could potentially be applied to pre-treatment prostate MRI scans at the clinic to provide clinicians with a decision support tool for assessing the risk of BCR, in turn allowing them to make better decisions for treatment management.

DATA AVAILABILITY STATEMENT

The data analyzed in this study are subject to the following licenses/restrictions: The dataset is under the IRB regulations between Case Western Reserve University, Cleveland Clinic, and the University Hospitals Cleveland Medical Center. Requests to access these datasets should be directed to rxs558@case.edu.

AUTHOR CONTRIBUTIONS

RS contributed to the experimental design, data acquisition, analysis, software, writing, and funding. SG provided software and contributed to the analysis. AmM helped with the data acquisition. LL provided software. IH contributed to the analysis and experiments. PF contributed to the statistical analysis. ST and AP contributed to the imaging interpretation, experimental design, and imaging analysis. LP contributed to the data acquisition, funding, and clinical interpretation. AnM contributed to the conception of the idea, experimental design, supervision, writing, and funding. All authors contributed to the article and approved the submitted version.

FUNDING

The research reported in this publication was supported by the National Cancer Institute under award numbers 1U24CA199374-01, R01CA249992-01A1, R01CA202752-01A1, R01CA208236-01A1, R01CA216579-01A1, R01CA220581-01A1, R01CA257612-01A1, 1U01CA239055-01, 1U01CA248226-01, and 1U54CA254566-01; the National Heart, Lung and Blood Institute (1R01HL15127701A1); the National Institute of Biomedical Imaging and Bioengineering (1R43EB028736-01); the National Center for Research Resources under award number 1 C06 RR12463-01; VA Merit Review Award IBX004121A from the United States Department of Veterans Affairs Biomedical Laboratory Research and Development Service; the Office of the Assistant Secretary of Defense for Health Affairs, through the Breast Cancer Research Program (W81XWH-19-1-0668); the Prostate Cancer Research Program (W81XWH-15-1-0558, W81XWH-20-1-0851); the Lung Cancer Research Program (W81XWH-18-1-0440, W81XWH-20-1-0595); the Peer-Reviewed Cancer Research Program (W81XWH-18-1-0404); the Kidney Precision Medicine Project (KPMP) Glue Grant; the Ohio Third Frontier Technology Validation Fund; the Clinical and Translational Science Collaborative of Cleveland (UL1TR0002548) from the National Center for Advancing Translational Sciences (NCATS) component of the National Institutes of Health and NIH Roadmap for Medical Research; The Wallace H. Coulter Foundation Program in the Department of Biomedical Engineering at Case Western Reserve University; the DoD Prostate Cancer Research Program Idea Development Award W81XWH-18-1-0524; and the Clinical and Translational Science Collaborative (CTSC) Cleveland Annual Pilot Award 2020 UL1TR002548. This study also had sponsored research agreements from Bristol Myers-Squibb, Boehringer-Ingelheim, and AstraZeneca.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.841801/full#supplementary-material

REFERENCES

1. Boorjian SA, Thompson RH, Tollefson MK, Rangel LJ, Bergstralh EJ, Blute ML, et al. Long-Term Risk of Clinical Progression After Biochemical Recurrence Following Radical Prostatectomy: The Impact of Time From Surgery to Recurrence. *Eur Urol* (2011) 59(6):893–9. doi: 10.1016/j.eururo.2011.02.026
2. Antonarakis ES, Feng Z, Trock BJ, Humphreys EB, Carducci MA, Partin AW, et al. The Natural History of Metastatic Progression in Men With Prostate-Specific Antigen Recurrence After Radical Prostatectomy: Long-Term Follow-Up. *BJU Int* (2012) 109(1):32–9. doi: 10.1111/j.1464-410X.2011.10422.x
3. Freedland SJ, Humphreys EB, Mangold LA, Eisenberger M, Dorey FJ, Walsh PC, et al. Risk of Prostate Cancer-Specific Mortality Following Biochemical Recurrence After Radical Prostatectomy. *JAMA* (2005) 294(4):433–9. doi: 10.1001/jama.294.4.433
4. Brajtbord JS, Leapman MS, Cooperberg MR. The CAPRA Score at 10 Years: Contemporary Perspectives and Analysis of Supporting Studies. *Eur Urol* (2017) 71(5):705–9. doi: 10.1016/j.eururo.2016.08.065
5. Cooperberg MR, Hilton JF, Carroll PR. The CAPRA-S Score: A Straightforward Tool for Improved Prediction of Outcomes After Radical Prostatectomy. *Cancer* (2011) 117(22):5039–46. doi: 10.1002/cncr.26169
6. Kim HL, Li P, Huang H-C, Deheshi S, Marti T, Knudsen B, et al. Validation of the Decipher Test for Predicting Adverse Pathology in Candidates for

Prostate Cancer Active Surveillance. *Prostate Cancer Prostatic Dis* (2019) 22 (3):399–405. doi: 10.1038/s41391-018-0101-6
7. McKay RR, Berchuck J, Kwak L, Xie W, Silver R, Bubley GJ, et al. Outcomes of Post-Neoadjuvant Intense Hormone Therapy and Surgery for High Risk Localized Prostate Cancer: Results of a Pooled Analysis of Contemporary Clinical Trials. *J Urol* (2021) 205(6):1689–97. doi: 10.1097/JU.0000000000001632
8. Hegde JV, Mulkern RV, Panych LP, Fennessy FM, Fedorov A, Maier SE, et al. Multiparametric MRI of Prostate Cancer: An Update on State-of-the-Art Techniques and Their Performance in Detecting and Localizing Prostate Cancer. *J Magn Reson Imaging* (2013) 37(5):1035–54. doi: 10.1002/jmri.23860
9. Rouvière O, Vitry T, Lyonnet D. Imaging of Prostate Cancer Local Recurrences: Why and How? *Eur Radiol* (2010) 20(5):1254–66. doi: 10.1007/s00330-009-1647-4
10. Vargas HA, Wassberg C, Akin O, Hricak H. Mr Imaging of Treated Prostate Cancer. *Radiology* (2012) 262(1):26–42. doi: 10.1148/radiol.11101996
11. Panebianco V, Barchetti F, Grompone MD, Colarieti A, Salvo V, Cardone G, et al. Magnetic Resonance Imaging for Localization of Prostate Cancer in the Setting of Biochemical Recurrence. *Urol Oncol* (2016) 34(7):303–10. doi: 10.1016/j.urolonc.2016.01.004
12. Ho R, Siddiqui MM, George AK, Frye T, Kilchevsky A, Fascelli M, et al. Preoperative Multiparametric Magnetic Resonance Imaging Predicts Biochemical Recurrence in Prostate Cancer After Radical Prostatectomy. *PloS One* (2016) 11(6):e0157313. doi: 10.1371/journal.pone.0157313
13. McKenna DA, Coakley FV, Westphalen AC, Zhao S, Lu Y, Webb EM, et al. Prostate Cancer: Role of Pretreatment MR in Predicting Outcome After External-Beam Radiation Therapy—Initial Experience. *Radiology* (2008) 247 (1):141–6. doi: 10.1148/radiol.2471061982
14. Riaz N, Afaq A, Akin O, Pei X, Kollmeier MA, Cox B, et al. Pretreatment Endorectal Coil Magnetic Resonance Imaging Findings Predict Biochemical Tumor Control in Prostate Cancer Patients Treated With Combination Brachytherapy and External-Beam Radiotherapy. *Int J Radiat Oncol Biol Phys* (2012) 84(3):707–11. doi: 10.1016/j.ijrobp.2012.01.009
15. Jambor I, Kähkönen E, Taimen P, Merisaari H, Saunavaara J, Alanen K, et al. Prebiopsy Multiparametric 3T Prostate MRI in Patients With Elevated PSA, Normal Digital Rectal Examination, and No Previous Biopsy. *J Magn Reson Imaging* (2015) 41(5):1394–404. doi: 10.1002/jmri.24682
16. Jambor I, Boström PJ, Taimen P, Syvänen K, Kähkönen E, Kallajoki M, et al. Novel Biparametric MRI and Targeted Biopsy Improves Risk Stratification in Men With a Clinical Suspicion of Prostate Cancer (IMPROD Trial). *J Magn Reson Imaging* (2017) 46(4):1089–95. doi: 10.1002/jmri.25641
17. Li L, Shiradkar R, Leo P, Algohary A, Fu P, Tirumani SH, et al. A Novel Imaging Based Nomogram for Predicting Post-Surgical Biochemical Recurrence and Adverse Pathology of Prostate Cancer From Pre-Operative Bi-Parametric MRI. *EBioMedicine* (2021) 63:103163. doi: 10.1016/j.ebiom.2020.103163
18. Litjens GJS, Barentsz JO, Karssemeijer N, Huisman HJ. Clinical Evaluation of a Computer-Aided Diagnosis System for Determining Cancer Aggressiveness in Prostate MRI. *Eur Radiol* (2015) 25(11):3187–99. doi: 10.1007/s00330-015-3743-y
19. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are More Than Pictures, They Are Data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169
20. Algohary A, Viswanath S, Shiradkar R, Ghose S, Pahwa S, Moses D, et al. Radiomic Features on MRI Enable Risk Categorization of Prostate Cancer Patients on Active Surveillance: Preliminary Findings. *J Magn Reson Imaging* (2018) 48(6):1626–36. doi: 10.1002/jmri.25983
21. Viswanath SE, Palumbo D, Chappelow J, Patel P, Bloch BN, Rofsky NM, et al. Empirical Evaluation of Bias Field Correction Algorithms for Computer-Aided Detection of Prostate Cancer on T2w MRI. In: *Spie Medical Imaging*. SPIE. Houston, TX, USA: Proceeings of the SPIE Conference on Medical Imaging (2011). p. 79630V. Available at: http://link.aip.org/link/?PSI/7963/79630V/1.
22. Gnep K, Fargeas A, Gutiérrez-Carvajal RE, Commandeur F, Mathieu R, Ospina JD, et al. Haralick Textural Features on T2 -Weighted MRI are Associated With Biochemical Recurrence Following Radiotherapy for Peripheral Zone Prostate Cancer. *J Magn Reson Imaging* (2016). doi: 10.1002/jmri.25335
23. Shiradkar R, Ghose S, Jambor I, Taimen P, Ettala O, Purysko AS, et al. Radiomic Features From Pretreatment Biparametric MRI Predict Prostate Cancer Biochemical Recurrence: Preliminary Findings. *J Magn Reson Imaging* (2018). doi: 10.1002/jmri.26584
24. Ginsburg SB, Rusu M, Kurhanewicz J, Madabhushi A. *Computer Extracted Texture Features on T2w MRI to Predict Biochemical Recurrence Following Radiation Therapy for Prostate Cancer*. San Diego, CA, USA: Proceedings of the SPIE Conference on Medical Imaging (2014). p. 903509. doi: 10.1117/12.2043937.
25. Palumbo D, Yee B, O'Dea P, Leedy S, Viswanath S, Madabhushi A. Interplay Between Bias Field Correction, Intensity Standardization, and Noise Filtering for T2-Weighted MRI. In: *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference*. Boston MA, USA: Annu Int Conf IEEE Eng Med Biol Soc (2011). pp. 5080–3.
26. Chirra P, Leo P, Yim M, Bloch BN, Rastinehad AR, Purysko A, et al. Multisite Evaluation of Radiomic Feature Reproducibility and Discriminability for Identifying Peripheral Zone Prostate Tumors on MRI. *J Med Imaging (Bellingham)* (2019) 6(2):024502. doi: 10.1117/1.JMI.6.2.024502
27. Gordon V. Measuring the Mechanical Stress Induced by an Expanding Multicellular Tumor System: A Case Study. *Exp Cell Res* (2003) 289(1):58–66. doi: 10.1016/S0014-4827(03)00256-8
28. Jain RK, Martin JD, Stylianopoulos T. The Role of Mechanical Forces in Tumor Growth and Therapy. *Annu Rev BioMed Eng* (2014) 16:321–46. doi: 10.1146/annurev-bioeng-071813-105259
29. Lorenzo G, Hughes TJR, Dominguez-Frojan P, Reali A, Gomez H. Computer Simulations Suggest That Prostate Enlargement Due to Benign Prostatic Hyperplasia Mechanically Impedes Prostate Cancer Growth. *Proc Natl Acad Sci* (2019) 116(4):1152–61. doi: 10.1073/pnas.1815735116
30. Rusu M, Purysko AS, Verma S, Kiechle J, Gollamudi J, Ghose S, et al. Computational Imaging Reveals Shape Differences Between Normal and Malignant Prostates on MRI. *Sci Rep* (2017) 7:41261. doi: 10.1038/srep41261
31. Ghose S, Shiradkar R, Rusu M, Mitra J, Thawani R, Feldman M, et al. Prostate Shapes on Pre-Treatment MRI Between Prostate Cancer Patients Who do and do Not Undergo Biochemical Recurrence are Different: Preliminary Findings. *Sci Rep* (2017) 7(1):15829. doi: 10.1038/s41598-017-13443-8
32. Ghose S, Shiradkar R, Rusu M, Mitra J, Thawani R, Feldman M, et al. Field Effect Induced Organ Distension (Forge) Features Predicting Biochemical Recurrence From Pre-treatment Prostate MRI. In: M Descoteaux, L Maier-Hein, A Franz, P Jannin, DL Collins, S Duchesne, editors. *Medical Image Computing and Computer-Assisted Intervention – Miccai 2017*. Cham: Springer International Publishing (2017). p. 442–9. doi: 10.1007/978-3-319-66185-8_50
33. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magn Reson Imaging* (2012) 30(9):1323–41. doi: 10.1016/j.mri.2012.05.001
34. Turkbey B, Choyke PL. Pirads 2.0: What is New? *Diagn Interv Radiol* (2015) 21(5):382–4. doi: 10.5152/dir.2015.15099
35. Deppe M, Marinell J, Krämer J, Duning T, Ruck T, Simon OJ, et al. Increased Cortical Curvature Reflects White Matter Atrophy in Individual Patients With Early Multiple Sclerosis. *NeuroImage Clin* (2014) 6:475–87. doi: 10.1016/j.nicl.2014.02.012
36. Jain AK, Farrokhnia F. Unsupervised Texture Segmentation Using Gabor Filters. *Pattern Recognit* (1991) 24(12):1167–86. doi: 10.1016/0031-3203(91)90143-S
37. Laws KI. *Rapid Texture Identification*. TF Wiener, editor (1980), pp. 376–81.
38. Haralick RM. Statistical and Structural Approaches to Texture. *Proc IEEE* (1979) 67(5):786–804. doi: 10.1109/PROC.1979.11328
39. Prasanna P, Tiwari P, Madabhushi A. Co-Occurrence of Local Anisotropic Gradient Orientations (Collage): A New Radiomics Descriptor. *Sci Rep* (2016) 6:37241. doi: 10.1038/srep37241
40. Unler A, Murat A, Chinnam RB. Mr2PSO: A Maximum Relevance Minimum Redundancy Feature Selection Method Based on Swarm Intelligence for Support Vector Machine Classification. *Inf Sci* (2011) 181(20):4625–41. doi: 10.1016/j.ins.2010.05.037
41. Park SY, Oh YT, Jung DC, Cho NH, Choi YD, Rha KH, et al. Prediction of Biochemical Recurrence After Radical Prostatectomy With PI-RADS Version 2 in Prostate Cancers: Initial Results. *Eur Radiol* (2016) 26(8):2502–9. doi: 10.1007/s00330-015-4077-5

42. Wong NC, Lam C, Patterson L, Shayegan B. Use of Machine Learning to Predict Early Biochemical Recurrence After Robot-Assisted Prostatectomy. *BJU Int* (2019) 123(1):51–7. doi: 10.1111/bju.14477
43. Gandhi J, Afridi A, Vatsia S, Joshi G, Joshi G, Kaplan SA, et al. The Molecular Biology of Prostate Cancer: Current Understanding and Clinical Implications. *Prostate Cancer Prostatic Dis* (2018) 21(1):22–36. doi: 10.1038/s41391-017-0023-8
44. Martini A, Cumarasamy S, Gupta A, Falagario UG, Shah QN, Beksac AT, et al. Clinical Implications of Prostatic Capsular Abutment or Bulging on Multiparametric Magnetic Resonance Imaging. *Minerva Urol Nefrol* (2019) 71(5):502–7. doi: 10.23736/S0393-2249.19.03370-8
45. Ball MW, Partin AW, Epstein JI. Extent of Extraprostatic Extension Independently Influences Biochemical Recurrence-Free Survival: Evidence for Further pT3 Subclassification. *Urology* (2015) 85(1):161–4. doi: 10.1016/j.urology.2014.08.025
46. Takamatsu K, Matsumoto K, Shojo K, Tanaka N, Takeda T, Morita S, et al. The Prognostic Value of Zonal Origin and Extraprostatic Extension of Prostate Cancer for Biochemical Recurrence After Radical Prostatectomy. *Urol Oncol* (2019) 37(9):575.e19–575.e25. doi: 10.1016/j.urolonc.2019.03.012
47. Hagiwara A, Hori M, Cohen-Adad J, Nakazawa M, Suzuki Y, Kasahara A, et al. Linearity, Bias, Intrascanner Repeatability, and Interscanner Reproducibility of Quantitative Multidynamic Multiecho Sequence for Rapid Simultaneous Relaxometry at 3 T: A Validation Study With a Standardized Phantom and Healthy Controls. *Invest Radiol* (2019) 54 (1):39–47. doi: 10.1097/RLI.0000000000000510
48. Hiremath A, Shiradkar R, Merisaari H, Prasanna P, Ettala O, Taimen P, et al. Test-Retest Repeatability of a Deep Learning Architecture in Detecting and Segmenting Clinically Significant Prostate Cancer on Apparent Diffusion Coefficient (ADC) Maps. *Eur Radiol* (2020). doi: 10.1007/s00330-020-07065-4
49. Merisaari H, Taimen P, Shiradkar R, Ettala O, Pesola M, Saunavaara J, et al. Repeatability of radiomics and machine learning for DWI: Short-term repeatability study of 112 patients with prostate cancer. *Magn Reson Med* (2019). doi: 10.1002/mrm.28058
50. Jambor I, Falagario U, Ratnani P, Perez IM, Demir K, Merisaari H, et al. Prediction of Biochemical Recurrence in Prostate Cancer Patients Who Underwent Prostatectomy Using Routine Clinical Prostate Multiparametric MRI and Decipher Genomic Score. *J Magn Reson Imaging* (2020) 51(4):1075–85. doi: 10.1002/jmri.26928
51. Roux V, Eyraud R, Brureau L, Gourtaud G, Senechal C, Fofana M, et al. [Impact of Gleason Score on Biochemical Recurrence Free Survival After Radical Prostatectomy With Positive Surgical Margins]. *Prog Urol* (2017) 27 (8–9):467–73. doi: 10.1016/j.purol.2017.05.003
52. Cooperberg MR, Davicioni E, Crisan A, Jenkins RB, Ghadessi M, Karnes RJ. Combined Value of Validated Clinical and Genomic Risk Stratification Tools for Predicting Prostate Cancer Mortality in a High-risk Prostatectomy Cohort. *Eur Urol* (2015) 67(2):326–33. doi: 10.1016/j.eururo.2014.05.039

Author Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the U.S. Department of Veterans Affairs, the Department of Defense, or the United States Government.

Conflict of Interest: AnM is an equity holder in Elucid Bioimaging and Inspirata Inc. He is also a scientific advisory consultant for Inspirata Inc. In addition, he has served as a scientific advisory board member for Inspirata Inc., Astra**Z**eneca, Bristol Meyers-Squibb, and Merck. He also has sponsored research agreements with Philips and Inspirata Inc. His technology has been licensed to Elucid Bioimaging and Inspirata Inc. He is also involved in an NIH U24 grant with PathCore Inc. and three different R01 grants with Inspirata Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Shiradkar, Ghose, Mahran, Li, Hubbard, Fu, Tirumani, Ponsky, Purysko and Madabhushi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 12 August 2022
DOI 10.3389/fonc.2022.794555

Check for updates

OPEN ACCESS

EDITED BY
Tingying Peng,
Helmholtz zentrum
München, Germany

REVIEWED BY
Alaa Bessadok,
Helmholtz AI, Germany
Guang Yang,
Imperial College London,
United Kingdom

*CORRESPONDENCE
Jianling Cui
jianlingcui@sina.com
Han Zhang
zhanghan2@shanghaitech.edu.cn.

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 13 October 2021
ACCEPTED 18 July 2022
PUBLISHED 12 August 2022

CITATION
Duan L, Huang H, Sun F, Zhao Z, Wang M, Xing M, Zang Y, Xiu X, Wang M, Yu H, Cui J and Zhang H (2022) Comparing the blood oxygen level–dependent fluctuation power of benign and malignant musculoskeletal tumors using functional magnetic resonance imaging.
Front. Oncol. 12:794555.
doi: 10.3389/fonc.2022.794555

COPYRIGHT
© 2022 Duan, Huang, Sun, Zhao, Wang, Xing, Zang, Xiu, Wang, Yu, Cui and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Comparing the blood oxygen level–dependent fluctuation power of benign and malignant musculoskeletal tumors using functional magnetic resonance imaging

Lisha Duan[1,2], Huiyuan Huang[3,4], Feng Sun[1,2], Zhenjiang Zhao[1,2], Mengjun Wang[1,2], Mei Xing[1,2], Yufeng Zang[3,5], Xiaofei Xiu[6], Meng Wang[7], Hong Yu[1], Jianling Cui[1,2*†] and Han Zhang[3,8*†]

[1]Department of Radiology, the Third Hospital of Hebei Medical University, Shijiazhuang, China, [2]Key Laboratory of Biomechanics of Hebei Province, Shijiazhuang, China, [3]Center for Cognition and Brain Disorders, Hangzhou Normal University, Hangzhou, China, [4]School of Public Health and Management, Guangzhou University of Chinese Medicine, Guangzhou, China, [5]Zhejiang Key Laboratory for Research in Assessment of Cognitive Impairments, Hangzhou, China, [6]Department of Pathology, The Third Hospital of Hebei Medical University, Shijiazhuang, China, [7]Department of Radiology, the First Hospital of Hebei Medical University, Shijiazhuang, China, [8]School of Biomedical Engineering, ShanghaiTech University, Shanghai, China

Purpose: The aim of this study is to compare the blood oxygen level–dependent (BOLD) fluctuation power in 96 frequency points ranging from 0 to 0.25 Hz between benign and malignant musculoskeletal (MSK) tumors *via* power spectrum analyses using functional magnetic resonance imaging (fMRI).

Materials and methods: BOLD-fMRI and T1-weighted imaging (T1WI) of 92 patients with benign or malignant MSK tumors were acquired by 1.5-T magnetic resonance scanner. For each patient, the tumor-related BOLD time series were extracted, and then, the power spectrum of BOLD time series was calculated and was then divided into 96 frequency points. A two-sample *t*-test was used to assess whether there was a significant difference in the powers (the "power" is the square of the BOLD fluctuation amplitude with arbitrary unit) of each frequency point between benign and malignant MSK tumors. The receiver operator characteristic (ROC) analysis was used to assess the diagnostic capability of distinguishing between benign and malignant MSK tumors.

Results: The result of the two-sample *t*-test showed that there was significant difference in the power between benign and malignant MSK tumor at frequency points of 58 (0.1508 Hz, P = 0.036), 59 (0.1534 Hz, P = 0.032), and 95 (0.247 Hz, P = 0.014), respectively. The ROC analysis of mean power of three frequency points showed that the area of under curve is 0.706 (P = 0.009), and the cutoff value is 0.73130. If the power of the tumor greater than or equal to 0.73130 is considered the possibility of benign tumor, then the

diagnostic sensitivity and specificity values are 83% and 59%, respectively. The *post hoc* analysis showed that the merged power of 0.1508 and 0.1534 Hz in benign MSK tumors was significantly higher than that in malignant ones (P = 0.014). The ROC analysis showed that, if the benign MSK tumor was diagnosed with the power greater than or equal to the cutoff value of 1.41241, then the sensitivity and specificity were 67% and 68%, respectively.

Conclusion: The mean power of three frequency points at 0.1508, 0.1534, and 0.247 Hz may potentially be a biomarker to differentiate benign from malignant MSK tumors. By combining the power of 0.1508 and 0.1534 Hz, we could better detect the difference between benign and malignant MSK tumors with higher specificity.

KEYWORDS

blood oxygen level-dependent, musculoskeletal tumors, functional magnetic resonance imaging, benign, malignant

Introduction

Differentiating malignant from benign musculoskeletal (MSK) tumors prior to clinical treatment is especially essential, as it may assist in formulating an appropriate treatment strategy, such as chemotherapy, radiation therapy, or surgery. Imaging evaluation of MSK tumors often involves a combination of modalities, such as plain radiography (1–3), computed tomography (4, 5), magnetic resonance imaging (MRI) (6, 7), scintigraphy (8), or ultrasonography (9, 10), and some functional MRI techniques (11), like diffusion-weighted imaging (DWI) (12), perfusion-weighted imaging (PWI) (13), and magnetic resonance spectroscopy (MRS) (14). These examinations can distinguish between most benign and malignant MSK tumors, but it is still difficult to identify some bone tumors and many soft tissue tumors. Although new developed state-of-the-art methods such as DWI, PWI, and MRS can help to differentiate between benign and malignant tumors, their sensitivity and specificity are low when only one of these techniques is used (15–18). Even biopsy findings may be inconclusive (19, 20), resulting in a delay in diagnosis and curative therapy. Hence, it is still of great value to find a new imaging feature to identify the malignancy of MSK tumors.

The structure of malignant tumor vasculature that is different from the benign one is typically disorganized, with blood vessels that are tortuous, dilated, elongated and leaky and with saccular and dead-end formations (21–24). This kind of vascular architecture may contribute to the presence of regions of persistent low oxygen tension and the temporally variable oxygenation.

Many indexes are used to detect the pathophysiological status of blood flow, one of which is vasomotion (25, 26). Vasomotion is the oscillation of vascular tone, and caliber with frequency points in the range from 1 to 20 min(-1) is seen in most vascular beds (27), which may arise from the activity of the local myogenic mechanism (28).

Intravital microscopy (IVM), laser Doppler flowmetry, and blood oxygen level–dependent (BOLD) can be used to detect fluctuations in blood flow (29–31). The fluctuations can be transformed into frequency domain by either Fourier or wavelet analysis to analyze the fluctuation power (32).

Deoxyhaemoglobin content time series in blood flow could be easily detected by BOLD with high temporal and spatial resolution (33–35). Deoxyhaemoglobin has the T2* effect. The spatial and temporal heterogeneity of spontaneous T2* MR signal fluctuations was first observed in an implanted fibrosarcoma mouse model by Baudelet et al. (36, 37) and later confirmed in tumor xenograft models of colorectal carcinoma by Gonçalves and colleagues (38). However, the occurrence of spontaneous BOLD signal fluctuations has not been well demonstrated in human MSK tumors except in our previous published work (39).

In our previous work, we extracted BOLD signal fluctuations of the 48 MSK tumors by manually selecting the ROIs (regions of interest) in the peripheral and central regions of tumor. The results showed that mean power (the "power" is the square of the BOLD fluctuation amplitude with arbitrary unit) of BOLD fluctuation at a frequency band of 0.073–0.198 Hz was stronger in the peripheral than central regions of the malignant tumors ($P < 0.05$), whereas no such difference for the benign tumors were there (39), which may be due to the vasculature difference between

the malignant and benign MSK tumors. However, it still did not allow us to differentiate malignant from benign MSK tumors just according to this feature (39).

In the current study, we hypothesized that benign MSK tumors have higher fluctuation power of the BOLD signal on some specific frequency than malignant tumors, specifically focused on the frequency band of 0.073–0.198 Hz. MSK tumors were analyzed by using independent component analysis (ICA) that is a method of blind source separation without defining ROI. ICA can decompose the observed data into statistically independent components (40), and it is capable of extracting desired tumor BOLD fluctuation signal from noise (e.g., scanner, physiological, and motion artifacts). After that, we analyzed the tumor BOLD fluctuations *via* power spectrum analyses on 96 specific frequency points ranging from 0 to 0.25 Hz especially in the band of 0.073–0.198 Hz to see if any fluctuation power can be used to differentiate benign from malignant MSK tumor.

Materials and methods

Participants

Ninety-two patients (58 men and 34 women; age, 37.4 ± 18.2 years, 11–73 years) who were hospitalized from March 2009 to May 2014 with primary MSK tumors were included in this study. This sample size included 43 cases from our previous study (28 men and 15 women; age, 36.3 ± 17.6 years, 14–73 years). The final diagnosis of the tumors was determined by histological testing after resection or needle biopsy. All the patients were drug-naive and had no treatment prior to MRI. The inclusion criteria were as follows: 1) the tumor had to be predominately solid; 2) the tumor dimension had to be larger than 3 cm in each of the axial, sagittal, and coronal plains; and 3) the tumor had no extensive intratumor necrosis or bleeding. This study was approved by the Institutional Research Ethics Board in the Third Hospital of Hebei Medical University, China, and all patients signed a written informed consent before the study was carried out.

Magnetic resonance imaging

MRI was performed on a 1.5-T Siemens MR scanning system (Avanto, Siemens, Erlangen, Germany). Patients were imaged in the supine position and were asked to relax. Depending on which body part was under study, image acquisition used one of the following coils: a large body matrix coil, an eight-channel knee coil, or an eight-channel body array coil. T1-weighted imaging (3D T1WI) is based on the three-dimensional turbo fast low-angled shot (Figure 1A) (repetition time/echo time, 1,900/2.97 ms; flip angle, 15°; number of slices, 176; slice thickness/gap, 1/0.5 mm; acquisition matrix, 256 × 246; field of view, 220 × 220 mm), and BOLD fMRI is based on the two-dimensional echo planar imaging (Figure 1B) with 20

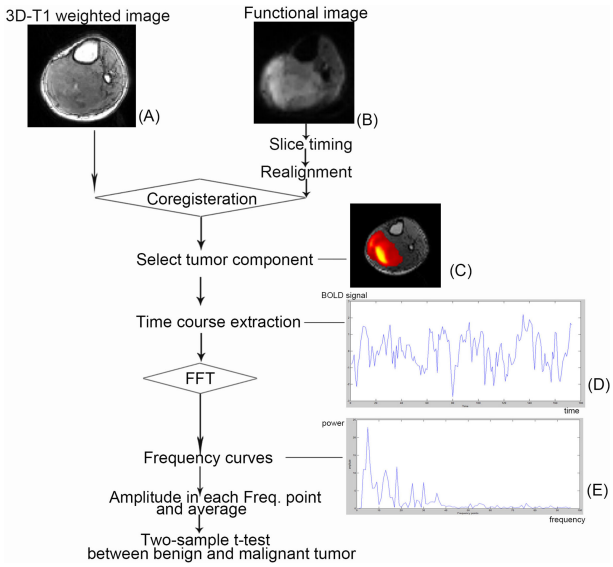


FIGURE 1
Schematic illustration of the course to data processing and analysis. The selected patient is a 57-year-old man with primitive neuroectodermal tumor in his right calf. **(A)** A functional image. **(B)** A 3D T1-weighted image. **(C)** The tumor component selected in independent component analysis. **(D)** The time course after preprocessing and independent component analysis. **(E)** Power spectrum using fast Fourier transformation (FFT).

axial slices (repetition time/echo time, 2,000/40 ms; slice thickness/gap, 5/1 mm; field of view, 220 × 220 mm; acquisition matrix, 64 × 64; voxel size, 3.44 × 3.44 × 6.0 mm; dummy scans, 3; scanning time, 6 min, total of 177 scans).

Data preprocessing

Preprocessing and analysis of fMRI data were carried out using DPARSFA v2.3 (41) and REST v1.8 (42) based on SPM8 (https://www.fil.ion.ucl.ac.uk/spm) and MATLAB 2012a (the MathWorks, Inc., Natick, MA). Figure 2 shows the flowchart of the whole analysis procedure. The following prestatistics processing was applied: 1) the fMRI data was converted to Neuroimaging Informatics Technology Initiative format; 2) the first four frames (i.e., data acquired in first 8 s) were discarded in consideration of machine equilibrium and subjects' adaptation to the scanning, leaving 173 frames for further analysis; 3) slice timing correction was applied to correct within-scan acquisition time differences between slices by using sinc interpolation; 4) patient's body motion was corrected using a six-parameter rigid-body transformation, and the patients that were found to have excessive body motion (>3 mm or >3°) were excluded from further analysis; 5) the 3D high-resolution T1 image of each patient was co-registered to their own averaged BOLD fMRI image to spatially match the two modalities; and 6) finally, the images were smoothed using a 6-mm full-width-at-half-maximum isotropic Gaussian kernel to improve signal-to-noise ratio (SNR).

Independent component analysis

After the preprocessing steps mentioned above, individual-level ICA, a data-driven method, was used to identify the component related to tumor tissue from resting-state fMRI (rs-fMRI) data for each patient. With optimized total component number (i.e., model order) set to 10, individual-level ICA was applied to the preprocessed rs-fMRI data using MICA toolbox (22). In MICA, each subject's preprocessed rs-fMRI data were fed into a one-stage principal component analysis to perform data dimension reduction temporally. After that, ICA decomposition was performed by using an Infomax algorithm. Such an analysis was performed for 100 times, each time with randomized initial value to produce consistent and reliable ICA decompositions (43, 44). Individual subject components were derived from back-reconstruction and z-score transformation. After that, the software (MICA toolbox) took the 10 signal fluctuation components in the form of color pictures. Then, that would be matched to its T1WI image. The color pictures that can basically cover the T1WI images of the tumor were considered to represent the frequency signal of tumor (tumor-related component). This work was done by two experts engaged in medical imaging and data analysis (Lisha Duan and Huiyuan Huang) (Figure 1C).

Power spectrum analyses

For each patient in two groups, the tumor-related BOLD time series was extracted and transformed from temporal to frequency domain with a fast Fourier transformation using MATLAB (Figures 1D, E) (i.e., frequency power spectrum, where the "power" is the square of the BOLD fluctuation amplitude). Because there were 173 time points in the temporal domain, the power spectrum finally accounted for 96 frequency points.

Statistical analyses

A two-sample *t*-test was used to assess whether there was a significant difference in the fluctuation powers at each frequency

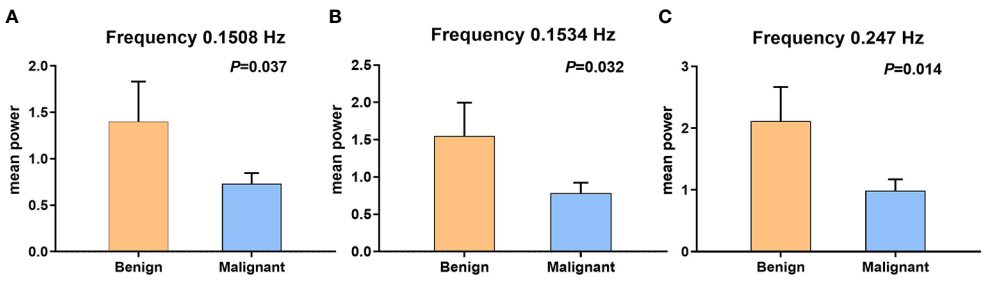


FIGURE 2
The error bars of two sample *t*-test between benign and malignant tumors, which have statistical significance. **(A)** Fifty-eighth frequency point (0.1508 Hz); *P*-value is 0.037. **(B)** Fifty-ninth frequency point (0.1534 Hz); *P*-value is 0.032. **(C)** Ninety-fifth frequency point (0.247 Hz); *P*-value is 0.014.

point (total of 96) between benign and malignant MSK tumors. $P < 0.05$ was considered to indicate a significant difference.

Receiver *operator characteristic analyses*

The difference of the powers at each frequency point between the benign and malignant MSK tumors was tested by using two-sample *t*-test, and then, the receiver operator characteristic (ROC) analysis was performed on SPSS v13.0 software (SPSS Inc., Chicago, IL, USA) to assess the diagnostic capability of distinguish between benign and malignant MSK tumors.

Results

Participants and histopathology

Of the 92 MSK tumor patients, four cases with excessive body motions (translation > 3 mm or rotation > 3°) were excluded for further study. Fourteen cases after ICA were not found clear tumor-related component and thus were removed from further analyses. The failed cases included nine cases from our previous study (five malignant tumors and four benign tumors). The data from the remained 74 patients were put into further analyses (including 34 cases from previous study; six benign tumors and 28 malignant tumors) (Table 1). Histological results showed that 56 cases were malignant tumors and 18 cases were benign tumors.

Statistical analyses

The BOLD fluctuation power of benign MSK tumors was significantly higher than that of malignant ones on the frequency points 0.1508 Hz ($P = 0.037$), 0.1534 Hz ($P = 0.032$), and 0.247 Hz ($P = 0.014$), respectively, by two-sample *t*-test analysis (Figure 2 and Table 2).

The difference of the powers at each frequency point between the benign and malignant MSK tumors was tested by using two-sample *t*-test, and then, the ROC analysis was used to assess the diagnostic capability of distinguish between benign and malignant MSK tumors. The ROC analysis of mean power of three frequency points (0.1508, 0.1534, and 0.247 Hz) showed that the area of under ROC curve is 0.706 [$P = 0.009$; 95% confidence interval (CI), 0.563–0.850], and the cutoff value was 0.73130. If the benign MSK tumor was diagnosed with the mean power value greater than or equal to the cutoff value of 0.73130, then the sensitivity and the specificity were 83% and 59%, respectively (Figure 3).

These data were further processed with *post hoc* analysis. The two frequency points of 0.1508 and 0.1534 Hz may be related to the frequency of myogenic activity in the arteriole (45–

TABLE 1 Demographic and tumor information of 74 cases with malignant or benign musculoskeletal tumors.

Variables (malignant)[a]	Values	Variables (benign)[b]	Values
Cases	56	**Cases**	18
Age (yrs)		**Age (yrs)**	
Mean ± std	37.39 ± 19.89	Mean ± std	36.71 ± 13.34
Range	11–73	Range	14–57
Men-no. (%)	45 (70.31)	**Men-no. (%)**	11 (45.83)
Position-no. (%)		**Position-no. (%)**	
Femur	25 (39.06)	Femur	6 (25.00)
Tibia	12 (18.75)	Tibia	8 (33.33)
Fibula	2 (3.12)	Fibula	2 (8.33)
Humerus	3 (4.69)	Humerus	3 (12.50)
Sacrum	1 (1.56)	Sacrum	n.s.
Os Innominatum	4 (6.25)	Os Innominatum	n.s.
Calcaneus	1 (1.56)	Calcaneus	n.s.
Soft tissue	16 (25.00)	Soft tissue	5 (20.83)
Tumor volume *(cm^3)		**Tumor volume *(cm^3)**	
Mean ± std	163.02 ± 140.88	Mean ± std	159.13 ± 180.96
Range	23.403–546.60	Range	19.60–588.60

[a]Osteosarcoma, malignant fibrous histiocytoma, synovial sarcoma, alveolar sarcoma, Ewing sarcoma, liposarcoma, chondrosarcoma, metastatic adenocarcinoma, primitive neuroectodermal tumor, malignant chondroblastoma, epidermoid leiomyosarcoma, fibrosarcoma, malignant tenosynovial giant cell tumor, lymphoma, malignant peripheral nerve sheath tumor, plasma cell tumors, mesenchymal malignant tumor, pleomorphic sarcoma, and pleomorphic rhabdomyosarcoma. [b]Giant cell tumor of bone, neurofibroma, fibromatosis, giant cell tumor of bone associated with aneurysmal bone cyst, low-grade malignant fibrous histiocytoma, and atypical fibroblastic tumor. yrs, years; no, number. * Selected 37 tumors to calculate their volumes using formula: $(4/3)\pi abc$ (a, b, c = the three radiuses). N.s., not specified.

TABLE 2 Three frequency points possess statistical significance in comparing the difference fluctuation power between the malignant and benign tumors in all frequency points (sum of 96 points).

Freq. points	Benign (N = 18)		Malignant (N = 56)		t-value	*P*-value
	Mean	*Std. Error*	*Mean*	*Std. Error*		
58th (0.1508 Hz)	1.400	0.4310	0.7330	0.1133	2.127	0.0369
59th (0.1534 Hz)	1.550	0.4473	0.7872	0.1374	2.106	0.0321
95th (0.247 Hz)	2.117	0.5484	0.9857	0.1867	2.510	0.0143

Two-sample t-test. The test standard is 0.05.

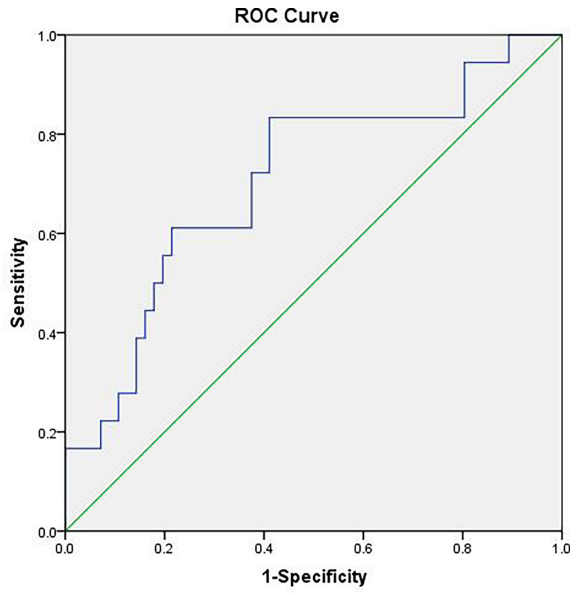


FIGURE 3
The receiver operating characteristic (ROC) curve of mean power of three frequency points (0.1508, 0.1534, and 0.247 Hz). The area under the ROC curve value was 0.706 (P = 0.009). The mean power of musculoskeletal tumor is 0.73130, which means that, if the power value is greater than or equal to 0.73130, then it is considered benign MSK tumor with sensitivity of 83% and specificity of 59%.

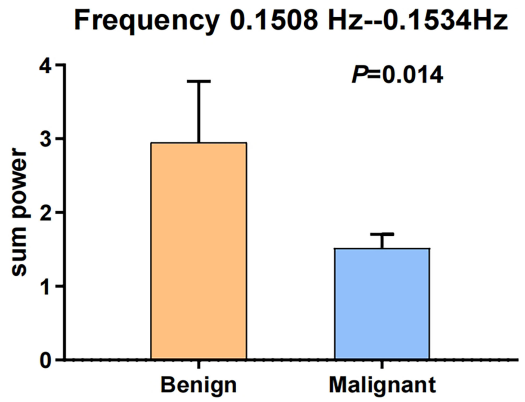


FIGURE 4
The error bars of two sample *t*-test after merging the two frequency points of 0.1508 and 0.1534 Hz. The sum power of benign MSK tumors was significantly higher than that of malignant tumors. *P*-value is 0.014.

49). Hence, we further merged the two frequency points of 0.1508 and 0.1534 Hz to compare the sum powers between benign and malignant MSK tumors. The difference of the sum powers between the benign and malignant MSK tumors was tested by using two-sample *t*-test. The result showed that the sum powers of benign MSK tumors was significantly higher than that of malignant ones (P = 0.014) (Figure 4). The ROC analysis of sum powers of two frequency points (0.1508 and 0.1534 Hz) showed that the area of under ROC curve is 0.661 (P = 0.041; 95% CI, 0.515–0.806), and the cutoff value was 1.41241. If the benign MSK tumor was diagnosed with the sum power value greater than or equal to the cutoff value of 1.41241, then the sensitivity and the specificity were 67% and 68%, respectively (Figure 5).

Discussion

In our previous study, we extracted BOLD signal fluctuations from different regions of the MSK tumor by manually selecting the ROIs, and we characterized them in benign and malignant MSK tumors *via* power spectrum analyses in pre-established low-frequency bands (band-1, 0.01–0.027 Hz; band-2, 0.027–0.073 Hz; band-3, 0.073–0.198 Hz; and band-4, 0.198–0.25 Hz). It was found that BOLD fluctuations at 0.073–0.198 Hz were stronger in the peripheral than central regions of the malignant tumors, and no such difference was observed for the benign tumors (13 benign tumors and 35 malignant tumors) (39).

In current study, the integral MSK tumor-related signal components were extracted by using ICA without defining ROI. After that, we analyzed the powers *via* power spectrum analyses on 96 specific frequency points ranging from 0 to 0.25 HZ. The result showed that there was significant difference in power at the frequency points of 0.1508 Hz (P = 0.037), 0.1534 Hz (P = 0.032), and 0.247 Hz (P = 0.014), respectively, between the benign and malignant MSK tumors after two-sample *t*-test analysis (Table 2).

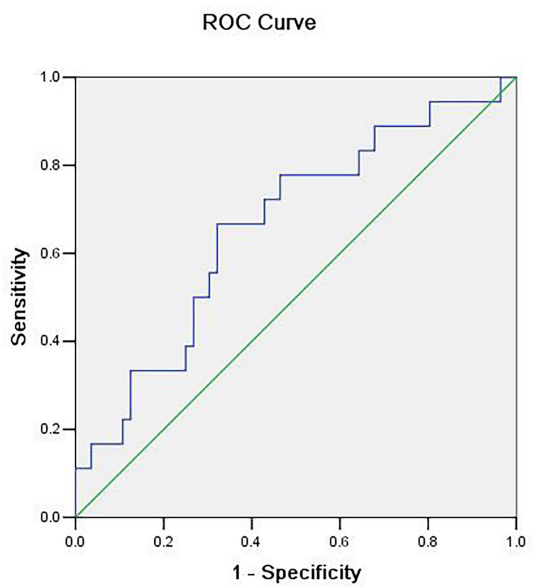


FIGURE 5
The receiver operating characteristic (ROC) curve of sum power of two frequency points (0.1508 and 0.1534 Hz). The area under the ROC curve value was 0.661 (P = 0.041). The cutoff value is 1.41241, which means that, if the sum power value is greater than or equal to 1.41241, then it is considered benign MSK tumor with sensitivity of 67% and specificity of 68%.

BOLD signal usually is very weak and can easily be masked by noise. Although we used the two different signal extraction methods and data analysis methods in our two studies, the significant fluctuation power of two frequency points of 0.1508 Hz (P = 0.037) and 0.1534 Hz (P = 0.032) between the benign and malignant MSK tumors in this study is consistent with that of the significant frequency band of 0.073–0.198 Hz (P < 0.05) in the previous study, which may indicate that the significant fluctuation power around 0.073–0.198Hz is really true.

ICA is a method of blind source separation, which linearly decomposes the observed data into statistically independent components and separates noise and artifacts effectively, making the tumor detection robust to imaging noise (50, 51). It is a multivariate signal processing method to explore the spatiotemporal properties of fMRI data (52–54). The test–retest reliability of ICA on functional component detection generally stays high (55, 56). In addition, previous BOLD fMRI studies have demonstrated the feasibility of rs-fMRI-based automatic tumor tissue identification by using individual ICA (57, 58).

The ICA method is better than the defined ROI method used by previous study, because ICA is capable of extracting desired tumor BOLD time series from noise (e.g., scanner, physiological, and motion artifacts). Moreover, the variability caused by the operator-defined ROI can be avoided by using the ICA method. In particular, the great heterogeneity in tumors may also mask the really signal by using defined ROI method.

Interestingly, the frequency band of 0.073–0.198 Hz in tumors has been suggested to be caused primarily by vascular myogenic activity in vasomotion (45–49). Vasomotion is a phenomenon of blood flow in normal tissue, which *in vivo* is associated with the rhythmic oscillations in vessel diameter (26, 59). These oscillations of the lumen diameter modify blood flow in a corresponding fashion and produce periodical fluctuations known as flowmotion (26, 60).

The frequency of myogenic activity in the arteriole (0.052–0.15Hz) is similar to the first two frequency points (0.1508 and 0.1534 Hz) related to MSK tumors in current study and the band of 0.073–0.198Hz in the previous study, which may indicate that different fluctuation power of frequency points between the malignant and benign MSK tumors is related to the different power of vasomotion between the malignant and benign MSK tumors (39).

The BOLD fluctuation power of benign MSK tumors was significantly higher than that of malignant ones on three frequency points. The difference in the power of the benign and malignant tumors may be explained by the differences of their blood vessel structure (24, 61). The blood vessels, especially smooth muscle cells, are relatively mature in benign tumor compared with the malignant one (21). Relative normal function of smooth muscle may explain higher power of frequency points (0.1508 and 0.1534 Hz) that indicate vasomotion status.

The ROC analyses of mean power of three frequency points (0.150, 0.1534, and 0.247 Hz) showed that the area under the ROC curve value was 0.706 (P = 0.009; 95% CI, 0.563–0.850), and the cutoff value was 0.73130, which means that, if the mean power value is greater than or equal to 0.73130, then it is considered benign MSK tumor, and the sensitivity and specificity are 83% and 59%, respectively. The AUC was 0.706, whose diagnosis accuracy is moderate. This indicates that the mean power of BOLD fluctuation at the frequency points of 0.1508, 0.1534, and 0.247 Hz, as a biomarker can be used to differentiate the benign from malignant MSK tumors with a high sensitivity but low specificity (Figure 4). Our current results are similar to new developed state-of-the-art methods such as DWI, PWI, and MRS for differentiating benign from malignant MSK tumors with low specificity (15–18).

The ROC analyses of sum power of two frequency points (0.1508 and 0.1534 Hz) showed that the cutoff value was 1.41241. If the benign MSK tumor was diagnosed with the sum power value greater than or equal to the cutoff value of 1.41241, then the sensitivity and the specificity were 67% and 68%, respectively. This specificity is higher than that analyzed by mean power of three frequency points (0.1508, 0.1534, and 0.247 Hz) above. It indicated that the frequency points of 0.1508 and 0.1534 Hz could better reflect the true difference between benign and malignant MSK tumors.

The ICA method used to detect tumor-related signal component has the potential to determine the boundary of the

TABLE 3 Advanced functional MRI techniques (PWI, MRS, and DWI) to differentiate benign and malignant musculoskeletal tumor and their sensitivities and specificities.

Advanced technique	Parameter	P-value	Cutoff value	Diagnostic performance		
				Accuracy (%)	Sensitivity (%)	Specificity (%)
PWI	Ktrans (62)	p = 0.028	0.19 min^{-1}	–	79☆	27☆
	Ve (62)	NS	0.37min^{-1}	–	79☆☆	38☆☆
	FP Slope (63)	P < 0.001	45%/s	76	74	77
MRS	Choline SNR (64)	P < 0.001	–	73.3	–	–
	PoA of cho (18)	NS	Presence/absence	–	50☆☆ 82.4☆	61.5☆☆ 64.3☆
DWI	ADC* (65)	P < 0.01	1.45 × 10^{-3} mm^2/s	–	90.9	60
	ADC** (66)	P = 0.004	1.132 × 10^{-9} mm^2/s	–	83	81

Perfusion-weighted imaging (PWI), magnetic resonance spectroscopy (MRS), diffusion-weighted imaging (DWI), the transfer rate of plasma contrast agent to the extracellular extra vascular space (Ktrans), the volume fraction of the extracellular extra vascular space (Ve), first pass (FP), signal-to-noise ratio (SNR), the presence or absence of a choline peak at 3.2 ppm (PoA of cho), apparent diffusion coefficient (ADC), soft tissue tumors (STT), and not statistically significant (NS). *Non-myxoid STT; **non-myxoid, nonhaemosiderin STT; ☆STT; ☆☆ bone tumor.

tumor more precisely and delineate the tumor tissue from the surrounding healthy tissues (57), which may be helpful for making a comprehensive presurgical planning, especially for tumor resection or other image-directed interventions.

The relatively high frequency (0.247 Hz, P = 0.014), whose pathophysiological significance we have not yet understood, may be a noise because of the many factors that affect BOLD imaging. Perhaps, it represents a different kind of pathophysiology, which needs to be explored.

This study had limitations. First, the substantially smaller sample size of benign tumors compared with malignant tumors may induce statistical bias. Second, the sensitivity (83%) and the specificity (59%) are low when our new biomarker is used to differentiate benign from malignant MSK tumors.

In the future, more research studies should be done to improving its sensitivity and specificity. For example, the BOLD signals unrelated to the MSK tumor should be further removed through advanced algorithm. In particular, the more precise methods to acquire the powers of BOLD fluctuations at the frequency points of 0.1508, 0.1534, and 0.247 Hz are needed, of which the first two frequency points may be related to vascular myogenic activity. The sensitivity and the specificity of other functional MRI techniques (PWI, MRS, and DWI) (62–66) (Table 3) in the literature were low too for differentiating benign from malignant MSK tumors. The research studies should be conducted to clarify if the sensitivity and specificity can be promoted when this new biomarker is associated with other biomarkers.

In conclusion, we found that the BOLD fluctuation power of the benign tumors is higher than that of malignant MSK tumors on the frequency points of 0.1508, 0.1534, and 0.247 Hz using the ICA method to extract tumor BOLD time series. By combining the power of 0.1508 and 0.1534 Hz, we could better detect the difference between benign and malignant MSK tumors with higher specificity. The meaning of this difference remains under investigation. However, our finding suggested that BOLD fMRI allows a non-invasive *in vivo* study that avoids contrast agent administration, with a potential to differentiate between malignant and benign MSK tumors. Furthermore, this result was more effective and clear to distinguish between benign and malignant MSK tumors than our previous study.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of the Third Hospital of Hebei Medical University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

LD, JC, and HZ contributed to conception and design of the study. LD wrote the first draft of the manuscript. HH wrote sections of the manuscript. LD, FS, ZZ, MW, and MX collected the required MRI data. XX provided and analyzed the pathological data required for the study. LD collected and assembled the total data. LD, HZ, HH, YZ, MW and HY

performed the statistical analysis and interpretation. LD and JC determined the selection of references and experimental standards and performed the MRI data analysis and interpretation. JC and YZ performed manuscript approval and modifification. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This work was supported by the Medical Science Research Foundation of Hebei Province (No.16277706D) and partially supported by the Key-Area Research and Development Program of Guangdong Province (No. 2021B0101420006), Shanghai Pilot Program for Basic Research-Chinese Academy of Science, Shanghai Branch (No. JCYJ-SHFY-2022-014), the China Postdoctoral Science Foundation (No. 2020M682682), Guangdong Province Philosophy and Social Science Foundation for Youths (No. GD21YXL02), Basic and Applied Basic Research Project of Guangzhou Basic Research Program (No. 202201011471).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Madewell, JE, Ragsdale, BD, Sweet, DE. Radiologic and pathologic analysis of solitary bone lesions. *Radiol Clin North Am* (1981) 4:715–814.

2. Greenspan, A, Jundt, G, Remagen, W. "Fibrogenic, fibroosseous, and fibrohistiocytic lesions,". In: K Barrett, editor. *Differential diagnosis of orthopaedic oncology*. Philadelphia, PA: LWW Press (2007). p. p.257–313.

3. Oudenhoven, LF, Dhondt, E, Kahn, S, Nieborg, A, Kroon, HM, Hogendoorn, PC, et al. Accuracy of radiography in grading and tissue-specific diagnosis-a study of 200 consecutive bone tumors of the hand. *Skeletal Radiol* (2006) 2:78–87. doi: 10.1007/s00256-005-0023-y

4. deSantos, LA, Bernardino, ME, Murray, JA. Computed tomography in the evaluation of osteosarcoma: experience with 25 cases. *AJR Am J Roentgenol* (1979) 4:535–40. doi: 10.2214/ajr.132.4.535

5. Egund, N, Ekelund, L, Sako, M, Persson, B. CT of soft-tissue tumors. *AJR Am J Roentgenol* (1981) 4:725–29. doi: 10.2214/ajr.137.4.725

6. Aisen, AM, Martel, W, Braunstein, EM, McMillin, KI, Phillips, WA, Kling, TF. MRI And CT evaluation of primary bone and soft-tissue tumors. *AJR Am J Roentgenol* (1986) 4:749–56. doi: 10.2214/ajr.146.4.749

7. Yu, H, Cui, JL, Cui, SJ, Sun, YC, Cui, FZ. Differentiating benign from malignant bone tumors using fluid-fluid level features on magnetic resonance imaging. *Korean J Radiol* (2014) 6:757–63. doi: 10.3348/kjr.2014.15.6.757

8. Sanders, TG, Parsons, TW3rd. Radiographic imaging of musculoskeletal neoplasia. *Cancer Control* (2001) 3:221–31. doi: 10.1177/107327480100800302

9. Van der Woude, HJ, Vanderschueren, G. Ultrasound in musculoskeletal tumors with emphasis on its role in tumor follow-up. *Radiol Clin North Am* (1999) 4:753–66. doi: 10.1016/s0033-8389(05)70127-5

10. Widmann, G, Riedl, A, Schoepf, D, Glodny, B, Peer, S, Gruber, H. State-of-the-art HR-US imaging findings of the most frequent musculoskeletal soft-tissue tumors. *Skeletal Radiol* (2009) 7:637–49. doi: 10.1007/s00256-008-0602-9

11. Fayad, LM, Jacobs, MA, Wang, X, Carrino, JA, Bluemke, DA. Musculoskeletal tumors: How to use anatomic, functional, and metabolic MR techniques. *Radiology* (2012) 2:340–56. doi: 10.1148/radiol.12111740

12. Padhani, AR, van Ree, K, Collins, DJ, D'Sa, S, Makris, A. Assessing the relation between bone marrow signal intensity and apparent diffusion coefficient in diffusion-weighted MRI. *AJR Am J Roentgenol* (2013) 1:163–70. doi: 10.2214/AJR.11.8185

13. Guo, J, Reddick, WE, Glass, JO, Ji, Q, Billups, CA, Wu, J, et al. Dynamic contrast-enhanced magnetic resonance imaging as a prognostic factor in predicting event-free and overall survival in pediatric patients with osteosarcoma. *Cancer* (2012) 15:3776–85. doi: 10.1002/cncr.26701

14. Subhawong, TK, Wang, X, Durand, DJ, Jacobs, MA, Carrino, JA, Machado, AJ, et al. Proton MR spectroscopy in metabolic assessment of musculoskeletal lesions. *AJR Am J Roentgenol* (2012) 1:162–72. doi: 10.2214/AJR.11.6505

15. Teixeira, PA, Gay, F, Chen, B, Zins, M, Sirveaux, F, Felblinger, J, et al. Diffusion-weighted magnetic resonance imaging for the initial characterization of non-fatty soft tissue tumors: Correlation between T2 signal intensity and ADC values. *Skeletal Radiol* (2016) 2:263–71. doi: 10.1007/s00256-015-2302-6

16. van Rijswijk, CS, Geirnaerdt, MJ, Hogendoorn, PC, Taminiau, AH, van Coevorden, F, Zwinderman, AH, et al. Soft-tissue tumors: Value of static and dynamic gadopentetate dimeglumine-enhanced MR imaging in prediction of malignancy. *Radiology* (2004) 2:493–502. doi: 10.1148/radiol.2332031110

17. Drapé, JL. Advances in magnetic resonance imaging of musculoskeletal tumours. *Orthop Traumatol Surg Res* (2013) 1:S115–23. doi: 10.1016/j.otsr.2012.12.005

18. Gondim Teixeira, PA, Ledrich, M, Kauffmann, F, Wamba, JM, Felblinger, J, Blum, A, et al. Qualitative 3-T proton MR spectroscopy for the characterization of musculoskeletal neoplasms: Update on diagnostic performance and indications. *AJR Am J Roentgenol* (2017) 6:1312–9. doi: 10.2214/AJR.16.17285

19. Huang, AJ, Kattapuram, SV. Musculoskeletal neoplasms: Biopsy and intervention. *Radiol Clin N Am* (2011) 6:1287–305. doi: 10.1016/j.rcl.2011.07.010

20. Hillen, TJ, Baker, JC, Jennings, JW, Wessell, DE. Image-guided biopsy and treatment of musculoskeletal tumors Travis. *Semin Musculoskelet Radiol* (2013) 2:189–202. doi: 10.1055/s-0033-1343096

21. Konerding, MA, Fait, E, Gaumann, A. 3D microvascular architecture of pre-cancerous lesions and invasive carcinomas of the colon. *Br J Cancer* (2001) 10:1354–62. doi: 10.1054/bjoc.2001.1809

22. Ma, LD, Frassica, FJ, McCarthy, EF, Bluemke, DA, Zerhouni, EA. Benign and malignant musculoskeletal masses: MR imaging differentiation with rim-to-center differential enhancement ratios. *Radiology* (1997) 3:739–44. doi: 10.1148/radiology.202.3.9051028

23. Cao, Y. Tumor angiogenesis and molecular targets for therapy. *Front Biosci (Landmark Ed)* (2009) 14:3962–73. doi: 10.2741/3504

24. Nakamura, G, Arita, S, Tanaka, K, Hioki, K, Tsubura, A, Shikata, N. Correlation between microvessel density and tumor cell proliferation with clinical factors in breast carcinomas. *Oncol Rep* (1999) 5:1023–7. doi: 10.3892/or.6.5.1023

25. Jones, TW. Discovery that the veins of the bat's wing (Which are furnished with valves) are endowed with rhythmical contractility, and that the onward flow of blood is accelerated by such contraction. *Edinb Med Surg J* (1853) 195:367–73.

26. Aalkjær, C, Boedtkjer, D, Matchkov, V. Vasomotion-what is currently thought? *Acta Physiol (Oxf)* (2011) 3:253–69. doi: 10.1111/j.1748-1716.2011.02320.x

27. Aalkjaer, C, Nilsson, H. Vasomotion: Cellular background for the oscillator and for the synchronization of smooth muscle cells. *Br J Pharmacol* (2005) 5:605–16. doi: 10.1038/sj.bjp.0706084

28. Ursino, M, Fabbri, G. Role of the myogenic mechanism in the genesis of microvascular oscillations (vasomotion): Analysis with a mathematical model. *Microvasc Res* (1992) 2:156–77. doi: 10.1016/0026-2862(92)90014-g

29. Ortiz, D, Briceño, JC, Cabrales, P. Microhemodynamic parameters quantification from intravital microscopy videos. *Physiol Meas* (2014) 3:351–67. doi: 10.1088/0967-3334/35/3/351

30. Bocchi, L, Evangelisti, A, Barrella, M, Scatizzi, L, Bevilacqua, M. Recovery of 0.1Hz microvascular skin blood flow in dysautonomic diabetic (type 2) neuropathy by using frequency rhythmic electrical modulation system (FREMS). *Med Eng Phys* (2010) 4:407–13. doi: 10.1016/j.medengphy.2010.02.004

31. Rossi, M, Bradbury, A, Magagna, A, Pesce, M, Taddei, S, Stefanovska, A. Investigation of skin vasoreactivity and blood flow oscillations in hypertensive patients: Effect of short-term antihypertensive treatment. *J Hypertens* (2011) 8:1569–76. doi: 10.1097/HJH.0b013e328348b653

32. Newman, JM, Dwyer, RM, St-Pierre, P, Richards, SM, Clark, MG, Rattigan, S. Decreased microvascular vasomotion and myogenic response in rat skeletal muscle in association with acute insulin resistance. *J Physiol* (2009) 11:2579–88. doi: 10.1113/jphysiol.2009.169011

33. Ogawa, S, Lee, TM, Kay, AR, Tank, DW. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci* (1990) 24:9868–72. doi: 10.1073/pnas.87.24.9868

34. Howe, FA, Robinson, SP, McIntyre, DJ, Stubbs, M, Griffiths, JR. Issues in flow and oxygenation dependent contrast (FLOOD) imaging of tumours. *NMR BioMed* (2001) 7-8:497–506. doi: 10.1002/nbm.716

35. Baudelet, C, Gallez, B. Current issues in the utility of blood oxygen level dependent MRI for the assessment of modulations in tumor oxygenation. *Curr Med Imaging Rev* (2005) 3:229–43. doi: 10.2174/157340505774574754

36. Baudelet, C, Ansiaux, R, Jordan, BF, Havaux, X, Macq, B, Gallez, B. Physiological noise in murine solid tumours using T2*-weighted gradient-echo imaging: A marker of tumour acute hypoxia? *Phys Med Biol* (2004) 15:3389–411. doi: 10.1088/0031-9155/49/15/006

37. Baudelet, C, Cron, GO, Ansiaux, R, Crokart, N, DeWever, J, Feron, O, et al. The role of vessel maturation and vessel functionality in spontaneous fluctuations of T2*-weighted GRE signal within tumors. *NMR BioMed* (2006) 1:69–76. doi: 10.1002/nbm.1002

38. Gonçalves, MR, Johnson, SP, Ramasawmy, R, Pedley, RB, Lythgoe, MF, Walker-Samuel, S. Decomposition of spontaneous fluctuations in tumour oxygenation using BOLD MRI and independent component analysis. *Br J Cancer* (2015) 8:1168–77. doi: 10.1038/bjc.2015.270

39. Duan, LS, Wang, MJ, Sun, F, Zhao, ZJ, Xing, M, Zang, YF, et al. Characterizing the blood oxygen level-dependent fluctuations in musculoskeletal tumours using functional magnetic resonance imaging. *Sci Rep* (2016) 6:36522. doi: 10.1038/srep36522

40. Zhang, H, Zuo, XN, Ma, SY, Zang, YF, Milham, MP, Zhu, CZ. Subject order-independent group ICA (SOI-GICA) for functional MRI data analysis. *Neuroimage* (2010) 4:1414–24. doi: 10.1016/j.neuroimage.2010.03.039

41. Yan, CG, Zang, YF. DPARSF: A MATLAB toolbox for "Pipeline" data analysis of resting-state fMRI. *Front Syst Neurosci* (2010) 4:13. doi: 10.3389/fnsys.2010.00013

42. Song, XW, Dong, ZY, Long, XY, Li, SF, Zuo, XN, Zhu, CZ, et al. REST: A toolkit for resting-state functional magnetic resonance imaging data processing. *PloS One* (2011) 9:e25031. doi: 10.1371/journal.pone.0025031

43. Song, S, Ma, X, Zhan, Y, Zhan, Z, Yao, L, Zhang, J. Bayesian Reconstruction of multiscale local contrast images from brain activity. *J Neurosci Methods* (2013) 1:39–45. doi: 10.1016/j.jneumeth.2013.08.020

44. Wang, L, Chen, D, Yang, X, Olson, JJ, Gopinath, K, Fan, T, et al. Group independent component analysis and functional MRI examination of changes in language areas associated with brain tumors at different locations. *PloS One* (2013) 3:e59657. doi: 10.1371/journal.pone.0059657

45. Meyer, JU, Borgstrom, P, Lindbom, L, Intaglietta, M. Vasomotion patterns in skeletal muscle arterioles during changes in arterial pressure. *Microvasc Res* (1988) 2:193–203. doi: 10.1016/0026-2862(88)90062-3

46. Delgado, E, Marques-Neves, C, Rocha, I, Sales-Luís, J, Silva-Carvalho, L. Myogenic oscillations in rabbit ocular vasculature are very low frequency. *Ophthalmic Res* (2013) 2:123–8. doi: 10.1159/000351629

47. Auer, LM, Gallhofer, B. Rhythmic activity of cat pial vessels *in vivo*. *Eur Neurol* (1981) 6:448–68. doi: 10.1159/000115278

48. Schmidt-Lucke, C, Borgström, P, Schmidt-Lucke, JA. Low frequency flowmotion/(vasomotion) during patho-physiological conditions. *Life Sci* (2002) 23:2713–28. doi: 10.1016/s0024-3205(02)02110-0

49. Slaaf, DW, Tangelder, GJ, Teirlinck, HC, Reneman, RS. Arteriolar vasomotion and arterial pressure reduction in rabbit tenuissimus muscle. *Microvasc Res* (1987) 1:71–80. doi: 10.1016/0026-2862(87)90008-2

50. Feldman, SC, Chu, D, Schulder, M, Pawar, R, Barry, M, Cho, E-, et al. The blood oxygen level-dependent functional MR imaging signal can be used to identify brain tumors and distinguish them from normal tissue. *Am J Neuroradiol* (2009) 2:389–95. doi: 10.3174/ajnr.A1326

51. Chamberland, M, Bernier, M, Fortin, D, Whittingstall, K, Descoteaux, M. 3D interactive tractography-informed resting-state fMRI connectivity. *Front Neurosci* (2015) 9:275. doi: 10.3389/fnins.2015.00275

52. Kiviniemi, V, Kantola, JH, Jauhiainen, J, Hyvärinen, A, Tervonen, O. Independent component analysis of nondeterministic fMRI signal sources. *Neuroimage* (2003):253–60. doi: 10.1016/s1053-8119(03)00097-1

53. van de Ven, VG, Formisano, E, Prvulovic, D, Roeder, CH, Linden, DE. Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. *Hum Brain Mapp* (2004) 3:165–78. doi: 10.1002/hbm.20022

54. Beckmann, CF, DeLuca, M, Devlin, JT, Smith, SM. Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc Lond B Biol Sci* (2005) 1457:1001–13. doi: 10.1098/rstb.2005.1634

55. Zuo, XN, Kelly, C, Adelstein, JS, Klein, DF, Castellanos, FX, Milham, MP. Reliable intrinsic connectivity networks: Test-retest evaluation using ICA and dual regression approach. *Neuroimage* (2010) 3:2163–77. doi: 10.1016/j.neuroimage.2009.10.080

56. Zuo, XN, Xing, XX. Test-retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: A systems neuroscience perspective. *Neurosci Biobehav Rev* (2014) 45:100–18. doi: 10.1016/j.neubiorev.2014.05.009

57. Huang, H, Lu, J, Wu, J, Ding, Z, Chen, S, Duan, L, et al. Tumor tissue detection using blood-Oxygen-Level-Dependent functional MRI based on independent component analysis. *Sci Rep* (2018) 1:1223. doi: 10.1038/s41598-017-18453-0

58. Eguizabal, A, Laughney, AM, García-Allende, PB, Krishnaswamy, V, Wells, WA, Paulsen, KD, et al. Direct identification of breast cancer pathologies using blind separation of label-free localized reflectance measurements. *BioMed Opt Express* (2013) 7:1104–18. doi: 10.1364/BOE.4.001104

59. Verbeuren, TJ, Vallez, MO, Lavielle, G, Bouskela, E. Activation of thromboxane receptors and the induction of vasomotion in the hamster cheek pouch microcirculation. *Br J Pharmacol* (1997) 5:859–66. doi: 10.1038/sj.bjp.0701464

60. Salerud, EG, Tenland, T, Nilsson, GE, Oberg, PA. Rhythmical variations in human skin blood flow. *Int J Microcirc Clin Exp* (1983) 2:91–102.

61. Begum, R, Douglas-Jones, AG, Morgan, JM. Radial intratumoral increase and correlation of microvessels and proliferation in solid breast carcinoma. *Histopathology* (2003) 3:244–53. doi: 10.1046/j.1365-2559.2003.01714.x

62. Leplat, C, Hossu, G, Chen, B, De Verbizier, J, Beaumont, M, Blum, A, et al. Contrast-enhanced 3-T perfusion MRI with quantitative analysis for the characterization of musculoskeletal tumors: Is it worth the trouble? *AJR Am J Roentgenol* (2018) 5:1092–8. doi: 10.2214/AJR.18.19618

63. Verstraete, KL, De Deene, Y, Roels, H, Dierick, A, Uyttendaele, D, Kunnen, M. Benign and malignant musculoskeletal lesions: dynamic contrast-enhanced MR imaging–parametric "first-pass" images depict tissue vascularization and perfusion. *Radiology* (1994) 3:835–43. doi: 10.1148/radiology.192.3.8058957

64. Wang, CK, Li, CW, Hsieh, TJ, Lin, CJ, Chien, SH, Tsai, KB, et al. *In vivo* 1H MRS for musculoskeletal lesion characterization: which factors affect diagnostic accuracy? *NMR BioMed* (2012) 2:359–68. doi: 10.1002/nbm.1758

65. Robba, T, Chianca, V, Albano, D, Clementi, V, Piana, R, Linari, A, et al. Diffusion-weighted imaging for the cellularity assessment and matrix characterization of soft tissue tumour. *Radiol Med* (2017) 11:871–79. doi: 10.1007/s11547-017-0787-x

66. Song, Y, Yoon, YC, Chong, Y, Seo, SW, Choi, YL, Sohn, I, et al. Diagnostic performance of conventional MRI parameters and apparent diffusion coefficient values in differentiating between benign and malignant soft-tissue tumours. *Clin Radiol* (2017) 8:691.e1–10k. doi: 10.1016/j.crad.2017.02.003

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership