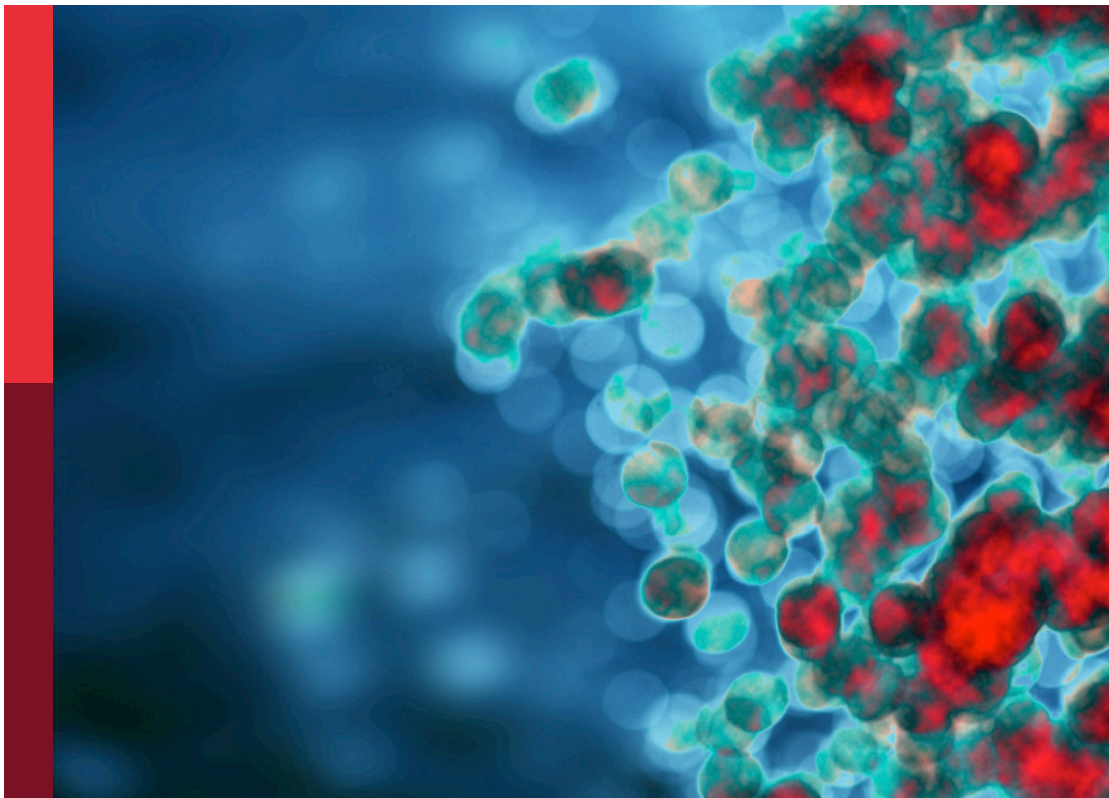
frontiers | Research Topics

Risk and protective factors in the natural history of autoimmunity

Edited by
Christine Gibson Parks, Esther Erdei and Frederick Miller

Published in
Frontiers in Immunology



ISSN 1664-8714
ISBN 978-2-8325-7323-5
DOI 10.3389/978-2-8325-7323-5

Generative AI statement
Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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

Risk and protective factors in the natural history of autoimmunity

Topic editors

Christine Gibson Parks — National Institute of Environmental Health Sciences (NIH), United States

Esther Erdei — University of New Mexico Health Sciences Center, United States

Frederick Miller — Division of Intramural Research, National Institute of Environmental Health Sciences (NIH), United States

Citation

Parks, C. G., Erdei, E., Miller, F., eds. (2026). *Risk and protective factors in the natural history of autoimmunity*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-7323-5

Table of contents

frontiers | Frontiers in Immunology

TYPE Editorial
PUBLISHED 16 January 2026
DOI 10.3389/fimmu.2026.1771091

OPEN ACCESS

EDITED AND REVIEWED BY
Betty Diamond,
Feinstein Institute for Medical Research,
United States

*CORRESPONDENCE
Christine G. Parks
Parks1@mail.nih.gov

RECEIVED 18 December 2025
ACCEPTED 05 January 2026
PUBLISHED 16 January 2026

CITATION
Parks CG, Erdei EO and Miller FW (2026) Editorial: Risk and protective factors in the natural history of autoimmunity.
Front. Immunol. 17:1771091.
doi: 10.3389/fimmu.2026.1771091

Editorial: Risk and protective factors in the natural history of autoimmunity

Christine G. Parks[1]*, Esther O. Erdei[2] and Frederick W. Miller[3]

[1]Chronic Disease Epidemiology Group, Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Durham, NC, United States, [2]University of New Mexico Health Sciences Center College of Pharmacy, Albuquerque, NM, United States, [3]Environmental Autoimmunity Group, Clinical Research Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Bethesda, MD, United States

KEYWORDS

autoimmune diseases, autoimmunity, environmental risk factors, genetic factors, sex differences, triggers and determinants

Editorial on the Research Topic
Risk and protective factors in the natural history of autoimmunity

Autoimmune diseases, the third-most common category after cancer and heart disease, affect at least 5% of the U.S. population (1) and are severe, chronic, and costly to individuals and society. Preclinical or asymptomatic autoimmunity may arise years before diagnosis, occurs in the general population, and appears to be increasing; an example is the rising prevalence of antinuclear antibodies in the U.S. in recent decades (2). However, only some individuals will develop symptoms and pathologies. The articles in this Research Topic focus on risk and protective factors for asymptomatic or preclinical autoimmunity and disease. The relationship between autoimmunity and other diseases, especially cancer and infections, also has important clinical implications. These questions take on greater urgency, given the apparent rise in rates and costs of many autoimmune diseases (3).

Autoimmunity

Clinical suspicion may lead to autoantibody testing; however, a low predictive probability can result in repeated, costly, and unnecessary testing. Barnado et al. addressed this problem using electronic health records of antinuclear antibody (ANA)-positive individuals, finding a greater likelihood of developing autoimmune diseases among those who were younger, female, with higher-titer ANAs, higher platelet counts, disease-specific autoantibodies, and more billing codes for relevant symptoms. In sum, this model is a useful clinical tool for identifying high-risk ANA-positive patients who should undergo further evaluation, while reassuring lower-risk individuals and reducing unnecessary referrals.

While autoantibodies are known to precede numerous autoimmune diseases, the majority of studies lack longitudinal sampling, and the factors that determine progression or regression are poorly understood. In children at risk of developing type 1

diabetes with disease-specific autoantibodies, Carry et al. found differences in DNA methylation, comparing those who progressed to disease, those who maintained autoantibodies, and those who sero-reverted. The candidate genes were related to diet, glucose levels, and immune and pancreatic beta cells. This suggests that environmental factors may contribute to disease risk. Further studies are needed that include exposure data and biomarkers in the progression of preclinical autoimmunity.

In a cross-sectional analysis of cotinine (a marker of cigarette smoke exposure) and ANA prevalence among a representative population sample of the U.S. population, Dinse et al. observed that, over the study periods (1988-1991, 1999-2004, and 2011-2012), the percentage of individuals with ANA was highest (13.3-19.2%) among nonsmokers but non-trending, lower (11.1-15.5%) for "passive" smokers but steadily increasing, and lowest for active smokers, increasing from 7.4% in 1999–2004 to 13.3% in 2011-2012. These findings imply the presence of unmeasured environmental influences on ANA prevalence.

Autoimmunity and cancer

In their review of the cancer risk associated with connective tissue disease, Tonutti et al. explored the multiple, complex interrelationships between these entities. The long-recognized increased cancer risk in many rheumatic conditions may develop for various reasons, including loss of immune tolerance due to oncogenesis, proinflammatory immune activation/autoimmunity that may promote oncogenesis, or immunosuppressive therapies that may decrease cancer surveillance. Conversely, autoimmunity may contribute to the removal of constantly generated neoplasms. Incomplete data support all these theories, and further research is needed. In response, Chen highlighted the need for multidisciplinary collaborations that synthesize different diseases and harmonize methods for detecting autoantibodies.

Sex differences in autoimmunity

Female sex is associated with ANA prevalence and an increased risk of many autoimmune diseases. Investigating a polygenetic risk score for juvenile idiopathic arthritis (JIA), Haftorn et al. examined scores in a population-based study of 238 JIA cases vs. over 73,000 controls. Their investigations into how to best model genetic susceptibilities revealed strong sex differences, suggesting that generalized additive models (GAM) should employ sex stratification, although general linear models can also be applied successfully.

Scofield et al. examined the mechanisms underlying sex differences in immune cells' Toll-Like Receptor (TLR7) signaling using published studies among subjects with SLE (along with other autoimmune diseases). The authors found that the sex bias among patients was explained by specific gene expressions, while inactivations of the X chromosome were also observed. Examined environmental factors included EBV infections and hormonal, mainly estrogen, effects on B cells, suggesting potential molecular pathways.

Environmental and genetic risk factors for autoimmune diseases

In their overview, Choi et al. highlighted diverse non-genetic risk and protective factors for systemic autoimmune rheumatic disorders and the complex interactions that may occur prior to disease development. These risk factors include airborne, waterborne, workplace/occupational, social, and behavioral factors, many of which have changed dramatically in recent decades, which may help explain the increase in autoimmunity and disease. Machine learning methods and multiomics have paved the way for a better understanding of these risk factors, and expansions of these and other new technologies could allow for better preventive approaches in the future.

In a study of JIA, Dåstøl et al. explored the role of seafood and dietary contaminants in the context of a polygenic risk score. While they did not find evidence of associations between estimated intakes of environmental contaminants and risk of JIA based on quantiles of fish intake or proxies for potential heavy metal exposure, patients with low genetic predisposition had stronger, significant associations with environmental toxicants, suggestive of environmentally induced JIA.

Some environmental factors may be considered triggers. Concerns have been raised that autoimmunity may develop following vaccine-specific immune activation and inflammatory responses. In their study of myositis patients, Alhassan et al., in the pre-COVID era, found genetic risk and protective factors for developing myositis within 6 months of vaccination. These factors included human leukocyte antigen (HLA) alleles and immunoglobulin (Ig) allotypes. Large-scale studies with greater genotyping and phenotyping are needed to personalize risk assessment and enhance vaccine safety.

Infections are also possible triggers. In a global network of 74 healthcare organizations and nearly 4 million patients, Hileman et al. investigated the incidence of autoimmune diseases up to 1 year after a diagnosed infection. They found an elevated risk of eight autoimmune diseases in patients diagnosed with COVID-19, especially cutaneous vasculitis, polyarteritis nodosa, and hypersensitivity angiitis. A positive ANA was also more likely and predictive of risk following infection. The authors concluded that SARS-CoV-2 may be a potential trigger for some autoimmune diseases, but the risk may diminish over time, as seen in this study following infection with Omicron variants.

Summary

Taken together, these studies highlight the importance of considering environmental factors and genetic susceptibility in the context of autoimmunity and disease. These contributions suggest the need for well-designed, multidisciplinary studies of

asymptomatic autoimmunity, exposome-genome interactions, and relationships with cancer and infections. The external exposome includes a broader range of features than represented here, including heavy metals, other xenobiotics, along with the psychosocial environment and natural disasters (4, 5), all of which warrant focused future research.

Author contributions

CP: Writing – original draft, Writing – review & editing, Conceptualization. EE: Conceptualization, Writing – review & editing, Writing – original draft. FM: Writing – review & editing, Conceptualization, Writing – original draft.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was supported in part by the Intramural Research Program of the National Institutes of Health, the National Institute of Environmental Health Sciences (Z01-ES049028). The contributions of the NIH author(s) were made as part of their official duties as NIH federal employees, comply with agency policy requirements, and are considered Works of the United States Government. However, the findings and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

1. Abend AH, He I, Bahroos N, Christianakis S, Crew AB, Wise LM, et al. Estimation of prevalence of autoimmune diseases in the United States using electronic health record data. *J Clin Invest*. (2024) 135:e178722. doi: 10.1172/JCI178722

2. Dinse GE, Parks CG, Weinberg CR, Co CA, Wilkerson J, Zeldin DC, et al. Increasing prevalence of antinuclear antibodies in the United States. *Arthritis Rheumatol*. (2022) 74:2032–41. doi: 10.1002/art.42330

3. Miller FW. The increasing prevalence of autoimmunity and autoimmune diseases: an urgent call to action for improved understanding, diagnosis, treatment, and prevention. *Curr Opin Immunol*. (2023) 80:102266. doi: 10.1016/j.coi.2022.102266

4. Kim Y, Koopman JJ, Choi M, Feldman CH, Costenbader KH. Environmental risk factors for systemic lupus erythematosus through the lens of social determinants of health. *Arthritis Care Res (Hoboken)*. (2025) 77:689–99. doi: 10.1002/acr.25497

5. Miller FW, Katsumoto TR. Overview of climate change, pollution, and sustainability in the rheumatic and autoimmune diseases. *Rheum Dis Clin North Am*. (2026) 52:1–12. doi: 10.1016/j.rdc.2025.08.003

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 08 February 2024
DOI 10.3389/fimmu.2024.1337406

OPEN ACCESS

EDITED BY
Frederick Miller,
South China Agricultural University, China

REVIEWED BY
Daniele Sola,
University of Eastern Piedmont, Italy
Giovanni Filocamo,
Fondazione IRCCS CàGranda Ospedale Maggiore Policlinico, Italy

*CORRESPONDENCE
Corrilynn O. Hileman
corrilynn.hileman@case.edu
Grace A. McComsey
grace.mccomsey@uhhospitals.org

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

RECEIVED 13 November 2023
ACCEPTED 19 January 2024
PUBLISHED 08 February 2024

CITATION
Hileman CO, Malakooti SK, Patil N, Singer NG and McComsey GA (2024) New-onset autoimmune disease after COVID-19. *Front. Immunol.* 15:1337406. doi: 10.3389/fimmu.2024.1337406

New-onset autoimmune disease after COVID-19

Corrilynn O. Hileman[1,2*†], Shahdi K. Malakooti[1,2†], Nirav Patil[3], Nora G. Singer[1,2] and Grace A. McComsey[1,3*]

[1]Case Western Reserve University School of Medicine, Cleveland, OH, United States, [2]Department of Medicine, MetroHealth Medical Center, Cleveland, OH, United States, [3]University Hospitals Cleveland Medical Center, Cleveland, OH, United States

Introduction: Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) may trigger autoimmune disease (AD) through initial innate immune activation with subsequent aberrations in adaptive immune cells leading to AD. While there are multiple reports of incident AD diagnosed after COVID-19, the risk in the context of key circulating strains is unknown.

Methods: TriNetX, a global, federated, health research network providing access to electronic medical records across 74 healthcare organizations, was utilized to define an adult cohort between January 1, 2020, and March 3, 2023. Exposure was defined as COVID-19 diagnosis (ICD-10 code or positive laboratory test). Age- and sex-propensity score-matched controls never had COVID-19 diagnosed. Outcomes were assessed 1 month to 1 year after the index date. Patients with AD prior to or within 1 month after the index date were excluded from the primary analysis. Incidence and risk ratios of each AD were assessed.

Results: A total of 3,908,592 patients were included. Of 24 AD patients assessed, adjusted risk ratios for eight AD patients who had COVID-19 were higher compared to those who had no COVID-19. Cutaneous vasculitis (adjusted hazard ratio (aHR): 1.82; 95% CI 1.55–2.13), polyarteritis nodosa (aHR: 1.76; 95% CI 1.15–2.70), and hypersensitivity angiitis (aHR: 1.64; 95% CI 1.12–2.38) had the highest risk ratios. Overall, psoriasis (0.15%), rheumatoid arthritis (0.14%), and type 1 diabetes (0.13%) had the highest incidence during the study period, and of these, psoriasis and diabetes were more likely after COVID-19. The risk of any AD was lower if COVID-19 was diagnosed when Omicron variants were the predominant circulating strains. A positive antinuclear antibody was more likely and predictive of AD after COVID-19.

Discussion: SARS-CoV-2 may be a potential trigger for some AD, but the risk for AD may decrease with time given the apparent lower risk after infection with Omicron variants.

KEYWORDS

autoimmune diseases, COVID-19, autoantibodies, risk factors, antinuclear antibodies

Introduction

Viral infections are often cited as important environmental triggers for autoimmune disease. In the setting of the global COVID-19 pandemic, this is highly relevant, as millions of individuals have been infected with severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). Indeed, there have been multiple reports of newly diagnosed autoimmune diseases after COVID-19 (1, 2). With the breadth of autoimmune disease manifestations, the rarity of many autoimmune diseases, and the lack of accumulated data in the context of COVID-19 variants up to this point, the overall risk of autoimmune disease after COVID-19 including recent key COVID-19 variants is not yet known.

The pathophysiology of autoimmune disease is complex, and the interplay of multiple factors, including genetic and environmental, likely contribute. Simplistically, the host immune response to viral infection has been postulated as a trigger for autoimmunity and includes the production of both interferons (especially alpha interferon), presentation of nuclear contents by "netting" neutrophils, and subsequent maturation of plasmacytoid dendritic cells that act as potent antigen-presenting cells. These virus-induced T cell-mediated autoimmune responses in the right host may lead to autoimmune disease *via* activation of the adaptive immune system resulting in B- and T-cell activations as evidenced first by autoantibodies and later by dysregulated T cells that contribute to overall loss of tolerance to self-antigen.

Interestingly, autoantibodies have been detected in patients with COVID-19 (3, 4). Further, some human proteins have homologous regions with SARS-CoV-2 peptides that could function as autoantigens (5). Additionally, it is clear that in some people with COVID-19, SARS-CoV-2 infection mediates a hyperinflammatory state. Dysregulated inflammasome activation has been implicated in autoimmune disease pathogenesis, and SARS-CoV-2 can activate the inflammasome (nod-like family, pyrin domain-containing 3, or NLRP3), which regulates the secretion of proinflammatory cytokines interleukin 1 beta (IL-1β) and IL-18 (6). More research is needed in this area; however, there is biological plausibility linking SARS-CoV-2 with autoimmunity.

The purpose of this study was to assess the risk of new-onset autoimmune disease within the first year after COVID-19 diagnosis in the context of the predominate circulating variants at the time of infection. We hypothesized that autoimmune disease diagnoses would be higher after COVID-19 infection than in age- and sex-matched controls and that risk would be attenuated when COVID-19 diagnosis occurred when the predominate circulating strains were the Omicron variants. While positive antinuclear antibodies (ANAs) are associated with a variety of autoimmune diseases, a positive ANA test alone is neither sufficient for rheumatologic diagnosis nor predictive of disease development. Therefore, our secondary aim was to evaluate the risk of ANA positivity after COVID-19 and how well ANA positivity predicted the development of new autoimmune diseases within the first year after COVID-19 diagnosis.

Materials and methods

This was a retrospective and population-based cohort study utilizing TriNetX. TriNetX is a global, federated, health research network providing access to electronic medical records including diagnoses, procedures, medications, laboratory values, and genomic information across large healthcare organizations. TriNetX provides de-identified data, transformed into a proprietary data schema, including an extensive data quality and accuracy assessment. This analysis was performed on data drawn from 74 healthcare organizations and completed on March 3, 2023. The study population was defined as adults 18 years of age or older, seen on or after January 1, 2020, with at least one follow-up visit after the index date. Patients with any of the autoimmune diseases evaluated as outcomes in this study diagnosed prior to the index date or within 1 month after the index date were excluded from the primary analysis. The exposure of interest was COVID-19 diagnosis defined by ICD-10 code or positive laboratory test (see Supplementary Table 1 for ICD-10 codes and laboratory tests included). Controls did not have COVID-19 diagnosis (defined by the same criteria) and were propensity score-matched to patients with COVID-19 by age and sex. The index date was defined as the date of COVID-19 diagnosis for the exposed group or first provider visit for any reason during the study period for controls. ANA positivity was defined as nuclear antibody presence in serum by immunofluorescence. This study was approved by the Institution Board Review Committee at Case Western Reserve University/University Hospitals Cleveland Medical Center (STUDY20231104). Written informed consent was waived, as the TriNetX system safeguards patients' privacy in reporting de-identified data.

Outcomes, i.e., incident autoimmune diseases, selected for inclusion were those previously reported in case reports and case series as well as additional autoimmune diseases to attempt to develop as complete a list as possible. Outcomes were defined by ICD-10 codes (see Supplementary Table 1 for ICD-10 codes utilized for each autoimmune disease included). Outcomes were assessed starting 1 month after the index date until 1 year after.

Statistical analysis

The two groups, the exposed or COVID-19 group and the controls or no COVID-19 group, were propensity score-matched by age and sex. Demographics were described by frequency and percent for categorical variables and by mean ± standard deviation for continuous variables for each group. Incidence of each autoimmune disease and risk ratios were assessed for each outcome, i.e., patients with outcome/total patients per group with 95% confidence intervals. Incidence and risk ratios were adjusted for age and sex through propensity score matching as described above. In the primary analyses, patients with any of the autoimmune diseases evaluated as outcomes in this study diagnosed prior to the index date or within 1 month after the index date were excluded. As part of the secondary analyses, patients with the specific outcome being analyzed were excluded

from the analysis for that outcome only. For the secondary analyses, the cohort was propensity score-matched by age and sex prior to excluding the patients with known disease.

Results

Data were available from 1,954,296 adults from January 1, 2020, to March 3, 2023, who lacked prior autoimmune disease and who were diagnosed with COVID-19. Adults without prior autoimmune disease and a diagnosis of COVID-19 during the same time period were propensity score-matched by age and sex at birth to these adults to generate a cohort of 3,908,592 people. Overall, the mean age ± standard deviation (SD) was 48.7 ± 17.9, and 57.7% were women. There were more people from racial and ethnic minorities among those who had COVID-19; however, there were also more people with unknown race and/or ethnicity among those who did not have COVID-19 (see Table 1).

Risk of incident autoimmune disease after COVID-19

The risk of being diagnosed with any autoimmune disease was higher within 1 year following COVID-19 compared to a similar time period in age- and sex-matched controls who did not have COVID-19 diagnosis (adjusted risk ratio (aRR) for any autoimmune disease 1.09 (95% confidence interval or CI 1.07–1.12)). In evaluating each type of autoimmune disease individually, one-third (8 out of 24) of the autoimmune diseases assessed were more likely to be diagnosed after COVID-19. Figure 1 shows adjusted risk ratios for each autoimmune disease assessed. Cutaneous vasculitis (aRR 1.82 (95% CI 1.55–2.13)), polyarteritis nodosa (aRR 1.76 (1.15–2.70)), and hypersensitivity angiitis (aRR 1.64 (1.12–2.38)) were associated with the highest risk. The three autoimmune diseases with the highest incidence during the study period were psoriasis (diagnosed in 5,690 or 0.15%), rheumatoid arthritis (5,618 or 0.14%), and type 1 diabetes mellitus (5,015 or 0.13%). Of these, both psoriasis (aRR 1.23 (95% CI 1.17–1.30)) and type 1 diabetes mellitus (aRR 1.38 (1.31–1.46)) were more common after COVID-19. Graves' disease (0.88 (0.80–0.97)), systemic lupus erythematosus (0.88 (0.80–0.97)), and Crohn's disease (0.84 (0.76–0.92)) were the only diseases less likely to be diagnosed after COVID-19. See Table 2 for the incidence of each autoimmune disease assessed overall as well as by group and adjusted risk ratios.

Of those with COVID-19, the risk of having been hospitalized within 10 days of COVID-19 diagnosis was higher for people who developed autoimmune disease after COVID-19 than people who did not (aRR for hospitalization 1.54 (95% CI 1.44–1.63)) (see Supplementary Figure 1).

In the secondary analysis, people with a specific autoimmune disease prior to or within 1 month after the index date were excluded from the analysis for that outcome. Overall, 4,407,892 individuals were included in this cohort. Supplementary Table 2 shows demographics overall and by COVID-19 exposure group, which were similar to the primary analysis. In this analysis, the risk of being diagnosed with 18 out of the 24 autoimmune diseases

TABLE 1 Demographics overall and by COVID-19 exposure group.

	Overall N = 3,908,592	COVID-19 n = 1,954,296	No COVID-19 n = 1,954,296	p-Value
Age (years) at index	48.7 ± 17.9	48.7 ± 17.9	48.7 ± 17.9	>0.99
Sex, n (%)				
Female	2,253,498 (57.7%)	1,126,749 (57.7%)	1,126,749 (57.7%)	>0.99
Male	1,654,160 (42.3%)	827,080 (42.3%)	827,080 (42.3%)	>0.99
Unknown	934 (<1%)	467 (<1%)	467 (<1%)	>0.99
Race, n (%)				
White	2,171,935 (55.6%)	1,139,355 (58.3%)	1,032,581 (52.8%)	<0.001
Black/African American	516,815 (13.2%)	280,842 (14.4%)	235,973 (12.1%)	<0.001
Asian	94,298 (2.4%)	42,225 (2.2%)	52,073 (2.7%)	<0.001
American Indian, Alaskan Native	13,447 (0.3%)	7,038 (0.4%)	6,409 (0.3%)	<0.001
Pacific Islander	4,947 (0.1%)	2,654 (0.1%)	2,293 (0.1%)	<0.001
Unknown	1,107,149 (28.3%)	482,182 (24.7%)	624,967 (32%)	<0.001
Ethnicity, n (%)				
Not Hispanic/Latino	2,133,897 (54.6%)	1,143,516 (58.5%)	990,381 (50.7%)	<0.001
Hispanic/Latino	296,358 (7.6%)	171,282 (8.8%)	125,076 (6.4%)	<0.001
Unknown	1,478,337 (37.8%)	639,498 (32.7%)	838,839 (42.9%)	<0.001

Groups are matched by propensity score. Propensity scoring included age, male sex, and female sex. People with any prevalent autoimmune diseases prior to or within 1 month after the index date were excluded prior to propensity score matching. Values shown are mean ± standard deviation for continuous variables and frequency (column percent) for categorical variables.

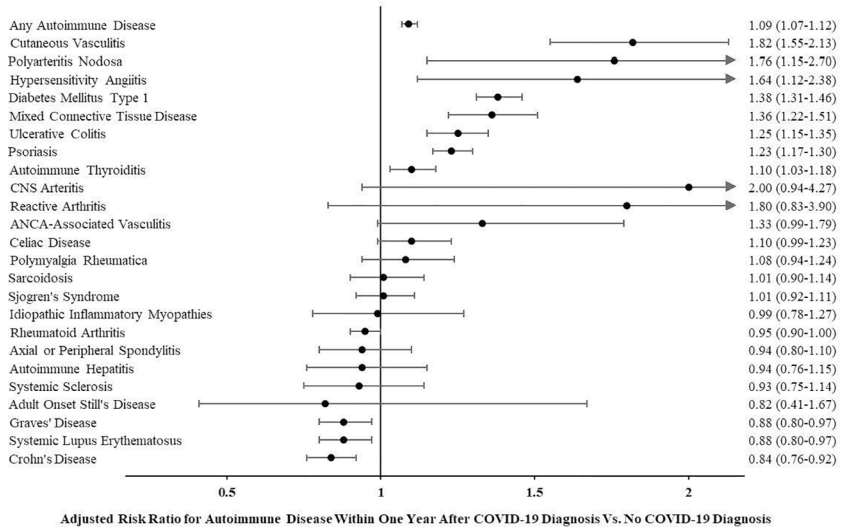


FIGURE 1
Adjusted risk ratio for autoimmune disease within 1 year after COVID-19 diagnosis *vs.* no COVID-19 diagnosis.

TABLE 2 Incident autoimmune diseases overall and by COVID-19 exposure group.

	Overall N = 3,908,592	COVID-19 n = 1,954,296	No COVID-19 n = 1,954,296	Adjusted risk ratio (95% CI)
Any autoimmune disease	31,052 (0.794%)	16,199 (0.829%)	14,853 (0.760%)	1.09 (1.07–1.12)
Autoimmune diseases more likely after COVID-19				
Cutaneous vasculitis	674 (0.017%)	435 (0.022%)	239 (0.012%)	1.82 (1.55–2.13)
Polyarteritis nodosa	91 (0.002%)	58 (0.003%)	33 (0.002%)	1.76 (1.15–2.70)
Hypersensitivity angiitis	116 (0.003%)	72 (0.004%)	44 (0.002%)	1.64 (1.12–2.38)
Type 1 diabetes mellitus	5,014 (0.128%)	2,908 (0.149%)	2,106 (0.108%)	1.38 (1.31–1.46)
Mixed connective tissue disease	1,407 (0.036%)	811 (0.041%)	596 (0.030%)	1.36 (1.22–1.51)
Ulcerative colitis	2,447 (0.063%)	1,359 (0.070%)	1,088 (0.056%)	1.25 (1.15–1.35)
Psoriasis	5,690 (0.146%)	3,137 (0.161%)	2,553 (0.131%)	1.23 (1.17–1.30)
Autoimmune thyroiditis	3,625 (0.093%)	1,902 (0.097%)	1,723 (0.088%)	1.10 (1.03–1.18)
Autoimmune diseases less likely after COVID-19				
Graves' disease	1,524 (0.039%)	713 (0.036%)	811 (0.041%)	0.88 (0.80–0.97)
Systemic lupus erythematosus	1,596 (0.041%)	746 (0.038%)	850 (0.043%)	0.88 (0.80–0.97)
Crohn's disease	1,737 (0.044%)	792 (0.041%)	945 (0.048%)	0.84 (0.76–0.92)
Autoimmune diseases with no associated increased or decreased risk after COVID-19				
CNS arteritis	30 (0.001%)	20 (0.001%)	≤10 (0.001%)	2.00 (0.94–4.27)
Reactive arthritis	28 (0.001%)	18 (0.001%)	≤10 (0.001%)	1.80 (0.83–3.90)
ANCA associated vasculitis	177 (0.005%)	101 (0.005%)	76 (0.004%)	1.33 (0.99–1.79)
Celiac disease	1,313 (0.034%)	689 (0.035%)	624 (0.032%)	1.10 (0.99–1.23)
Polymyalgia rheumatica	834 (0.021%)	433 (0.022%)	401 (0.021%)	1.08 (0.94–1.24)
Sarcoidosis	1,129 (0.029%)	568 (0.029%)	561 (0.029%)	1.01 (0.90–1.14)

(Continued)

10.3389/fimmu.2024.1337406

TABLE 2 Continued

	Overall N = 3,908,592	COVID-19 n = 1,954,296	No COVID-19 n = 1,954,296	Adjusted risk ratio (95% CI)
Sjögren's syndrome	1,811 (0.046%)	910 (0.047%)	901 (0.046%)	1.01 (0.92–1.11)
Idiopathic inflammatory myopathies	261 (0.007%)	130 (0.007%)	131 (0.007%)	0.99 (0.78–1.27)
Rheumatoid arthritis	5,618 (0.144%)	2,740 (0.140%)	2,878 (0.147%)	0.95 (0.90–1.00)
Axial or peripheral spondylitis	616 (0.016%)	298 (0.015%)	318 (0.016%)	0.94 (0.80–1.10)
Autoimmune hepatitis	370 (0.009%)	179 (0.009%)	191 (0.010%)	0.94 (0.76–1.15)
Systemic sclerosis	358 (0.009%)	172 (0.009%)	186 (0.010%)	0.93 (0.75–1.14)
Adult-onset Still's disease	31 (0.001%)	14 (0.001%)	17 (0.001%)	0.82 (0.41–1.67)

Groups are matched by propensity score. Propensity scoring included age, male sex, and female sex. People with any prevalent autoimmune diseases prior to or within 1 month after the index date were excluded from this analysis prior to propensity score matching.
CNS, central nervous system; ANCA, anti-neutrophil cytoplasmic antibodies.

evaluated was higher during the 1 year after COVID-19 diagnosis than during a similar time period in controls. The other autoimmune diseases had similar incidences over 1 year in both groups. See Supplementary Table 3 for the incidence of each autoimmune disease assessed overall and by group with adjusted risk ratios.

Effect of different timeframes on incident autoimmune disease risk after COVID-19

People diagnosed with COVID-19 from July 1, 2021, to November 30, 2021 (during which time the predominant circulating strain of SARS-CoV-2 was the Delta variant), as well as people diagnosed from January 1, 2020, to June 30, 2021 (pre-Delta variant timeframe), had a higher risk of any autoimmune disease when compared to people diagnosed with COVID-19 on or after December 1, 2021. Following December 1, 2021, Omicron SARS-CoV-2 variants were the predominant circulating strains in the USA. The adjusted risk ratio was 0.62 (95% CI 0.59–0.66) for incident autoimmune disease during Omicron *vs.* Delta variant timeframes and 0.66 (95% CI 0.64–0.69) during Omicron *vs.* pre-Delta variant timeframes. See Figure 2 for autoimmune diseases more commonly diagnosed during Delta and pre-Delta than Omicron variant timeframes. None of the autoimmune diseases were more likely to be diagnosed in the first year following COVID-19 infection when the predominant circulating strains were the Omicron variants.

The association of positive ANA test and incident autoimmune disease after COVID-19

In those without a history of autoimmune disease or a positive ANA test, the risk of having a positive ANA test was higher after COVID-19 (980 out of 1,949,921) than for those who did not have COVID-19 (578 out of 1,949,921), adjusting for age and sex (adjusted risk ratio 1.70 (95% CI 1.53–1.88)). Among those with COVID-19, the risk of developing an autoimmune disease was higher for those with a positive ANA test after COVID-19 diagnosis than those without a positive ANA test after adjusting for age and sex (adjusted risk ratio 11.90 (95% CI 6.28–22.55)) (see Supplementary Figure 2 for flowchart with absolute numbers). Specifically, a positive ANA test after COVID-19 was predictive of a new diagnosis for each of the following autoimmune diseases: systemic lupus erythematosus, rheumatoid arthritis, mixed connective tissue disease, Sjögren's syndrome, cutaneous vasculitis, hypersensitivity angiitis, autoimmune thyroiditis, Graves' disease, Crohn's disease, celiac disease, polymyalgia rheumatica, idiopathic inflammatory myopathies, autoimmune hepatitis, and systemic sclerosis (see Table 3).

Effect of any COVID-19 vaccination on incident autoimmune disease risk after COVID-19

Of 1,953,971 patients with COVID-19 and without a history of autoimmune disease, 159,306 (8.2%) had documentation of any COVID-19 vaccination in the TriNetX database. The adjusted risk ratio of any new autoimmune disease diagnosis within 1 year of the index date was 1.18 (95% CI 1.10–1.27) for those who received vaccination *vs.* those with no documentation of vaccination (see Supplementary Figure 3 for flowchart with absolute numbers). In assessing this for each separate autoimmune disease, increased risk post-COVID-19 in those vaccinated compared to those with no documentation of vaccination was only apparent for celiac disease (adjusted risk ratio 1.80 (95% CI 1.22–2.65)), autoimmune thyroiditis (1.70 (1.37–2.11)), Sjögren's syndrome (1.54 (1.16–2.04)), psoriasis (1.42 (1.21–1.66)), and ulcerative colitis (1.40 (1.09–1.80)). The risk of polymyalgia rheumatica was similar regardless of vaccination status. No autoimmune disease was less common post-COVID-19 in those who received vaccination when compared with those with no documentation of vaccination. See Supplementary Table 4 for the incidence of each autoimmune disease assessed overall and by group and adjusted risk ratios.

10.3389/fimmu.2024.1337406

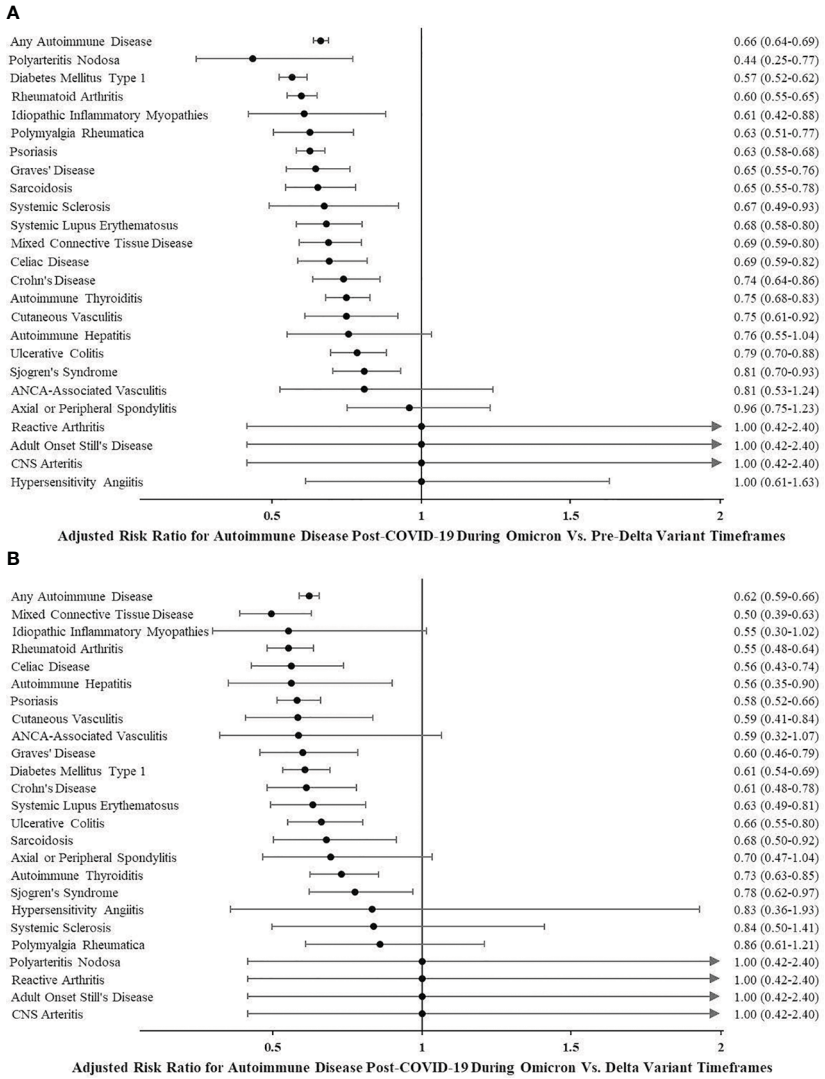


FIGURE 2
(A) Adjusted risk ratio for autoimmune disease post-COVID-19 during Omicron *vs.* pre-Delta variant timeframes. **(B)** Adjusted risk ratio for autoimmune disease post-COVID-19 during Omicron *vs.* Delta variant timeframes.

TABLE 3 Risk of autoimmune disease by ANA status.

	Positive ANA n = 991	Negative or no ANA n = 991	Adjusted risk ratio (95% CI)
Any autoimmune disease	119 (12.000%)	≤10 (1.009%)	11.90 (6.28–22.55)
Systemic lupus erythematosus	28 (2.825%)	0 (0%)	–
Rheumatoid arthritis	32 (3.229%)	≤10 (1.009%)	3.20 (1.58–6.47)
Mixed connective tissue disease	19 (1.917%)	0 (0%)	–
Sjögren's syndrome	17 (1.715%)	0 (0%)	–
Cutaneous vasculitis	≤10 (1.009%)	0 (0%)	–
Hypersensitivity angiitis	≤10 (1.009%)	0 (0%)	–
Autoimmune thyroiditis	≤10 (1.009%)	0 (0%)	–
Graves' disease	≤10 (1.009%)	0 (0%)	–
Crohn's disease	≤10 (1.009%)	0 (0%)	–

(Continued)

TABLE 3 Continued

(Continued)

10.3389/fimmu.2024.1337406

TABLE 3 Continued

	Positive ANA n = 991	Negative or no ANA n = 991	Adjusted risk ratio (95% CI)
Celiac disease	≤10 (1.009%)	0 (0%)	–
Polymyalgia rheumatica	≤10 (1.009%)	0 (0%)	–
Idiopathic inflammatory myopathies	≤10 (1.009%)	0 (0%)	–
Autoimmune hepatitis	≤10 (1.009%)	0 (0%)	–
Systemic sclerosis	≤10 (1.009%)	0 (0%)	–
Type 1 diabetes mellitus	≤10 (1.009%)	≤10 (1.009%)	1 (0.42–2.39)
Ulcerative colitis	≤10 (1.009%)	≤10 (1.009%)	1 (0.42–2.39)
Psoriasis	≤10 (1.009%)	≤10 (1.009%)	1 (0.42–2.39)
Sarcoidosis	≤10 (1.009%)	≤10 (1.009%)	1 (0.42–2.39)
Axial or peripheral spondylitis	≤10 (1.009%)	≤10 (1.009%)	1 (0.42–2.39)
Polyarteritis nodosa	0 (0%)	0 (0%)	–
CNS arteritis	0 (0%)	0 (0%)	–
Reactive arthritis	0 (0%)	0 (0%)	–
ANCA associated vasculitis	0 (0%)	0 (0%)	–
Adult-onset Still's disease	0 (0%)	0 (0%)	–

Groups are matched by propensity score. Propensity scoring included age, male sex, and female sex. People with any prevalent autoimmune diseases or positive ANA test prior to or within 1 month after the index date were excluded from this analysis prior to propensity score matching.
ANA, antinuclear antibody; CNS, central nervous system; ANCA, anti-neutrophil cytoplasmic antibodies.

Discussion

This is the first study of this magnitude of incident autoimmune disease including timeframes where circulating SARS-CoV-2 strains including more recent Omicron variants predominated. We demonstrate that COVID-19 diagnosis was associated with an increased risk of autoimmune disease in the year after infection, and notably, a positive ANA test was more likely after COVID-19 and predicted risk of new-onset autoimmune diseases.

Our finding of an increase in cutaneous vasculitis and polyarteritis after COVID-19 infection is not unexpected given that cutaneous small-vessel vasculitis or capillaritis such as leukocytoclastic vasculitis is frequently associated with perinuclear anti-neutrophil cytoplasmic antibodies (p-ANCA) and antibodies against myeloperoxidase (anti-MPO) and is seen after a variety of infections (streptococcal and hepatitis infection in Henoch–Schonlein purpura) and environmental insults (levamisole in therapeutic and illicit drugs) (7, 8). These diseases also may be accompanied by autoantibodies to cytoplasmic ANCA (c-ANCA) as well as anti-phospholipid antibodies (a major cause of clots following COVID-19 infection). Cutaneous and systemic polyarteritis have both been reported in association with genetic deficiency of adenosine deaminase-2 (DADA2). Adenosine deaminase-2 (ADA2) function(s) are not entirely known, but the protein does appear to contribute to vascular integrity. High levels of ADA2 have been reported in association with infectious and inflammatory illnesses (9) including macrophage activation syndrome in systemic-onset juvenile idiopathic arthritis (10). DADA2 also has been associated with the activation of alpha interferon-associated genes, but any interrelationship between these two states has not been described yet in COVID-19.

Of those with COVID-19, the risk of having been hospitalized within 10 days of COVID-19 diagnosis was higher for people who developed autoimmune disease after COVID-19 than people who did not develop autoimmune disease. This suggests that those who developed autoimmune disease may have had more severe manifestations of COVID-19 than people who did not develop autoimmune disease. Further, prior to vaccination and treatment availability, individuals with genetic risk factors for systemic lupus erythematosus (similar to those with pre-formed anti-cytokine antibodies) may have been at increased risk of life-threatening COVID-19 infection and mortality, potentially resulting in the underrepresentation of systemic lupus erythematosus in COVID-19 survivors when analyzing later timeframes in context of predominant SARS-CoV-2 circulating strains.

The effect of differing circulating strains on the advent of post-acute sequelae of SARS-CoV-2 (PASC) has been previously investigated. Whether PASC is defined by the persistence of symptoms months after a COVID-19 infection or by new-onset health conditions linked to COVID-19, such as new-onset diabetes and cardiovascular disease, Omicron variants appear to be associated with lesser risk than earlier strains (11–15). Our study extends these observations of the potentially less pathogenic nature of Omicron variants to new-onset autoimmune diseases following COVID-19.

Another interesting observation in our study is the apparent increased risk of certain autoimmune diseases after vaccination. In contrast to our observation of autoimmune disease, studies have shown that vaccination is protective against PASC symptoms and incident diabetes after COVID-19 infection (11, 16, 17). New-onset autoimmune phenomena have been described post-COVID-19 vaccination (including immune-mediated hepatitis after COVID-19 vaccination), not all of which have a clear causal relationship established (18–20). Using real-world electronic health record data is more prone to underreporting of vaccination status, which may explain the low vaccination numbers in our study. That said, more studies are needed to better define the risk of autoimmune disease after vaccination.

The finding that ANA positivity is more common after COVID-19 infection and is predictive of new-onset autoimmune disease is

noteworthy. In contrast to the often transient positivity of antiphospholipid antibodies, p-ANCA, anti-MPO, and autoantibodies to rheumatoid arthritis and systemic lupus erythematosus may be present for 8 years or more prior to the onset of incident autoimmune disease (21, 22). This implies that if autoantibodies are present at increased frequency, the incidence of autoantibody disease may rise over longer periods of time, and our estimates of the frequency of autoimmunity may vastly underestimate the effect of COVID-19 on incident autoimmunity in long-term studies. Further, if Omicron variants overly induce lower levels of innate immune activation and subsequently less stimulation of B and T cells, it may take longer to induce similar levels of autoantibodies and T-cell derangements. It is therefore impossible to exclude the possibility that there will be a longer lag in the onset of new autoimmune disease following infection with the Omicron variants compared to Alpha/Delta SARS-CoV-2 and that ultimately, the rates of autoimmune disease may be similar to those seen with all the SARS-CoV-2 variants. Our report undoubtedly includes some patients in the control group who were asymptomatic for COVID-19 and who were neither tested for COVID-19 nor recognized as having COVID-19. This could lead to type II error, as some patients who developed COVID-19-related autoimmune disease may have been misclassified as having been COVID-19 uninfected, leading to smaller effect sizes regarding the risk of autoimmune disease after COVID-19. Importantly, our study also differs from prior reports from TriNetX that required either a positive or negative polymerase chain reaction test to be available for the analyses and focused only on the pre-Omicron era of COVID-19 (January 2020–December 2021) (23). In that way, our results are more generalizable, as they reflect the aggregation of the effects of pre-Delta, Delta, and Omicron variants of COVID-19 with comparisons for incident autoimmune disease and can be re-run at intervals for many years to come.

In addition, ANAs have been classified historically using indirect immunofluorescence assays (IFAs) mostly on the human epidermoid carcinoma (Hep2) cell line to detect nuclear localization, and ANA by IFA was used to define ANA positivity in our study. However, many laboratories have switched to a multiplex assay to measure autoantibodies directly by the target antigen. Clinicians may conclude that an ANA is positive when autoantibodies measure an antigen in the cocktail and may or may not obtain concomitant or subsequent ANA by IFA on the Hep2 cell line to detect nuclear autoantibodies. This is important, as multiplex assays may result in overdiagnosis of autoimmune disease based on a single autoantibody specificity, as positive autoantibody status is sometimes equated to a clinical diagnosis of autoimmunity by non-rheumatologists. This is particularly relevant to anti-U1-ribonucleoprotein (anti-RNP), which accompanies a high-titer ANA by IFA for classification as mixed connective tissue disease but is seen frequently at low levels in the current multiplex technology used across multiple centers. The specificity of low-titer reactivity by multiplex as predictive of future autoimmune disease has never been established, but such low-titer antibodies are observed frequently after COVID-19. Whether any of the anti-SARS-CoV-2 protein antibodies cross-react with antigens in the multiplex assays and therefore wane over time also is a topic ripe for exploration.

A strength of this analysis included the use of TriNetX to analyze data from a large population encompassing 74 healthcare organizations throughout the globe. However, we should recognize that we were unable to adjust for all potential confounders. Limitations of our study are similar to other large studies using electronic health record-derived data and include reliance on ICD-10 diagnoses for disease classification, which may have led to some misclassification. Similarly, we relied on electronic health record data for diagnosis of COVID-19 infection, which may have misclassified some asymptomatic COVID-19 infections as uninfected controls. However, if the latter is true, the effect of COVID-19 infection on incident autoimmune disease may have been underestimated. Further, as with all studies assessing new health conditions after COVID-19, we cannot rule out the possibility that some apparent incident autoimmune diseases were actually flares of previously undiagnosed disease, nor can we rule out potential relation with disproportionate stress. Finally, while Omicron may relate to reduced pathogenesis regardless of vaccination and prior infection status when compared to Delta (23), the role of innate and adaptive immunity in new-onset autoimmune disease after COVID-19 in the context of key variants is yet to be determined. Despite these limitations, however, the major strengths of our study lie in the fact that we have carefully captured the emergence of new-onset autoimmune disease following COVID-19 in a large-scale study. Importantly, our study differs from a prior report from TriNetX that required either a positive or negative polymerase chain reaction test to be included in the analyses and focused only on the pre-Omicron era of COVID-19 (January 2020–December 2021) (24), whereas our report reflects comparison and aggregation of the effects of pre-Delta, Delta, and Omicron variants of COVID-19 on incident autoimmune disease.

In summary, several autoimmune diseases were more likely to be diagnosed within the first year after COVID-19 than in age- and sex-matched controls. The risk of new-onset autoimmune diseases after COVID-19 appears to be attenuated with the more recent Omicron strains. Positive ANA test is more common after COVID-19 and is predictive of incident autoimmune diseases. This suggests that SARS-CoV-2 may be a trigger for certain autoimmune diseases. Future work must focus on longer-term observational cohorts and should assess the persistence and predictive value of different measured autoantibodies.

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 humans were approved by Institution Board Review committee at Case Western Reserve University/ University Hospitals Cleveland Medical Center (STUDY20231104). The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because data from the TriNetX system safeguards patient's privacy in reporting deidentified data.

Author contributions

CH: Writing – original draft, Writing – review & editing. SM: Writing – original draft, Writing – review & editing. NP: Data curation, Formal analysis, Writing – review & editing. NS: Supervision, Writing – review & editing. GM: Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This publication was made possible through funding support of University Hospitals Cleveland Medical Center Clinical Research Center and the Clinical and Translational Science Collaborative of Cleveland which is funded by the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Science Award grant, UM1TR004528. The funding source had no role in the study design, collection, analysis, interpretation of the data, writing of the report or in the decision to submit the paper for publication.

Acknowledgments

Preliminary results were presented at the Conference on Retroviruses and Opportunistic Infections, held in Seattle, Washington, USA, February 19–22, 2023.

Conflict of interest

CH has served as consultant for Theratechnologies and Gilead and has received research grant support from Gilead. GM has served as consultant for Gilead, Merck, Theratechnologies, Janssen, GSK/ViiV, and has received research grants from Gilead, Merck, Janssen and Theratechnologies.

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 author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1337406/full#supplementary-material

References

1. Gracia-Ramos AE, Martin-Nares E, Hernandez-Molina G. New onset of autoimmune diseases following COVID-19 diagnosis. *Cells.* (2021) 10(12). doi: 10.3390/cells10123592

2. Ursini F, Ruscitti P, Addimanda O, Foti R, Raimondo V, Murdaca G, et al. Inflammatory rheumatic diseases with onset after SARS-CoV-2 infection or COVID-19 vaccination: a report of 267 cases from the COVID-19 and ASD group. *RMD Open* (2023) 9(2). doi: 10.1136/rmdopen-2023-003022

3. Chang SE, Feng A, Meng W, Apostolidis SA, Mack E, Artandi M, et al. New-onset IgG autoantibodies in hospitalized patients with COVID-19. *Nat Commun* (2021) 12 (1):5417. doi: 10.1038/s41467-021-25509-3

4. Jiang W, Johnson D, Adekunle R, Heather H, Xu W, Cong X, et al. COVID-19 is associated with bystander polyclonal autoreactive B cell activation as reflected by a broad autoantibody production, but none is linked to disease severity. *J Med Virol* (2023) 95(1):e28134. doi: 10.1002/jmv.28134

5. Mohkhedkar M, Venigalla SSK, Janakiraman V. Untangling COVID-19 and autoimmunity: Identification of plausible targets suggests multi organ involvement. *Mol Immunol* (2021) 137:105–13. doi: 10.1016/j.molimm.2021.06.021

6. Chen IY, Moriyama M, Chang MF, Ichinohe T. Severe acute respiratory syndrome coronavirus viroporin 3a activates the NLRP3 inflammasome. *Front Microbiol* (2019) 10:50. doi: 10.3389/fmicb.2019.00050

7. Ferlazzo B, Barresi G, Puglisi A. [Cutaneous necrotizing vasculitis caused by immunocomplexes during treatment with levamisole]. *Boll Ist Sieroter Milan.* (1983) 62 (1):107–11.

8. Graf J, Lynch K, Yeh CL, Tarter L, Richman N, Nguyen T, et al. Purpura, cutaneous necrosis, and antineutrophil cytoplasmic antibodies associated with levamisole-adulterated cocaine. *Arthritis Rheumatol* (2011) 63(12):3998–4001. doi: 10.1002/art.30590

9. Meyts I, Aksentijevich I. Deficiency of adenosine deaminase 2 (DADA2): updates on the phenotype, genetics, pathogenesis, and treatment. *J Clin Immunol* (2018) 38 (5):569–78. doi: 10.1007/s10875-018-0525-8

10. Lee PY, Schulert GS, Canna SW, Huang Y, Sundel J, Li Y, et al. Adenosine deaminase 2 as a biomarker of macrophage activation syndrome in systemic juvenile idiopathic arthritis. *Ann Rheum Dis* (2020) 79(2):225–31. doi: 10.1136/annrheumdis-2019-216030

11. Thaweethai T, Jolley SE, Karlson EW, Levitan EB, Levy B, McComsey GA, et al. Development of a definition of postacute sequelae of SARS-coV-2 infection. *JAMA.* (2023) 329(22):1934–46. doi: 10.1001/jama.2023.8823

12. Kahlert CR, Strahm C, Gusewell S, Cusini A, Brucher A, Goppel S, et al. Post-acute sequelae after severe acute respiratory syndrome coronavirus 2 infection by viral variant and vaccination status: A multicenter cross-sectional study. *Clin Infect Dis* (2023) 77(2):194–202. doi: 10.1093/cid/ciad143

13. Durstenfeld MS, Peluso MJ, Peyser ND, Lin F, Knight SJ, Djibo A, et al. Factors associated with long COVID symptoms in an online cohort study. *Open Forum Infect Dis* (2023) 10(2):ofad047. doi: 10.1093/ofid/ofad047

14. Ghantous E, Shetrit A, Hochstadt A, Banai A, Lupu L, Levi E, et al. Cardiologic manifestations in omicron-type versus wild-type COVID-19: A systematic echocardiographic study. *J Am Heart Assoc* (2023) 12(3):e027188. doi: 10.1161/JAHA.122.027188

15. Kwan AC, Ebinger JE, Botting P, Navarrette J, Claggett B, Cheng S. Association of COVID-19 vaccination with risk for incident diabetes after COVID-19 infection. *JAMA Netw Open* (2023) 6(2):e2255965. doi: 10.1001/jamanetworkopen.2022.55965

16. Zisis SN, Durieux JC, Mouchati C, Perez JA, McComsey GA. The protective effect of coronavirus disease 2019 (COVID-19) vaccination on postacute sequelae of COVID-19: A multicenter study from a large national health research network. *Open Forum Infect Dis* (2022) 9(7):ofac228. doi: 10.1093/ofid/ofac228

17. Brannock MD, Chew RF, Preiss AJ, Hadley EC, Redfield S, McMurry JA, et al. Long COVID risk and pre-COVID vaccination in an EHR-based cohort study from the RECOVER program. *Nat Commun* (2023) 14(1):2914. doi: 10.1038/s41467-023-38388-7

18. Efe C, Kulkarni AV, Terziroli Beretta-Piccoli B, Magro B, Stattermayer A, Cengiz M, et al. Liver injury after SARS-CoV-2 vaccination: Features of immune-mediated hepatitis, role of corticosteroid therapy and outcome. *Hepatology.* (2022) 76(6):1576–86. doi: 10.1002/hep.32572

19. Chen Y, Xu Z, Wang P, Li XM, Shuai ZW, Ye DQ, et al. New-onset autoimmune phenomena post-COVID-19 vaccination. *Immunology.* (2022) 165(4):386–401. doi: 10.1111/imm.13443

20. Jara LJ, Vera-Lastra O, Mahroum N, Pineda C, Shoenfeld Y. Autoimmune post-COVID vaccine syndromes: does the spectrum of autoimmune/inflammatory syndrome expand? *Clin Rheumatol* (2022) 41(5):1603–9. doi: 10.1007/s10067-022-06149-4

21. Majka DS, Deane KD, Parrish LA, Lazar AA, Baron AE, Walker CW, et al. Duration of preclinical rheumatoid arthritis-related autoantibody positivity increases in subjects with older age at time of disease diagnosis. *Ann Rheum Dis* (2008) 67 (6):801–7. doi: 10.1136/ard.2007.076679

22. Eriksson C, Kokkonen H, Johansson M, Hallmans G, Wadell G, Rantapaa-Dahlqvist S. Autoantibodies predate the onset of systemic lupus erythematosus in northern Sweden. *Arthritis Res Ther* (2011) 13(1):R30. doi: 10.1186/ar3258

23. Thiruvengadam R, Rizvi ZA, Raghavan S, Murugesan DR, Gosain M, Dandotiya J, et al. Clinical and experimental evidence suggest omicron variant of SARS-CoV-2 is inherently less pathogenic than delta variant independent of previous immunity. *Eur J Med Res* (2023) 28(1):421. doi: 10.1186/s40001-023-01373-3

24. Chang R, Yen-Ting Chen T, Wang SI, Hung YM, Chen HY, Wei CJ. Risk of autoimmune diseases in patients with COVID-19: A retrospective cohort study. *EClinicalMedicine.* (2023) 56:101783. doi: 10.1016/j.eclinm.2022.101783

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 20 March 2024
DOI 10.3389/fimmu.2024.1384229

OPEN ACCESS

EDITED BY
Frederick Miller,
National Institute of Environmental Health Sciences (NIH), United States

REVIEWED BY
Edward K. L. Chan,
University of Florida, United States
Kathryn Connelly,
Monash University, Australia

*CORRESPONDENCE
April Barnado
april.barnado@vumc.org

RECEIVED 08 February 2024
ACCEPTED 08 March 2024
PUBLISHED 20 March 2024

CITATION
Barnado A, Moore RP, Domenico HJ, Green S, Camai A, Suh A, Han B, Walker K, Anderson A, Caruth L, Katta A, McCoy AB and Byrne DW (2024) Identifying antinuclear antibody positive individuals at risk for developing systemic autoimmune disease: development and validation of a real-time risk model.
Front. Immunol. 15:1384229.
doi: 10.3389/fimmu.2024.1384229

Identifying antinuclear antibody positive individuals at risk for developing systemic autoimmune disease: development and validation of a real-time risk model

April Barnado[1,2]*, Ryan P. Moore[3], Henry J. Domenico[3], Sarah Green[1], Alex Camai[1], Ashley Suh[1], Bryan Han[1], Katherine Walker[1], Audrey Anderson[1], Lannawill Caruth[1], Anish Katta[1], Allison B. McCoy[2] and Daniel W. Byrne[2,3]

[1]Division of Rheumatology & Immunology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, United States, [2]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, United States, [3]Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN, United States

Objective: Positive antinuclear antibodies (ANAs) cause diagnostic dilemmas for clinicians. Currently, no tools exist to help clinicians interpret the significance of a positive ANA in individuals without diagnosed autoimmune diseases. We developed and validated a risk model to predict risk of developing autoimmune disease in positive ANA individuals.

Methods: Using a de-identified electronic health record (EHR), we randomly chart reviewed 2,000 positive ANA individuals to determine if a systemic autoimmune disease was diagnosed by a rheumatologist. *A priori*, we considered demographics, billing codes for autoimmune disease-related symptoms, and laboratory values as variables for the risk model. We performed logistic regression and machine learning models using training and validation samples.

Results: We assembled training (n = 1030) and validation (n = 449) sets. Positive ANA individuals who were younger, female, had a higher titer ANA, higher platelet count, disease-specific autoantibodies, and more billing codes related to symptoms of autoimmune diseases were all more likely to develop autoimmune diseases. The most important variables included having a disease-specific autoantibody, number of billing codes for autoimmune disease-related symptoms, and platelet count. In the logistic regression model, AUC was 0.83 (95% CI 0.79-0.86) in the training set and 0.75 (95% CI 0.68-0.81) in the validation set.

Conclusion: We developed and validated a risk model that predicts risk for developing systemic autoimmune diseases and can be deployed easily within the EHR. The model can risk stratify positive ANA individuals to ensure high-risk individuals receive urgent rheumatology referrals while reassuring low-risk individuals and reducing unnecessary referrals.

KEYWORDS

antinuclear antibodies, electronic health record, risk model, autoimmune disease, rheumatology

1 Introduction

Positive antinuclear antibodies (ANAs) cause diagnostic dilemmas for clinicians across multiple specialties (1–3). Currently, no clinically available or validated tools exist to help clinicians determine the significance of a positive ANA. While a positive ANA serves as a diagnostic criterion for multiple autoimmune diseases, the test alone only has a 11% positive predictive value for systemic autoimmune disease (4). In US studies, rates of positive ANAs in the general population without autoimmune disease range from 14% to 27% (5, 6).

Frequent, inappropriate ordering of ANA testing has been recognized as a clinical problem by the American Board of Internal Medicine and the American College of Rheumatology in their "Choosing Wisely" campaign. Specifically, it is recommended to not order an ANA test unless specific symptoms for an autoimmune disease are present (7, 8). Up to 22% of all rheumatology referrals are for a positive ANA (1, 9). Only 11-20% of individuals with a positive ANA have an autoimmune disease diagnosed at referral (4, 10–13). Frequent ANA referrals in the setting of an international shortage of pediatric and adult rheumatologists (14–16) contribute to inefficient use of limited resources and lengthen wait times for rheumatology consultation (1, 9, 12).

Triage systems and electronic consultations have attempted to tackle the problem of frequent ANA referrals with limited success (12, 17–20). Risk models have been developed for systemic lupus erythematosus (SLE) (21, 22) but not for multiple systemic autoimmune diseases associated with a positive ANA. We aimed to develop and validate a robust risk model for use in the rheumatology clinic that uses readily available data in the electronic health record (EHR) to identify which individuals with a positive ANA are at high and low risk for developing systemic autoimmune disease.

2 Methods

2.1 Data source and patient selection

After receiving approval from the Vanderbilt University Medical Center (VUMC) IRB (#210189), we used the Synthetic Derivative, a de-identified version of the EHR that contains billing code and clinical data on over 3.6 million individuals spanning across three decades (23). Records from outside VUMC are not available.

We assembled all individuals within the Synthetic Derivative who had a positive ANA, defined as a titer ≥ 1:80 (Supplementary Figure 1). For ANA testing, the Hep-2 immunofluorescence assay was used for the entire study period (Appendix). We selected a random sample of 2,000 individuals with a positive ANA to perform chart review to assess for the model outcome and collect covariates. Model outcome was defined as developing a systemic autoimmune disease diagnosed by a rheumatologist, as EHR notes often lack systematic documentation of disease criteria (24). We performed chart review for development of systemic autoimmune disease from time of first positive ANA up to ten years later or individual's last EHR interaction. We allowed up to ten years, as individuals with autoimmune diseases can face significant diagnostic delays (25). Systemic autoimmune diseases are listed in Supplementary Table 1. In addition to diseases classically associated with a positive ANA (i.e., SLE, Sjogren's, systemic sclerosis, mixed connective tissue disease, and idiopathic inflammatory myopathies), we included other systemic autoimmune diseases such as rheumatoid arthritis (RA) and seronegative conditions (i.e., psoriatic arthritis, ankylosing spondylitis). Since the risk model will be used for triage to the rheumatology clinic, we aimed to include individuals with systemic autoimmune diseases who would be followed in that setting. While the ANA is not part of clinical criteria for these conditions, the ANA test is still frequently ordered in the evaluation of symptoms for these conditions (26). We excluded individuals with organ-specific autoimmune diseases such as autoimmune thyroiditis and autoimmune hepatitis, who would not be primarily managed by a rheumatologist. Individuals diagnosed outside of VUMC were included only if notes documented the individual was seen by an outside rheumatologist. For our primary analysis, we only analyzed individuals who were incident cases, defined as newly diagnosed with systemic autoimmune diseases at VUMC.

2.2 Model development

Based on clinical relevance and published SLE risk models (21, 22), prespecified predictors included demographics, laboratory values, and billing codes up to the time of first positive ANA

(Supplementary Table 2). Specifically, billing codes captured signs and symptoms for autoimmune diseases. A collection timeline for model covariates and outcome is detailed in Figure 1. Model outcome was developing a systemic autoimmune disease diagnosed by a rheumatologist within 10 years of first positive ANA (25).

Age was defined as age at first positive ANA documented at VUMC. The Synthetic Derivative defines race and ethnicity using a mixture of self-report and administrative entry with a fixed set of categories in accordance with NIH terminology. Studies have validated that these race and ethnicity assignments reflect self-report and genetic ancestry (27). For our primary analysis, race was initially excluded from the model as it was not significant in univariate analyses. Studies have shown that risk models that include race could potentially disadvantage high-risk groups from receiving appropriate care (28, 29). We performed a sensitivity analysis where race was included in the model, as studies demonstrate an increased risk of developing autoimmune disease in racial and ethnic underserved populations (1, 5).

We examined laboratory values one year prior to the date of the first positive ANA to allow for adequate data capture for individuals in the EHR and up to one month after to ensure capture of send-out studies such as the myositis antibody panel. We included autoantibodies associated with multiple autoimmune diseases (Supplementary Table 3). Autoantibodies were measured via enzyme-linked immunosorbent assays with manufacturer values to determine positivity (Appendix). We selected white blood cell count, platelet count, and serum creatinine as leukopenia, thrombocytopenia, and elevated serum creatinine have all been associated with autoimmune diseases (22, 30, 31). In SLE risk models (21, 22) and studies assessing presence of autoimmune diseases in positive ANA individuals (30, 31), leukopenia and thrombocytopenia were important predictors. Therefore, when examining multiple laboratory values for an individual, we selected the lowest white blood cell and platelet counts within the study period. For serum creatinine, we used the highest value within the study period to simulate how a rheumatologist might review lab trends. These values were treated as continuous variables. For missing laboratory values, we used median value imputation, as this method has been shown to be comparable to multiple imputation and is more feasible in real-time predictive models (32). We included ANA titer, as higher ANA titers are associated with risk of developing autoimmune disease (9, 30). Reporting of ANA titers are detailed in the Appendix. Briefly, ANA titer was dichotomized to 1:80 and ≥ 1:160 categories due to limited reporting of titers in some of the historical data. While different ANA patterns may have associations with different systemic autoimmune diseases (33), we did not include ANA pattern. ANA patterns are not reported in a standardized fashion at our institution according to the International Consensus on ANA patterns (33). Multiple or inconsistent patterns are often reported, particularly in the setting of changing technology over the study period. Further, as pattern is reported as a text variable, extraction from the EHR in real-time to input into the risk model would be challenging.

We used both ICD-9 and ICD-10-CM billing codes to capture signs and symptoms for systemic autoimmune diseases (Supplementary Table 4). These codes were significant in a UK SLE risk model (21) and were expanded upon to ensure capture of signs and symptoms for multiple autoimmune diseases in addition to SLE. Similar to the UK model, we searched for billing codes up to five years prior to the date of first positive ANA (21). In model development, we had an insufficient sample size to fit a model with a unique predictor for each billing code, so we created a single aggregated variable (Supplementary Table 5).

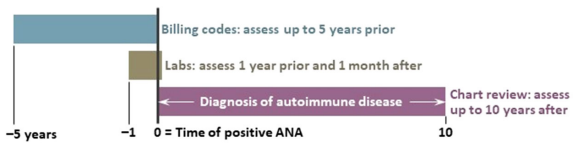


FIGURE 1
Timeline of model covariates. We assessed billing codes up to 5 years prior to the first positive antinuclear antibody (ANA) test. Laboratory values were assessed up to 1 year and 1 month after the ANA test. We conducted chart review for the model outcome of developing a systemic autoimmune disease diagnosed by a rheumatologist up to 10 years after the first positive ANA test.

2.3 Statistical analysis

We derived separate training and validation sets using 2,000 positive ANA individuals. We estimated that 10-15% of our 2,000 positive ANA individuals would have an incident autoimmune disease (4, 10–13), leading to 200-300 cases for the training and validation sets combined. To prevent overfitting and applying the rule of 10-15 outcomes per one degree of freedom (34), we fit a logistic regression model with 13 degrees of freedom. Prespecified variables are shown in Supplementary Table 2. Total number of visits, white blood cell count, and serum creatinine were collinear with included model variables and were removed from the final model. We performed logistic regression using the following predictors: age at time of first positive ANA, sex, ANA titer, platelet count, and billing codes. Final model formula is in Supplementary Figure 2. We also performed machine learning methods including extreme gradient boosting (XGB) (35–37) and neural networks. Hyperparameters are in the Appendix. We assessed model performance in the training and validation sets using c-statistic, Brier score, and calibration curves.

2.4 Model validation

We conducted an internal validation of the logistic regression model using a bootstrap with 200 replications (38, 39). The bootstrap validation can test the stability of a model across different samples. In addition, a random selection of individuals, separate from the training set, was set aside as a “hold-out” for model validation (Supplementary Figure 1). Specifically, we estimated needing 100-200 incident autoimmune disease cases to avoid overfitting our model. To achieve this sample, we used 1384

individuals of which 1030 incident individuals were used for analysis, resulting in 152 incident cases. We then used the remainder of the original 2,000 set for a validation set with 616 individuals, of which 449 incident individuals were used for analysis, resulting in 74 incident cases.

2.5 Sensitivity analyses and deployment feasibility assessment

For our primary analysis, we excluded subjects with "unclear" autoimmune diagnoses. In a sensitivity analysis, we treated "unclear" subjects as not cases. We also included a sensitivity analysis where race was included with categories of White, Black, and Other. To account for longitudinal and censored data, we conducted a Cox proportional-hazard model using the same variables as the logistic regression model. Outcome was time from first positive ANA to either autoimmune disease diagnosis or last EHR follow-up (Appendix). We initially dichotomized ANA titer to 1:80 and ≥ 1:160 categories due to historical reporting in some of our data (Appendix). We then conducted a sensitivity analysis using more recent data (2017-2021) that incorporated multiple categories for the ANA titer (1:80, 1:160, 1:320, 1:640, 1:1280, and ≥ 1:2560). We also conducted sensitivity analyses where seronegative conditions were not counted as a case (Appendix).

We applied our logistic regression model to data extracted from our EHR-provided data warehouse (Epic Clarity) to assess feasibility of deploying the model in real-time. We calculated risk probabilities for systemic autoimmune disease for individuals with a positive ANA from 2017-2021. This time period captured the updated ANA titer reporting to the most current data available at time of analysis.

3 Results

3.1 Individual characteristics

Training (n = 1030) and validation (n = 449) sets are compared in Table 1 with individuals having similar characteristics. In the training set, 15% (n = 152) of individuals with a positive ANA developed a systemic autoimmune disease. Individuals with systemic autoimmune diseases were younger (41.8 ± 21.5 vs. 47.9 ± 19.3 years, p = 0.003), more likely to be female (84% vs. 70%, p <

TABLE 1 Characteristics of incident positive ANA individuals in training and validation sets.

Characteristics	Training set n = 1030	Validation set n = 449	*p* value*
Autoimmune disease % (n)	15% (152)	16% (74)	0.40
Age at positive ANA, years mean ± SD	47.0 ± 19.8	48.0 ± 20.3	0.44
Race % (n)† White	85% (807)	85% (355)	0.88
African American	12% (113)	12% (50)	
Asian	2% (16)	1% (5)	
Other	1% (11)	1% (5)	
Ethnicity† Hispanic	3% (32)	3% (11)	0.46
Not Hispanic or Latino/a	97% (889)	97% (397)	
Sex			
Female	72% (739)	74% (333)	0.34
ANA titer‡			
1:80	20% (202)	19% (87)	0.92
≥ 1:160	80% (828)	81% (362)	
White blood cell count† K/uL, Mean ± SD	6.9 ± 3.4	6.9 ± 2.9	0.88
Platelet count† K/uL, Mean ± SD	235 ± 100	233 ± 92	0.58
Serum creatinine† mg/dL, Mean ± SD	1.1 ± 0.9	1.2 ± 1.4	0.25
Ever present autoantibody§ **% (n)**	15% (155)	15% (68)	0.96

(Continued)

TABLE 1 Continued

Characteristics	Training set n = 1030	Validation set n = 449	*p* value*
Total any billing codes mean ± SD	30 ± 60	37 ± 71	0.27
Count of specific billing codes‖ mean ± SD	0.7 ± 0.8	0.8 ± 0.9	0.01
Alopecia % (n)	2% (21)	1% (6)	0.35
Arthritis	26% (264)	31% (140)	0.03
Fatigue	20% (207)	23% (104)	0.18
Interstitial Lung Disease	1% (14)	2% (11)	0.14
Pulmonary Hypertension	1% (11)	1% (6)	0.66
Rash	9% (97)	9% (42)	0.97
Raynaud's	2% (19)	3% (12)	0.31
Serositis	4% (40)	5% (23)	0.28
Sicca	0.3% (3)	1% (5)	0.05

*Mann-Whitney U test for continuous variables and chi-square test for categorical variables. P values calculated with excluding missing observations.
†Race, ethnicity, and lab values have missing data with 81 (8%) for race, 109 (11%) for ethnicity, 201 (20%) for white blood cell count, 211 (20%) for platelet count, and 210 (20%) for serum creatine in the training set. In the validation set, 32 (7%) for race, 41 (9%) for ethnicity, 91 (20%) for white blood cell count, 95 (21%) for platelet count, and 100 (22%) for serum creatine.
‡For ANA titer, up until July 1, 2016, titers were reported as 1:40 (negative), 1:80, and ≥ 1:160. After this date, titers were then reported as 1:40 (negative), 1:80, 1:160, 1:320, 1:640, 1:1280, and 1:2560.
§Presence of other autoantibodies included rheumatoid factor, cyclic citrullinated peptide, SSA (Ro), SSB (La), scl-70, centromere, RNP, Smith, dsDNA, ANCA, Jo-1, or any antibody from the myositis antibody panel.
‖See Supplementary Table 4 for full list of ICD-9 and ICD-10-CM billing codes and Supplementary Table 5 for details on scoring. For each individual, we counted if any billing code was ever present (1 for present, 0 for absent) for each of the nine categories (i.e., arthritis, fatigue) and then summed this up across the nine prespecified billing code categories for a maximum score of nine.

0.001), have a higher ANA titer (≥1:160 vs. 1:80) (90% vs. 79%, p = 0.002), lower serum creatinine (0.9 ± 0.6 vs. 1.2 ± 1.0 mg/dL, p < 0.001), higher platelet count (274 ± 113 vs. 229 ± 96 K/uL, p < 0.001), and a disease-specific autoantibody (51% vs. 9%, p < 0.001) (Table 2). No significant differences were found in race, ethnicity, or white blood cell count in individuals with vs. without systemic autoimmune diseases. Individuals with systemic autoimmune disease had a higher count of the nine billing code categories (scale 0 to 9) compared to individuals without disease (0.9 ± 0.9 vs. 0.6 ± 0.8, p < 0.001). Individuals with systemic autoimmune

TABLE 2 Characteristics of positive ANA individuals with vs. without systemic autoimmune disease in the training set.

Characteristics	No systemic autoimmune disease n = 878	Systemic autoimmune disease n = 152	Proportion with systemic autoimmune disease*	*p* value†
Age at positive ANA, years, mean ± SD	47.9 ± 19.3	41.8 ± 21.5	..	0.003
Race % (n)‡ White	85% (680)	85% (127)	16%	0.26
African American	12% (94)	13% (19)	17%	
Asian	2% (16)	0% (0)	0%	
Native American	0.1% (1)	1% (1)	50%	
Other	1% (10)	1% (1)	9%	
Ethnicity‡ Hispanic	4% (30)	1% (2)	6%	0.13
Not Hispanic or Latino/a	96% (744)	99% (145)	16%	
Sex Female	70% (612)	84% (127)	17%	< 0.001
Male	30% (266)	16% (25)	9%	

(Continued)

TABLE 2 Continued

Characteristics	No systemic autoimmune disease n = 878	Systemic autoimmune disease n = 152	Proportion with systemic autoimmune disease*	*p* value†
ANA titer§ 1:80	21% (186)	11% (16)	8%	0.002
≥ 1:160	79% (692)	90% (136)	16%	
White blood cell count‡				
K/uL, mean ± SD	6.9 ± 3.4	7.1 ± 3.2	..	0.49
Platelet count‡ K/uL, mean ± SD	229 ± 96	274 ± 113	..	<0.001
Serum creatinine‡ mg/dL, mean ± SD	1.2 ± 1.0	0.9 ± 0.6	..	<0.001
Ever present autoantibody‖				
No	91% (800)	49% (75)	9%	<0.001
Yes	9% (78)	51% (77)	50%	
Total any billing codes, mean ± SD	32 ± 62	23 ± 43	..	0.02
Count of specific billing codes,¶ mean ± SD	0.6 ± 0.8	0.9 ± 0.9	..	< 0.001
Alopecia	2% (16)	3% (5)	24%	0.24
Arthritis	23% (203)	40% (61)	23%	< 0.001
Fatigue	19% (169)	25% (38)	18%	0.10
Interstitial Lung Disease	2% (13)	1% (1)	7%	0.42
Pulmonary Hypertension	1% (9)	1% (2)	18%	0.26
Rash	9% (81)	11% (16)	17%	0.61
Raynaud's	1% (12)	5% (7)	37%	0.006
Serositis	4% (34)	4% (6)	15%	0.97
Sicca	0.3% (3)	0% (0)	0%	0.47

*Overall percentage of individuals with systemic autoimmune disease is 14.8%. P values calculated with excluding missing observations.
†Mann-Whitney U test for continuous variables and chi-square test for categorical variables.
‡Race, ethnicity, and lab values have missing data with 81 (8%) for race, 109 (11%) for ethnicity, 201 (20%) for white blood cell count, 211 (20%) for platelet count, and 210 (20%) for serum creatine.
§For ANA titer, up until July 1, 2016, titers were reported as 1:40 (negative), 1:80, and ≥ 1:160. After this date, titers were then reported as 1:40 (negative), 1:80, 1:160, 1:320, 1:640, 1:1280, and 1:2560.
‖Presence of other autoantibodies included rheumatoid factor, cyclic citrullinated peptide, SSA (Ro), SSB (La), scl-70, centromere, RNP, Smith, dsDNA, ANCA, Jo-1, or any antibody from the myositis antibody panel.
¶See Supplementary Table 4 for full list of ICD-9 and ICD-10-CM billing codes and Supplementary Table 5 for details on scoring. For each individual, we counted if any billing code was ever present (1 for present, 0 for absent) for each of the nine categories (i.e., arthritis, fatigue) and then summed this up across the nine prespecified billing code categories for a maximum score of nine.

disease were more likely to have billing codes for arthritis (40% vs. 23%, p < 0.001) and Raynaud's phenomenon (5% vs. 1%, p = 0.006) but not the other seven code categories.

Of the 152 individuals with systemic autoimmune diseases, the most frequent diagnoses were SLE at 18% (n = 28) followed by other at 16% (n = 24), undifferentiated connective tissue disease at 16% (n = 24), and RA at 15% (n = 22) (Supplementary Table 6). Other consisted of psoriatic arthritis, unspecified inflammatory arthritis, and inflammatory bowel disease (Supplementary Table 6). Individuals with unclear diagnoses of systemic autoimmune disease (n = 66) were excluded from the primary analysis but are described in Supplementary Table 7. For individuals without systemic autoimmune diseases, when available alternative diagnoses were documented by rheumatologists, the most frequent diagnoses were fibromyalgia (n = 18), osteoarthritis (n = 11), and gout (n = 6) (Supplementary Table 8).

3.2 Model description and validation

The final model included age at first positive ANA, sex, ANA titer, presence of another autoantibody, platelet count, and billing code category count. Age was fit with a three-knot restricted cubic spline and interacted with sex and was prespecified based on prior

literature (21). Our data demonstrated a higher probability of systemic autoimmune disease in female vs. male individuals at younger ages but a similar probability at older ages (Supplementary Figure 3). The most important variables in the model were presence of another autoantibody (i.e., dsDNA), billing code category count, and platelet count (Figure 2). Model AUC was 0.83 (95% CI 0.79-0.86) (Figure 3A) with a Brier score of 0.10 and calibration shown in Figure 3B. XGBoost resulted in an AUC of 0.94 (95% CI 0.91-0.95) and neural networks with an AUC of 0.83 (95% CI 0.79-0.87).

Based on the internal bootstrap validation, the logistic regression model was stable and robust (Appendix). For the validation set (n = 449), 16% of individuals had systemic autoimmune disease (Supplementary Table 9). For the logistic regression model, AUC was 0.75 (95% CI 0.68-0.81) (Figure 3C) with a Brier score of 0.12 with calibration shown in Figure 3D. XGBoost resulted in an AUC of 0.72 (95% CI 0.65-0.78) and neural networks with an AUC of 0.74 (95% CI 0.68-0.81).

3.3 Sensitivity analyses

Race was included in the model with categories of White, Black, and Other resulting in an AUC of 0.83 (95% CI 0.79-0.87). When individuals of unclear case status for systemic autoimmune disease were counted as non-cases, model AUC was 0.80 (95% CI 0.76-0.83). When these unclear individuals were counted as cases, model AUC was 0.74 (95% CI 0.71-0.77). The distribution of model risk scores for these unclear individuals most closely matched individuals who were not cases (Supplementary Figure 4). For the Cox model with the outcome time to autoimmune diagnosis, model predictors behaved similarly to the logistic regression model (Supplementary Figure 5).

To reflect more updated ANA titer reporting, we used a cohort of individuals with a positive ANA from 2017 to 2021 (n = 584) (Appendix) to perform additional sensitivity analyses. For the 2017-2021 cohort, there was a significant difference in the distribution of ANA titers between cases and non-cases ($p < 0.001$). Of the cases, 40% had an ANA titer greater than 1:640, while 18% of non-cases had a titer greater than 1:640 (Supplementary Table 10). In this cohort, using a dichotomized ANA titer (1:80 vs. ≥1:160), model AUC was 0.85 (95% CI 0.81 – 0.90). For the model with full ANA titer reporting (i.e., 1:80, 1:160, 1:320, 1:640, 1:1280, ≥ 1:2560), model AUC was 0.89 (95% CI 0.84 – 0.92). Lastly, we assessed if a higher ANA titer cutoff would impact model performance using the above 2017-2021 cohort. We fit a model using an ANA cutoff at 1:160, which had an AUC of 0.83 (95% CI 0.78-0.87), identical to the performance of the model using the original ANA cutoff at 1:80 (AUC of 0.83 (95% CI 0.78-0.87)).

For using an alternative case definition for systemic autoimmune disease that did not count seronegative conditions (i.e., psoriatic arthritis, ankylosing spondylitis) as cases, model AUC was 0.86 (95% CI 0.83-0.89).

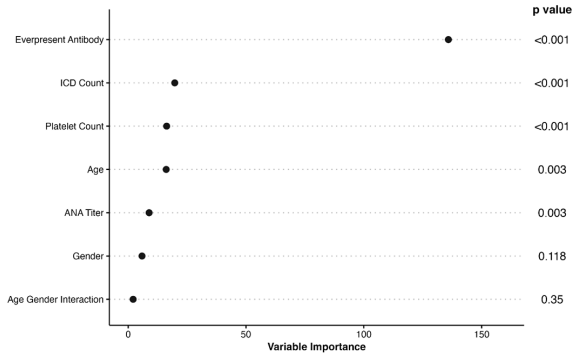


FIGURE 2
Importance of Variables in ANA Risk Model. The list of variables in the final ANA risk model are shown to the left with p values to the right. The x axis shows variable importance using a Wald statistic. Ever-present antibody refers to having a disease-specific autoantibody such as a rheumatoid factor or dsDNA. ICD count refers to billing code category count that ranges from 0 to 9.

3.4 Distribution of risk scores by type of autoimmune disease

We examined the distribution of model risk scores by type of autoimmune disease (Supplementary Figure 6). Individuals with SLE had the highest risk scores with a median of 0.481 and IQR of 0.312-0.685 followed by RA with 0.423 (0.144-0.582). Individuals labeled as other, with predominantly seronegative conditions, had the lowest median risk score of 0.107 (0.061-0.269). Seronegative conditions included psoriatic arthritis, and inflammatory bowel disease. Individuals with seropositive diseases had a higher median risk score compared to individuals with seronegative diseases (0.385 vs. 0.107, difference in medians = 0.278, 95% CI 0.195 – 0.332, $p < 0.001$).

3.5 Deployment feasibility

We assessed the feasibility of implementing the logistic regression risk model in our Epic EHR using data for all individuals with a positive ANA from 2017-2021 (n = 22,234). We observed a similar distribution of risk scores in Epic compared to our training set that used a de-identified EHR database (Synthetic Derivative) (Supplementary Figure 7). A demonstration of how the risk model works can be accessed at https://cqs.app.vumc.org/shiny/AutoimmuneDiseasePrediction/ (Figure 4). A disclaimer is included that the application is not intended for clinical practice.

4 Discussion

We developed and validated a risk model that predicts risk for developing systemic autoimmune disease in individuals with a positive ANA. The model is important because it utilizes readily available clinical data in the EHR, can be deployed easily within clinical practice, and helps risk stratify individuals with a positive ANA, a source of frequent rheumatology referrals. Our risk model

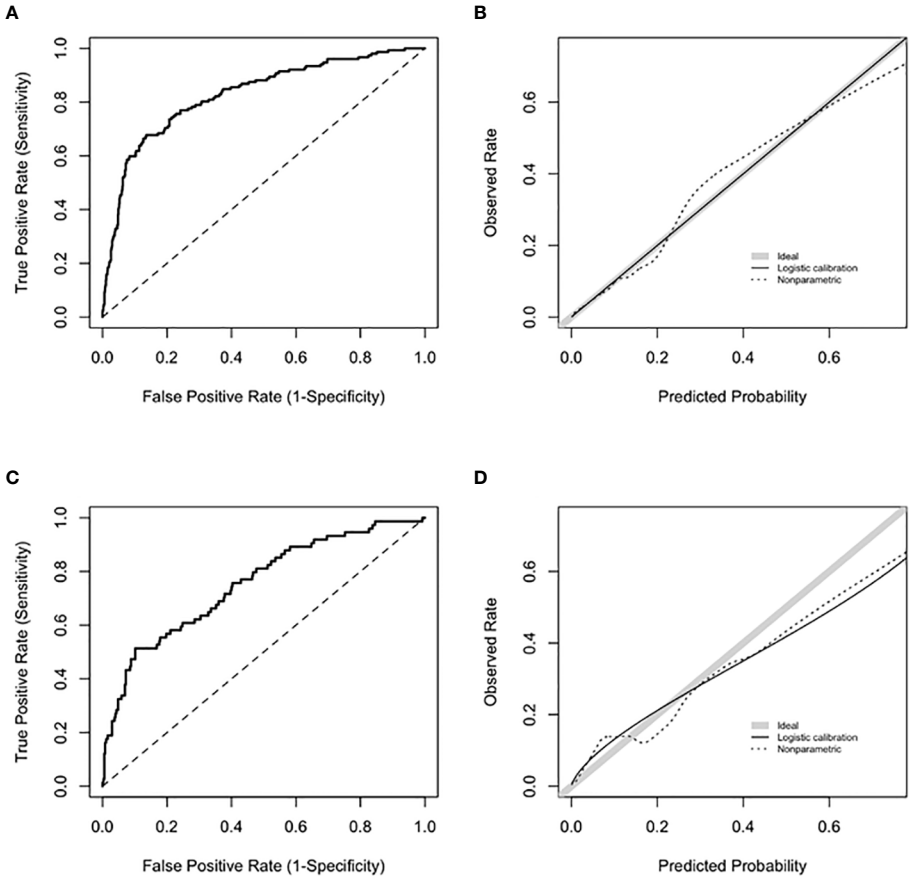


FIGURE 3
Model performance for training and validation sets. **(A)** shows ROC for the training set with an AUC 0.83 (95% CI 0.79-0.86). **(B)** shows calibration curve with a slope of 1 and intercept of 0 for the training set. Slopes that approach 1, as shown by the shaded grey line, demonstrate ideal calibration, agreement between predicted risk for systemic autoimmune disease and observed rate. **(C)** shows ROC for the validation set with an AUC 0.75 (95% CI 0.68-0.81). **(D)** shows calibration curve for the validation set. Calibration slope was equal to 0.71 and intercept was equal to 0.08.

identifies high-risk individuals, who are most likely to develop a systemic autoimmune disease, to ensure they are seen urgently for prompt diagnosis and treatment. Our risk model also identifies low-risk individuals who could be reassured, reducing unnecessary rheumatology referrals.

To our best knowledge, a risk model that focuses on individuals with a positive ANA and predicts risk for multiple systemic autoimmune diseases does not currently exist. One SLE risk model used UK EHR data (21) but did not focus on positive ANA individuals or examine risk for other autoimmune diseases.

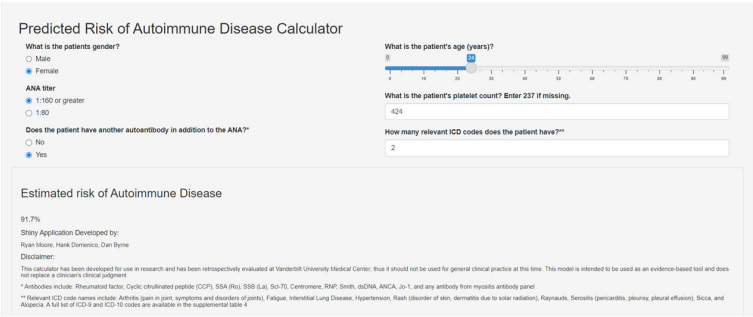


FIGURE 4
Screenshot of Shiny app for risk model for systemic autoimmune disease. The screenshot shows the risk model covariates used to estimate risk for systemic autoimmune disease. This app demonstrates how the risk score is calculated and is not intended for clinical practice. The Shiny app can be accessed at the following link: https://cqs.app.vumc.org/shiny/AutoimmuneDiseasePrediction/.

In this model, billing codes such as arthritis, rash, sicca, and fatigue were most significantly associated with risk of developing SLE along with female sex, younger age, and a higher number of clinic visits. We found similar results in our model and used similar billing codes but expanded our codes to identify not just SLE but also other systemic autoimmune diseases. Similar to the UK SLE model, we used a non-linear age and an age-sex interaction term. Despite its strengths, the UK SLE model had limited performance with a positive predictive value of 7-9%, a sensitivity of 24-34%, and an AUC of 0.75. Further, this model was not deployed in the EHR. Our model attained a higher AUC of 0.83 and can be easily deployed in real-time in the EHR.

Another SLE risk model from a Greek center (22) used random forests and Lasso-LR models. Not surprisingly, clinical items from the ACR SLE classification criteria accurately identified SLE cases with a high model AUC. While this study had a relatively large sample and a validation set, the model was developed using rheumatology clinic individuals and not in a general practice setting where there is often diagnostic dilemma. This model would be challenging to deploy in the EHR as it relies on SLE diagnostic criteria that may not be documented systematically, even in rheumatology notes (24).

The most important variable in our model was having another autoantibody in addition to the positive ANA, which is more specific for autoimmune diseases (1–3). Individuals with disease-specific autoantibodies may have a higher pretest probability for autoimmune disease by simply having these tests ordered. We tried to mitigate this bias by only including incident positive ANA individuals without established diagnoses of systemic autoimmune disease. Further, our institution conducts reflex testing where disease-specific autoantibodies are sent if an ANA is positive. Disease-specific autoantibodies may not be available fully in real-time at centers that do not perform reflex testing with a positive ANA, which may impact the performance of the model. The next most important variable was count of the nine prespecified billing code categories. *A priori*, we selected billing codes that captured signs and symptoms for autoimmune diseases and were significant in the UK SLE risk model (21). As expected, a higher count of these billing codes was predictive for systemic autoimmune disease. While billing codes may not always adequately capture an individual's symptoms, ICD billing codes allow for automation of the risk model in real-time and allow for portability of the model to other EHRs and databases that use common data models. Platelet count was also an important variable in our model. We originally hypothesized that a lower platelet count would be associated with systemic autoimmune disease. Prior SLE risk models identified thrombocytopenia as an important model predictor (21, 22), and other studies demonstrated an association of thrombocytopenia with autoimmune disease in positive ANA individuals (30, 31). Instead, we found a higher value of an individual's lowest platelet count was associated with systemic autoimmune disease. Higher platelet counts have been observed in individuals with RA and correlate with increased disease activity (40) and may also signal inflammation (41). *A priori*, we elected to not include inflammatory markers such as sedimentation rate (ESR) and C-reactive protein (CRP), as we had significant missingness of these values in the EHR. Further, these markers are nonspecific and can fluctuate widely in an individual (42–44). Elevations in these markers can be unrelated to an underlying systemic autoimmune disease, for example, in the setting of infection and malignancy (42–45).

A priori, we included race and ethnicity in our risk model. African American and Hispanic individuals have higher frequencies of positive ANAs compared to White individuals and are at higher risk of developing autoimmune disease, particularly SLE (1, 5). In univariate analysis, neither race nor ethnicity were significantly associated with systemic autoimmune disease, so race and ethnicity were not initially included. Studies have shown that risk models that include race could potentially disadvantage high-risk groups from receiving appropriate care (28, 29). For our model, this could include Black individuals. In a sensitivity analysis, we included race and found a similar model AUC of 0.83.

Our logistic regression model demonstrated robustness in both an internal bootstrap validation and a separate validation set. A successful bootstrap validation demonstrates the model can hold up when it encounters different samples. With predicting a clinically complex outcome where no current tools or risk models exist, our model validation demonstrated an improvement over usual care. To assess alternative approaches, we developed models using XGBoost and neural networks. XGBoost had a higher apparent AUC compared to the training set logistic regression model, likely due to overfitting, but did not hold up in validation. Neural networks performed similarly to the logistic regression model but with added complexity that would limit interpretability and deployment in the EHR.

While we developed, validated, and deployed a robust risk model to predict risk of systemic autoimmune disease in positive ANA individuals, our study has limitations. Our model was developed at a single academic medical center with more complex patients being evaluated, so may not generalize to other practice settings. Further, our study population was predominantly White, so it may not generalize to individuals with different race and ethnicity backgrounds and in other geographic areas. Our data encompasses an almost 30-year study period that included changes in ANA titer reporting. As a result, our primary analysis for the risk model included dichotomized reporting of the ANA titer to capture historical data. Sensitivity analyses using a more recent cohort of positive ANA individuals using both the dichotomized and full reporting of the ANA titer had similar model AUCs with overlapping confidence intervals. For future versions of the risk model, full reporting of the ANA titer can be used. We purposely defined systemic autoimmune disease based on a rheumatologist's diagnosis instead of classification criteria, as classification criteria are not systematically documented in clinical notes (24). Case definition by a rheumatologist could contribute to heterogeneity of cases (i.e. calling an individual with mild SLE and SLE nephritis both SLE).

Interestingly, our model did not perform as well in individuals with seronegative conditions not typified by autoantibodies, as

presence of these autoantibodies was the strongest predictor in our model. This limitation should be considered when interpreting risk scores. Seronegative conditions encompass overlapping diseases including plaque psoriasis, psoriatic arthritis, and inflammatory bowel diseases. These conditions have different HLA-based risk alleles, disease mechanisms, and disease presentations compared to seropositive conditions (46). While these seronegative conditions are not classically associated with a positive ANA, individuals with these conditions can have higher rates of ANA positivity compared to the general population (47–49) and often have an ANA test ordered as part of their clinical evaluation (26). In a sensitivity analysis, not counting the individuals with seronegative conditions as cases did not greatly impact the performance of the model.

Our model achieved a robust AUC of 0.83, but it does not discriminate perfectly between individuals with and without systemic autoimmune diseases. We found this AUC to be an improvement over usual care, where no current risk models exist to help risk stratify positive ANA individuals. The risk model was not designed to diagnose systemic autoimmune disease but to serve as a tool to identify positive ANA individuals who are at risk of developing systemic autoimmune disease within the next 10 years. The risk model can complement the clinician's judgment as well as the patient history and physical exam. The risk model could also assist the ordering physician in identifying individuals at lower risk that may not need rheumatology referral. This reassurance may reduce unnecessary referrals and expenses to the healthcare system. We purposefully created a continuous risk score, which is more rigorous than commonly used dichotomous or "cut-off" scores. Without a "cut-off score," we cannot currently estimate a positive predictive value. We are currently conducting a prospective validation of the risk model in real-time in the EHR to inform which individuals are low vs. high risk. While we created an application to demonstrate how the model incorporates variables and calculates a risk score, this application is not intended to be used in clinical practice yet or identify individuals as low vs. high risk.

In summary, we developed, validated, and deployed a risk model to identify which positive ANA individuals will develop systemic autoimmune disease. This risk model can be automated and deployed in real-time with no input needed from a clinician. In the setting of an international shortage of rheumatologists (14–16), a risk-stratifying tool for positive ANA individuals is critical. For future directions, we are assessing our risk model in real-time in the EHR prospectively and its impact on time to diagnosis and treatment for autoimmune diseases. Pending prospective validation, we envision our risk model would predict risk of autoimmune diseases within 10 years of a positive ANA similar to the FRAX that predicts 10-year fracture risk (50) or the ASCVD risk algorithm that predicts 10-year cardiovascular event risk (51). Risk scores from our model could then directly inform management of individuals with positive ANAs. High-risk individuals could be seen urgently by rheumatologists to ensure prompt diagnosis and treatment, and low-risk individuals could be reassured, reducing unnecessary rheumatology referrals.

Data availability statement

Raw data and R code used in analyses will be available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Vanderbilt University Medical Center. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

AB: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. RM: Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. HD: Conceptualization, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. SG: Data curation, Investigation, Project administration, Supervision, Writing – original draft, Writing – review & editing. AC: Data curation, Investigation, Writing – original draft, Writing – review & editing. AS: Data curation, Investigation, Project administration, Supervision, Writing – original draft, Writing – review & editing. BH: Data curation, Investigation, Writing – original draft, Writing – review & editing. KW: Data curation, Investigation, Writing – original draft, Writing – review & editing. AA: Data curation, Investigation, Writing – original draft, Writing – review & editing. LC: Data curation, Investigation, Writing – original draft, Writing – review & editing. AK: Data curation, Investigation, Writing – original draft, Writing – review & editing. AM: Investigation, Methodology, Project administration, Resources, Software, Writing – original draft, Writing – review & editing. DB: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Institutes of Health/National Institute of Arthritis and Musculoskeletal and Skin Diseases (1K08 AR072757-01, R01 AR080629, Barnado), National Institutes of Health/National Center for Research Resources (UL1 RR024975, VUMC), National Institutes of Health/National Center

for Advancing Translational Sciences (ULTR000445, VUMC), Vanderbilt University Medical Center Department of Biomedical Informatics Catalyzing Informatics Innovation Program.

Acknowledgments

The authors would like to thank Leslie J. Crofford, MD for review of the manuscript.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1384229/full#supplementary-material

References

1. Olsen NJ, Karp DR. Finding lupus in the ANA haystack. *Lupus Sci Med*. (2020) 7: e000384. doi: 10.1136/lupus-2020-000384

2. Pisetsky DS. Antinuclear antibody testing - misunderstood or misbegotten? *Nat Rev Rheumatol*. (2017) 13:495–502. doi: 10.1038/nrrheum.2017.74

3. Olsen NJ, Choi MY, Fritzler MJ. Emerging technologies in autoantibody testing for rheumatic diseases. *Arthritis Res Ther*. (2017) 19:172. doi: 10.1186/s13075-017-1380-3

4. Slater CA, Davis RB, Shmerling RH. Antinuclear antibody testing. A study of clinical utility. *Arch Intern Med*. (1996) 156:1421–5. doi: 10.1001/archinte.1996.00440120079007

5. Wandstrat AE, Carr-Johnson F, Branch V, Gray H, Fairhurst AM, Reimold A, et al. Autoantibody profiling to identify individuals at risk for systemic lupus erythematosus. *J Autoimmun*. (2006) 27:153–60. doi: 10.1016/j.jaut.2006.09.001

6. Satoh M, Chan EK, Ho LA, Rose KM, Parks CG, Cohn RD, et al. Prevalence and sociodemographic correlates of antinuclear antibodies in the United States. *Arthritis Rheum*. (2012) 64:2319–27. doi: 10.1002/art.34380

7. Qaseem A, Alguire P, Dallas P, Feinberg LE, Fitzgerald FT, Horwitch C, et al. Appropriate use of screening and diagnostic tests to foster high-value, cost-conscious care. *Ann Intern Med*. (2012) 156:147–9. doi: 10.7326/0003-4819-156-2-201201170-00011

8. Yazdany J, Schmajuk G, Robbins M, Daikh D, Beall A, Yelin E, et al. Choosing wisely: the American College of Rheumatology's Top 5 list of things physicians and patients should question. *Arthritis Care Res (Hoboken)*. (2013) 65:329–39. doi: 10.1002/acr.21930

9. McGhee JL, Kickingbird LM, Jarvis JN. Clinical utility of antinuclear antibody tests in children. *BMC Pediatr*. (2004) 4:13. doi: 10.1186/1471-2431-4-13

10. Dinser R, Braun A, Jendro MC, Engel A. Increased titres of anti-nuclear antibodies do not predict the development of associated disease in the absence of initial suggestive signs and symptoms. *Scand J Rheumatol*. (2007) 36:448–51. doi: 10.1080/03009740701406577

11. Soto ME, Hernandez-Becerril N, Perez-Chiney AC, Hernandez-Rizo A, Telich-Tarriba JE, Juarez-Orozco LE, et al. Predictive value of antinuclear antibodies in autoimmune diseases classified by clinical criteria: Analytical study in a specialized health institute, one year follow-up. *Results Immunol*. (2015) 5:13–22. doi: 10.1016/j.rinim.2013.10.003

12. Patel V, Stewart D, Horstman MJ. E-consults: an effective way to decrease clinic wait times in rheumatology. *BMC Rheumatol*. (2020) 4:54. doi: 10.1186/s41927-020-00152-5

13. Abeles AM, Abeles M. The clinical utility of a positive antinuclear antibody test result. *Am J Med*. (2013) 126:342–8. doi: 10.1016/j.amjmed.2012.09.014

14. Correll CK, Ditmyer MM, Mehta J, Imundo LF, Klein-Gitelman MS, Monrad SU, et al. 2015 american college of rheumatology workforce study and demand projections of pediatric rheumatology workforce, 2015-2030. *Arthritis Care Res (Hoboken)*. (2022) 74:340–8. doi: 10.1002/acr.24497

15. Battafarano DF, Ditmyer M, Bolster MB, Fitzgerald JD, Deal C, Bass AR, et al. 2015 american college of rheumatology workforce study: supply and demand projections of adult rheumatology workforce, 2015-2030. *Arthritis Care Res (Hoboken)*. (2018) 70:617–26. doi: 10.1002/acr.23518

16. Miloslavsky EM, Marston B. The challenge of addressing the rheumatology workforce shortage. *J Rheumatol*. (2022) 49:555–7. doi: 10.3899/jrheum.220300

17. Speed CA, Crisp AJ. Referrals to hospital-based rheumatology and orthopaedic services: seeking direction. *Rheumatol (Oxford)*. (2005) 44:469–71. doi: 10.1093/rheumatology/keh504

18. Rostom K, Smith CD, Liddy C, Afkham A, Keely E. Improving access to rheumatologists: use and benefits of an electronic consultation service. *J Rheumatol*. (2018) 45:137–40. doi: 10.3899/jrheum.161529

19. Vimalananda VG, Gupte G, Seraj SM, Orlander J, Berlowitz D, Fincke BG, et al. Electronic consultations (e-consults) to improve access to specialty care: a systematic review and narrative synthesis. *J Telemed Telecare*. (2015) 21:323–30. doi: 10.1177/1357633X15582108

20. Saxon DR, Kaboli PJ, Haraldsson B, Wilson C, Ohl M, Augustine MR. Growth of electronic consultations in the Veterans Health Administration. *Am J Manag Care*. (2021) 27:12–9. doi: 10.37765/ajmc.2021.88572

21. Rees F, Doherty M, Lanyon P, Davenport G, Riley RD, Zhang W, et al. Early clinical features in systemic lupus erythematosus: can they be used to achieve earlier diagnosis? A risk prediction model. *Arthritis Care Res (Hoboken)*. (2017) 69:833–41. doi: 10.1002/acr.23021

22. Adamichou C, Genitsaridi I, Nikolopoulos D, Nikoloudaki M, Repa A, Bortoluzzi A, et al. Lupus or not? SLE Risk Probability Index (SLERPI): a simple, clinician-friendly machine learning-based model to assist the diagnosis of systemic lupus erythematosus. *Ann Rheum Dis*. (2021) 80:758–66. doi: 10.1136/annrheumdis-2020-219069

23. Roden DM, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin Pharmacol Ther*. (2008) 84:362–9. doi: 10.1038/clpt.2008.89

24. Barnado A, Casey C, Carroll RJ, Wheless L, Denny JC, Crofford LJ. Developing electronic health record algorithms that accurately identify patients with systemic lupus erythematosus. *Arthritis Care Res (Hoboken)*. (2017) 69:687–93. doi: 10.1002/acr.22989

25. Sloan M, Harwood R, Sutton S, D'Cruz D, Howard P, Wincup C, et al. Medically explained symptoms: a mixed methods study of diagnostic, symptom and support experiences of patients with lupus and related systemic autoimmune diseases. *Rheumatol Adv Pract*. (2020) 4:rkaa006. doi: 10.1093/rap/rkaa006

26. Paknikar SS, Crowson CS, Davis JM, Thanarajasingam U. Exploring the role of antinuclear antibody positivity in the diagnosis, treatment, and health outcomes of patients with rheumatoid arthritis. *ACR Open Rheumatol*. (2021) 3:422–6. doi: 10.1002/acr2.11271

27. Dumitrescu L, Ritchie MD, Brown-Gentry K, Pulley JM, Basford M, Denny JC, et al. Assessing the accuracy of observer-reported ancestry in a biorepository linked to electronic medical records. *Genet Med*. (2010) 12:648–50. doi: 10.1097/GIM.0b013e3181efe2df

28. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight - reconsidering the use of race correction in clinical algorithms. *N Engl J Med*. (2020) 383:874–82. doi: 10.1056/NEJMms2004740

29. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. (2019) 366:447–53. doi: 10.1126/science.aax2342

30. Wang KY, Yang YH, Chuang YH, Chan PJ, Yu HH, Lee JH, et al. The initial manifestations and final diagnosis of patients with high and low titers of antinuclear antibodies after 6 months of follow-up. *J Microbiol Immunol Infect*. (2011) 44:222–8. doi: 10.1016/j.jmii.2011.01.019

31. Li X, Liu X, Cui J, Song W, Liang Y, Hu Y, et al. Epidemiological survey of antinuclear antibodies in healthy population and analysis of clinical characteristics of positive population. *J Clin Lab Anal*. (2019) 33:e22965. doi: 10.1002/jcla.22965

32. Berkelmans GFN, Read SH, Gudbjornsdottir S, Wild SH, Franzen S, van der Graaf Y, et al. Population median imputation was noninferior to complex approaches for imputing missing values in cardiovascular prediction models in clinical practice. *J Clin Epidemiol*. (2022) 145:70–80. doi: 10.1016/j.jclinepi.2022.01.011

33. Damoiseaux J, Andrade LEC, Carballo OG, Conrad K, Francescantonio PLC, Fritzler MJ, et al. Clinical relevance of HEp-2 indirect immunofluorescent patterns: the International Consensus on ANA patterns (ICAP) perspective. *Ann Rheum Dis*. (2019) 78:879–89. doi: 10.1136/annrheumdis-2018-214436

34. Harrell JFE. *Regression Modeling Strategies : With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. Cham: Springer International Publishing : Imprint: Springer (2015).

35. Fernandez-Delgado M, Sirsat MS, Cernadas E, Alawadi S, Barro S, Febrero-Bande M. An extensive experimental survey of regression methods. *Neural Netw*. (2019) 111:11–34. doi: 10.1016/j.neunet.2018.12.010

36. Sheridan RP, Wang WM, Liaw A, Ma J, Gifford EM. Extreme gradient boosting as a method for quantitative structure-activity relationships. *J Chem Inf Model*. (2016) 56:2353–60. doi: 10.1021/acs.jcim.6b00591

37. Xu Y, Yang X, Huang H, Peng C, Ge Y, Wu H, et al. Extreme gradient boosting model has a better performance in predicting the risk of 90-day readmissions in patients with ischaemic stroke. *J Stroke Cerebrovasc Dis*. (2019) 28:104441. doi: 10.1016/j.jstrokecerebrovasdis.2019.104441

38. Steyerberg EW, Harrell FE Jr., Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. (2001) 54:774–81. doi: 10.1016/s0895-4356(01)00341-9

39. Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal-external, and external validation *J Clin Epidemiol*. (2016) 69:245–7 doi: 10.1016/j.jclinepi.2015.04.005

40. Ertenli I, Kiraz S, Ozturk MA, Haznedaroglu I, Celik I, Calguneri M. Pathologic thrombopoiesis of rheumatoid arthritis. *Rheumatol Int*. (2003) 23:49–60. doi: 10.1007/s00296-003-0289-0

41. Gasparyan AY, Ayvazyan L, Mukanova U, Yessirkepov M, Kitas GD. The platelet-to-lymphocyte ratio as an inflammatory marker in rheumatic diseases. *Ann Lab Med*. (2019) 39:345–57. doi: 10.3343/alm.2019.39.4.345

42. Bitik B, Mercan R, Tufan A, Tezcan E, Kucuk H, Ilhan M, et al. Differential diagnosis of elevated erythrocyte sedimentation rate and C-reactive protein levels: a rheumatology perspective. *Eur J Rheumatol*. (2015) 2:131–4. doi: 10.5152/eurjrheum.2015.0113

43. Costenbader KH, Chibnik LB, Schur PH. Discordance between erythrocyte sedimentation rate and C-reactive protein measurements: clinical significance. *Clin Exp Rheumatol*. (2007) 25:746–9.

44. Brigden ML. Clinical utility of the erythrocyte sedimentation rate. *Am Fam Physician*. (1999) 60:1443–50.

45. Daniels LM, Tosh PK, Fiala JA, Schleck CD, Mandrekar JN, Beckman TJ. Extremely elevated erythrocyte sedimentation rates: associations with patients' Diagnoses, demographic characteristics, and comorbidities. *Mayo Clin Proc*. (2017) 92:1636–43. doi: 10.1016/j.mayocp.2017.07.018

46. Kirino Y, Remmers EF. Genetic architectures of seropositive and seronegative rheumatic diseases. *Nat Rev Rheumatol*. (2015) 11:401–14. doi: 10.1038/nrrheum.2015.41

47. Wei Q, Jiang Y, Xie J, Lv Q, Xie Y, Tu L, et al. Analysis of antinuclear antibody titers and patterns by using HEp-2 and primate liver tissue substrate indirect immunofluorescence assay in patients with systemic autoimmune rheumatic diseases. *J Clin Lab Anal*. (2020) 34:e23546. doi: 10.1002/jcla.23546

48. Romero-Alvarez V, Acero-Molina DA, Beltran-Ostos A, Bello-Gualteros JM, Romero-Sanchez C. Frequency of ANA/DFS70 in relatives of patients with rheumatoid arthritis compared to patients with rheumatoid arthritis and a healthy population, and its association with health status. *Reumatol Clin (Engl Ed)*. (2021) 17:67–73. doi: 10.1016/j.reuma.2019.02.003

49. Zhang JF, Ye XL, Duan M, Zhou XL, Yao ZQ, Zhao JX. Clinical and laboratory characteristics of rheumatoid arthritis with positive antinuclear antibody. *Beijing Da Xue Xue Bao Yi Xue Ban*. (2020) 52:1023–8. doi: 10.19723/j.issn.1671-167X.2020.06.006

50. Kanis JA, Johnell O, Oden A, Johansson H, McCloskey E. FRAX and the assessment of fracture probability in men and women from the UK. *Osteoporos Int*. (2008) 19:385–97. doi: 10.1007/s00198-007-0543-5

51. Goff DC Jr., Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. (2014) 129:S49–73. doi: 10.1161/01.cir.0000437741.48606.98

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 10 June 2024
DOI 10.3389/fimmu.2024.1345494

OPEN ACCESS

EDITED BY
Christine Gibson Parks,
National Institute of Environmental Health Sciences (NIH), United States

REVIEWED BY
Essi Laajala,
University of Helsinki, Finland
Jaclyn Goodrich,
University of Michigan, United States

*CORRESPONDENCE
Patrick M. Carry
patrick.carry@cuanschutz.edu

RECEIVED 27 November 2023
ACCEPTED 21 May 2024
PUBLISHED 10 June 2024

CITATION
Carry PM, Vanderlinden LA, Johnson RK, Buckner T, Steck AK, Kechris K, Yang IV, Fingerlin TE, Fiehn O, Rewers M and Norris JM (2024) Longitudinal changes in DNA methylation during the onset of islet autoimmunity differentiate between reversion versus progression of islet autoimmunity. *Front. Immunol.* 15:1345494. doi: 10.3389/fimmu.2024.1345494

Longitudinal changes in DNA methylation during the onset of islet autoimmunity differentiate between reversion versus progression of islet autoimmunity

Patrick M. Carry[1,2,3]*, Lauren A. Vanderlinden[2], Randi K. Johnson[2,3], Teresa Buckner[2,4], Andrea K. Steck[5], Katerina Kechris[2,3,6], Ivana V. Yang[3,7], Tasha E. Fingerlin[2,6,8], Oliver Fiehn[9], Marian Rewers[5] and Jill M. Norris[2]

[1]Colorado Program for Musculoskeletal Research, Department of Orthopedics, University of Colorado, Aurora, CO, United States, [2]Department of Epidemiology, Colorado School of Public Health, Aurora, CO, United States, [3]Department of Biomedical Informatics, School of Medicine, University of Colorado, Aurora, CO, United States, [4]Department of Kinesiology, Nutrition, and Dietetics, University of Northern Colorado, Greeley, CO, United States, [5]Barbara Davis Center, Department of Pediatrics, University of Colorado, Aurora, CO, United States, [6]Department of Biostatistics and Informatics, Colorado School of Public Health, Aurora, CO, United States, [7]Department of Medicine, University of Colorado, Aurora, CO, United States, [8]Department of Immunology and Genomic Medicine, National Jewish Health, Aurora, CO, United States, [9]University of California Davis West Coast Metabolomics Center, Davis, CA, United States

Background: Type 1 diabetes (T1D) is preceded by a heterogenous pre-clinical phase, islet autoimmunity (IA). We aimed to identify pre vs. post-IA seroconversion (SV) changes in DNAm that differed across three IA progression phenotypes, those who lose autoantibodies (reverters), progress to clinical T1D (progressors), or maintain autoantibody levels (maintainers).

Methods: This epigenome-wide association study (EWAS) included longitudinal DNAm measurements in blood (Illumina 450K and EPIC) from participants in Diabetes Autoimmunity Study in the Young (DAISY) who developed IA, one or more islet autoantibodies on at least two consecutive visits. We compared *reverters* - individuals who sero-reverted, negative for all autoantibodies on at least two consecutive visits and did not develop T1D (n=41); *maintainers* - continued to test positive for autoantibodies but did not develop T1D (n=60); *progressors* - developed clinical T1D (n=42). DNAm data were measured before (pre-SV visit) and after IA (post-SV visit). Linear mixed models were used to test for differences in pre- vs post-SV changes in DNAm across the three groups. Linear mixed models were also used to test for group differences in average DNAm. Cell proportions, age, and sex were adjusted for in all models. Median follow-up across all participants was 15.5 yrs. (interquartile range (IQR): 10.8-18.7).

Results: The median age at the pre-SV visit was 2.2 yrs. (IQR: 0.8-5.3) in progressors, compared to 6.0 yrs. (IQR: 1.3-8.4) in reverters, and 5.7 yrs. (IQR: 1.4-9.7) in maintainers. Median time between the visits was similar in reverters 1.4

yrs. (IQR: 1-1.9), maintainers 1.3 yrs. (IQR: 1.0-2.0), and progressors 1.8 yrs. (IQR: 1.0-2.0). Changes in DNAm, pre- vs post-SV, differed across the groups at one site (cg16066195) and 11 regions. Average DNAm (mean of pre- and post-SV) differed across 22 regions.

Conclusion: Differentially changing DNAm regions were located in genomic areas related to beta cell function, immune cell differentiation, and immune cell function.

KEYWORDS

DNA methylation, type 1 diabetes (T1D), DAISY, islet autoimmunity, reversion

1 Introduction

T1D is an autoimmune disorder with significant long-term morbidity. The pre-clinical phase is defined by the appearance of autoantibodies against pancreas cell antigens, termed islet autoimmunity (IA). There is strong evidence to support autoantibodies as a biomarker of T1D risk (1). However, IA is dynamic. While progression to T1D or multiple autoantibodies has been well characterized, a subset of individuals lose autoantibody positivity (2) and revert back to an autoantibody negative state. Autoantibody reversion was first described by Spencer et al (3) in a cohort of 685 individuals with a first degree relative affected by T1D. After 5 years, 7/20 developed T1D, 1 remained AB positive and 12/20 reverted. Transient autoantibody positivity has been described in several additional studies (4–6). However, these historical studies describing the transient nature of autoantibodies are difficult to interpret due to the development of more accurate autoantibody tests as well as differences in the definition of reversion. Vehik et al (2) conducted the most comprehensive and rigorous study of reversion in current literature. Among 596 individuals enrolled in The Environmental Determinants of Diabetes in the Young (TEDDY) study who developed one or more persistent autoantibodies, 21% reverted to an antibody negative state. Seroreversion was associated with significantly decreased risk of T1D (hazard ratio: 0.14, 95% CI: 0.04-0.59). Understanding the unique protective mechanisms occurring prior to or following IA that lead to IA reversion may have important implications for development of interventions that delay or prevent progression to T1D.

Genetic variation is a well-established risk factor for T1D (7). However, heterogeneity in disease concordance among monozygotic twins (8) as well as temporal changes in both T1D incidence (9) and age at T1D onset (10) in population studies have created a strong interest in the role of the environment in the etiology of T1D. Epigenetic modifications such as DNA methylation (DNAm) may represent a mechanistic pathway between genetic susceptibility, environmental exposures, and progression or reversion of IA. Epigenetics broadly describes a class of modifiable mechanisms that can regulate gene expression and are sensitive to external stimuli (11). DNAm is a frequently studied epigenetic biomarker that is postulated to play a role in autoimmune diseases as epigenetic mechanisms are important regulators of immune cell differentiation, plasticity and function (12, 13). DNAm changes prior to and during the IA phase may provide key information about underlying epigenetic profiles that explain progression or reversion from IA.

Previous epigenome wide studies have identified significant associations between DNAm and T1D (14–17). However, associations have been inconsistent and many of the studies have focused on static and/or post-T1D differences in DNAm between cases and controls (14–16). Although important in understanding the etiology of T1D, DNAm differences obtained from a single time point are difficult to interpret as it is not possible to determine when the changes occurred and moreover, whether they are the cause or consequence of the disease process. Understanding the timing of the changes is key to identifying external factors that cause these changes and therefore, may be amenable to preventative interventions. The purpose of this study was to test DNAm obtained before and after IA seroconversion (SV) in the Diabetes Autoimmunity Study in the Young (DAISY). We aimed to identify pre vs. post-SV changes in DNAm that differed across three distinct IA progression phenotypes, those who lose autoantibodies (reverters), progress to clinical T1D (progressors), or maintain autoantibody levels (maintainers).

2 Materials and methods

2.1 Study population

We reviewed individuals from the Diabetes Autoimmunity Study in the Young (DAISY) who developed islet autoimmunity (IA) between February 1994 and February 2019. DAISY is a longitudinal birth cohort study that includes n=2544 children at high risk for T1D. Subjects are recruited from two high risk populations, those with a first degree relative (FDR) with T1D or

those with a high-risk genotype, [defined as DRB1*04, DQB1*0302/DRB1*0301, DQB1*0201 (DR3/4 DQ8)]. Subjects complete study visits at 9, 15, and 24 months. Following the 24-month visit, study visits occur annually. As described previously (18), radioimmunoassays were used to test serum samples for autoantibodies to insulin (IAA), GAD65 (GAA), and IA-2 (IA-2A). Prior to 2010, GADA and IA-2A were tested using a combined radioassay (19). The National Institute of Diabetes and Digestive and Kidney Diseases harmonized assay was used to test for GADA and IA-2A after 2010 (20). Serum samples from individuals positive for GAD65, IAA, or IA-2 were tested for ZnT8A following development and implementation of the ZnT8 assay (21). If autoantibodies are detected, participants return for study visits every 3-6 months.

Islet autoimmunity (IA) was defined as the presence of one or more autoantibodies (see above) on at least two consecutive visits 3-6 months apart. The first visit among these consecutive autoantibody positive visits designated the start of IA, referred to as seroconversion (SV) throughout the remainder of the manuscript. We defined the three autoimmune progression phenotypes based on the autoantibody testing. The *reverter group* was defined as individuals who reverted for all autoantibodies during two or more consecutive visits, did not develop T1D, and were autoantibody negative for all autoantibodies at their last DAISY visit. The *maintainer group* was defined as individuals who continued to test positive for islet autoantibodies and did not develop T1D at the time of their last visit. The *progressor group* was defined as individuals who developed clinical T1D.

Among individuals who developed IA during DAISY and underwent autoantibody testing for a minimum of two or more study visits (n=213), we excluded individuals for the following: missing a pre- or post-SV blood sample (n=54), onset of IA unclear due to gaps (>365 days) in study visits (n=2), missing study visit prior to initial pre-SV positive visit (n=14). The Colorado Multiple Institutional Review Board approved all DAISY protocols (COMIRB 92-080). Informed consent and assent, if appropriate, was obtained from the parents/legal guardians of all children prior to participation in any research related activities.

2.2 Methylation measurements

Methylation measurements were obtained from peripheral whole blood samples collected at multiple time-points in individuals from DAISY. The Infinium HumanMethylation 450K Beadchip platform (Illumina, San Diego, CA, USA) was used to obtain methylation measurements on a subset of samples. The 850K Infininium MethylationEPIC BeadChip (Illumina, San Diego, CA, USA) was used to obtain measurements on the remaining samples. Two platforms were used due to changes in technology during the course of the study. Samples were randomly assigned to the two platforms making sure all timepoints from the same individual were included on the same platform.

DNA was bisulfite converted using the Zymo EZ DNA Methylation kit (Zymo Research, CA, USA). The bisulfite-converted DNA was labeled with fluorescent dyes and hybridized to 450K and 850K DNAm arrays. Samples were arranged on the plates in a specific sequence to minimize within and between batch effects (plate effects are represented by first 11 digits of the array variable on GEO). The minfi (v1.12.0) package (22) in R (v3.5.2) was used to perform quality control (QC) checks at the sample level. The processing pipeline is described in greater detail in Vanderlinden et al (23).

The DNAm probes were annotated to the genome based on the hg19 genome build using the Illumina annotation manifest files. Non-autosomal CpGs or CpGs located within or near (<2 base pairs) known single nucleotide polymorphisms (SNPs) were excluded. CpG sites with a beta range <3% on both platforms were removed from analysis. A total of n=198,008 overlapping DNAm probes met our filtering criteria and were used in subsequent analyses. Normalized M-values (SeSAMe (v1.0.0) pipeline with Noob normalization) were used in all statistical analyses. We use the term DNAm probe and the probe identifier when referring to the data in the Methods and Results. However, each probe is designed to measure DNAm at a single CpG site which is used as a more general term in the Discussion. See Figure 1 for an overview of the study methods.

2.3 Overlapping gene expression measurements

Gene expression data were available in a subset of individuals (n=36) at the post-SV visit. RNA processing and quantification is described in greater detail in Carry et al (24). In brief, paired end sequencing was performed using the Illumina NovaSEQ 6000™ system and samples were quantified against the Ensembl reference transcriptome (hg19, version 87) using the RSEM algorithm (25). Data were quantile normalized using DESeq2 (26), re-normalized using RUV (27), and then transformed using the regularized log function (26). The transformed data were used in all subsequent statistical analyses.

2.4 Overlapping metabolomics measurements

Untargeted metabolomics data were available in a subset (n=110) of individuals at both the pre-SV and post-SV visits. Metabolomics processing and quantification is described in greater detail in Carry et al (28). In brief, non-fasting plasma samples were used to quantify metabolite levels using three untargeted panels, HILIC panel: HILIC-QTOF MS/MS (29), GCTOF panel: GC-TOF-MS (30), and Lipid panel: CSH-QTOF MS/MS (31). BinBase (32) was used to process and annotate the GC-TOF-MS data. MS-Dial (33) was used to process and annotate the liquid chromatography (LC), CSH-QTOF-MS and HILIC-QTOF-MS, data. LipidBlast (34) and Massbank of North America were also used to annotate the complex lipids (http://mona.fiehnlab.ucdavis.edu/). Metabolomic data were normalized using the systematic error removal using random forest (SERRF) algorithm (35). All metabolites were Box-Cox transformed prior to statistical analysis.

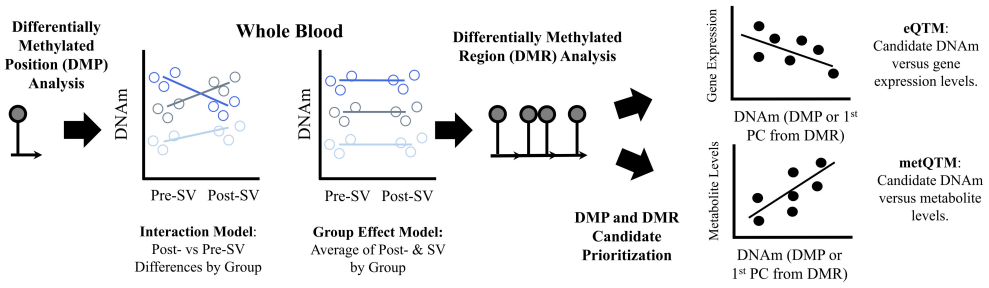


FIGURE 1
Summary of methods used to identify and prioritize DNAm candidates. Description: We used an epigenome wide association study design to identify differentially methylated positions (DMP) associated with the three islet autoimmunity progression phenotypes, reverters, maintainer, or progressors. We used two DMP models (1) an interaction model that tested whether changes in DNA methylation (DNAm) levels at single CpGs pre-IA versus post-IA differed across groups and (2) a group effect model that tested whether average methylation levels (pre- and post-IA) differed across groups. We also performed regional analyses (differentially methylated regions or DMRs) based on single CpG sites from the two models to identify regions with consistent methylation effects. We identified regions where average regional methylation levels differed between groups (μDMRs) as well as regions where changes in regional methylation levels pre- vs post-IA differed across groups (ΔDMRs). In order to prioritize regions, we tested whether the DNAm candidates identified in our analysis were associated with gene expression levels post-SV, an expression quantitative trait methylation analysis (eQTM). To account for the multiple CpGs within each DMR, we used a principal component analysis to capture common patterns across all CpGs included in the DMR. We identified cis-eQTMs (midpoint of region +/- 500 KB of the TSS of the gene) by testing the correlation between gene expression and the 1st principal component. We also tested the correlation between DNAm candidates and metabolite levels obtained from overlapping samples, a metabolite quantitative trait methylation analysis (metQTM). We used a principal component analysis to capture common patterns across all CpGs included in the candidate DMRs. We tested the correlation between metabolite levels and the 1st principal component. CpGs are represented by lollipop plots in the figure.

2.5 Genetic ancestry

Ancestry principal components (PC) were estimated for all study participants from genetic data collected in DAISY. Sample processing and genotyping were performed at the University of Virginia School of Medicine Center for Public Health Genomics based on exome genotyping (Illumina HumanCoreExome-24 BeadChip, N=283) or whole genome sequencing (N=162) from the larger DAISY population, see Buckner et al (36) for a more complete description of the genetic processing and calculation of the genetic ancestry PCs.

2.6 Statistical analyses

The overall methods workflow is summarized in Figure 1. Linear mixed models were used to test for differences in DNAm between the pre- and post-SV visit across reverters, maintainers, and progressors (autoimmune phenotype*visit interaction). Separate linear mixed models were also used to test for differences in average DNAm (mean of the DNAm levels at the pre- and post- SV visits) between the autoimmune phenotypes (group effect). Platform (EPIC vs 450K), age, sex, and cell proportions (estimated using the minfi (v1.12.0) package (22) implementation of the Houseman method) were adjusted for in all models. The group effect models were also adjusted for population ancestry (see Supplementary Material for complete description of ancestry data). Ancestry data (1st 2 PCs) were unavailable for 2 individuals in the group effect model and thus, these individuals were not included in this analysis. See Appendix 1 (Data Sheet 1) for the linear mixed model code. We did not adjust for ancestry in the interaction (autoimmune phenotype*visit) models because the interaction models test for within individual differences, and thus are less likely to be impacted by time invariant confounders such as population ancestry. The Benjamini Hochberg false discovery rate (FDR), was used to correct for multiple comparisons (37). Significance was assessed based on the FDR adjusted p-value <0.10. Model diagnostics are described in the Supplementary Files (Data Sheet 2), see Appendix 2, Figures A–C and Table A.

The comb-p python software package (38) was used to identify differentially methylated regions (DMRs). Within the comb-p pipeline, we used a seed p-value of 0.1 and then searched for adjacent probes within a window of 500 bases, using a step size of 50 bases. Comb-p combines probes within this window and then calculates an overall, spatially corrected p value for the entire region based on the Stouffer-Liptak method. The Sidak method is used to adjust the overall regional p values for multiple testing. Regional analyses were performed based on the individual DNAm probes from the interaction (post- vs pre-SV changes by autoimmune phenotype), referred to as differentially changing DMRs (ΔDMR) throughout the remainder of the manuscript. Regional analyses were also performed based on DNAm probes from the main effect model (differences in average of pre- and post-SV DNAm between groups), referred to as average DMRs (μDMR) throughout the remainder of the manuscript. For both regional analyses, we reviewed all regions with ≥4 DNAm probes that were significant at the combined Sidak adjusted region p value of 0.10. Because the interaction and group effect p values are based on a two degree of freedom test (numerator degrees of freedom for the overall F-test), it is possible for the DMR to capture a set of DNAm probes with similar p values but substantial heterogeneity in the directions of effect within the three groups. Therefore, for the ΔDMRs, we retained regions with a consistent direction of effect, defined as a region where the direction of change in DNAm between the two visits (hyper methylation or hypo methylation) was consistent across 100% of the DNAm probes within the region in one or more of the study groups. For the μDMRs, we retained regions where the direction of effect (hypo or

hypermethylation) for one or more of the pairwise group comparisons was consistent across 100% of the DNAm probes included in the region.

2.7 Expression quantitative trait methylation analysis: correlation between gene expression and DNAm candidates

In order to better understand our primary DNAm results, we tested the correlation between gene expression levels and our DNAm candidates, one DMP, 11 ΔDMRs, and 22 μDMRs in a subset of individuals (n=36, see Appendix 3, Table B) with methylation data pre- and post-SV as well as gene expression data post-SV. First, linear mixed models were used to regress out age, sex, platform, and cell proportions from the DNAm values at each of the candidate CpG sites. Ancestry PC1 and ancestry PC2 were also regressed out from all CpG sites included in the μDMRs candidate regions. Next, using the residuals from the linear mixed models, the within individual differences in DNAm (post-SV minus pre-SV) were used to represent changes in DNAm between the study visits for each of the CpG sites included in the ΔDMRs. The average residual values from the post-SV and pre-SV study visits were used to represent average methylation for each of the CpG sites within the μDMRs. Next, we performed a principal component analysis of DNAm levels across the region-specific CpG sets. For each DMR, the first PC was extracted for subsequent testing, allowing us to consider all CpG sites together rather than testing many individual sites separately. Linear regression models were then used to regress out the effects of age and sex from the gene expression levels. Finally, Spearman correlation coefficients were used to test the correlation between DNAm and gene expression residuals. We looked for cis-eQTMs, defined as genes significant at the FDR adjusted p value of 0.10 where transcription start site was +/- 500 KB of the midpoint of the DMR. FDR adjustment was based on the total number of DNAm cis-gene pairs (256 transcript DNAm pairs for the ΔDMR candidates and 544 transcript DNAm pairs for the μDMR candidates).

2.8 Metabolite quantitative trait methylation analysis: correlation between metabolite levels and DNAm candidates

We tested the correlation between DNAm and untargeted metabolite levels in a subset of our study population (n=110, see Appendix 3, Table B) with DNAm and metabolomics data available both pre- and post-SV. Only data from overlapping samples was included in this supplementary analysis. Linear models were used to regress age and sex from the Box-Cox transformed metabolite levels at each visit. Consistent with the DNAm methods, using the residuals from the linear mixed models, the difference between metabolite residuals at each visit (post-SV minus pre-SV residuals) was used to represent change in metabolites and the average residual values (average of post-SV and pre-SV residuals) were used to represent average metabolite values. For the ΔDMR candidates and the single DMP candidate, linear regression models were then used to test the correlation between the change in metabolites versus the ΔDMR PCs (described above) as well as the single DMP candidate. For the μDMR candidates, linear regression models were then used to test the correlation between average metabolite levels versus the μDMR PCs (described above). False discovery (FDR) rate adjusted p values were calculated for all individual metabolite DNAm candidate pairs according to methods described by Benjamini and Hochberg (37). FDR adjusted p values were calculated separately for each platform. Only annotated metabolites from the HILIC (81 metabolites), Lipid (373 metabolites), and GC-TOF (98 metabolites) panels were evaluated in subsequent analyses. Metabolites were evaluated at an FDR adjusted p value of 0.10.

3 Results

3.1 Study population

The final study population included 60 individuals in the maintainer group, 42 individuals in the progressor group, and 41 individuals in the reverter group. At both the pre-SV and post-SV visits, age differed by group, and the estimated cell proportions differed by group at the post-SV visit (Table 1). At the time of data analysis, duration of follow-up, defined as median time from the initial visit to the development of T1D or last study visit, was 9.3 years (IQR: 6.1 to 12.3 years) for the progressors, 16.5 years for the maintainers (IQR: 14.3 to 20.9 years) and 16.6 years for the reverters (IQR: 15.2 to 20.2 years).

The specific autoantibody subgroups present at the onset of seroconversion in the three groups are described in greater detail in Appendix 4 (Data Sheet 4), Table C. As expected, the prevalence of multiple autoantibodies at serconversion was higher in progressors (31%) relative to maintainers (18%) and reverters (0%). Across the entire islet autoimmunity follow-up period, the occurrence of multiple autoantibodies at one or more study visit(s) following IA seroconversion was also higher in progressors (86%) compared to maintainers (58%). Among reverters, 10% developed multiple autoantibodies at one of more study visit(s) during the time period between seroconversion (IA onset) and seroreversion.

3.2 Differentially methylated position analysis

Change in methylation at the DNAm site cg16066195 on chr 7 was significantly (FDR adjusted p value=0.0174) different across groups. The reverter group was characterized by an increase in DNAm between pre- and post-SV visits (ie, a positive slope) whereas the progressor and maintainer groups were characterized by no change or a decrease in DNAm (Figure 2). This site is an island CpG site (CpG island *chr7:73703458-73704127)* that maps to an area near the *CLIP2* gene.

We also tested whether average DNAm (mean of DNAm levels pre- and post-SV) differed across groups. No DNAm probe was significant at the FDR adjusted alpha level of 0.10.

10.3389/fimmu.2024.1345494

TABLE 1 Demographics and clinical characteristics.

	Maintainer n=60		Progressor n=42		Reverter n=41		P Value
	Median | Freq	IQR | %	Median | Freq	IQR | %	Median | Freq	IQR | %	
Pre-Islet Autoimmunity Visit							
Age at Visit, median (IQR)	5.7	1.4-9.7	2.2	0.8-5.3	6.0	1.3-8.4	0.0079
CD8T, median (IQR)	13.3%	9.4-16.6%	14.6%	11.8-15.9%	12.2%	9.7-16.1%	0.3864
CD4T, median (IQR)	22.0%	15.6-26.1%	23.4%	17.3-31.8%	19.3%	16.1-25.5%	0.1959
NK, median (IQR)	1.4%	0.0-4.7%	0.0%	0.0-1.5%	1.3%	0.0-3.1%	0.0653
Bcell, median (IQR)	15.3%	10.6-18.5%	17.9%	13.4-22.6%	14.9%	10.2-19.7%	0.1599
Mono, median (IQR)	8.3%	6.9-10.3%	7.5%	5.2-9.4%	7.6%	6.2-9.5%	0.3390
Gran, median (IQR)	38.5%	30.6-50.9%	35.5%	24.7-44.6%	42.8%	32.0-52.0%	0.2205
Post-Islet Autoimmunity Visit							
Age at Visit, median (IQR)	8.0	5.2-11.3	4.9	2.4-9.4	7.1	3.1-10.0	0.0087
CD8T, median (IQR)	11.8%	9.5-15.6%	14.6%	11.3-16.7%	12.3%	8.9-16.4%	0.1183
CD4T, median (IQR)	17.6%	13.1-22.1%	21.7%	17.3-26.9%	17.6%	13.0-21.7%	0.0061
NK, median (IQR)	2.7%	0.0-6.0%	0.0%	0.0-3.5%	1.3%	0.0-4.1%	0.0018
Bcell, median (IQR)	11.2%	8.6-15.0%	16.5%	12.7-19.7%	13.1%	8.3-16.7%	0.0011
Mono, median (IQR)	9.1%	7.8-10.8%	7.8%	4.8-9.3%	8.5%	7.0-10.1%	0.0293
Gran, median (IQR)	46.0%	39.6-52.7%	37.9%	28.6-44.4%	47.8%	38.4-53.6%	0.0025
Non-Hispanic White Ethnicity, freq (%)	43	71.7%	38	90.5%	29	70.7%	0.0458
Female Sex, freq (%)	34	56.7%	19	45.2%	21	51.2%	0.5224
HLDR3/4 High Risk Genotype, freq (%)	16	26.7%	19	45.2%	10	24.4%	0.0711
First Degree Relative with T1D, freq (%)	38	63.3%	25	59.5%	19	46.3%	0.2242

IQR, interquartile range; CD8T, cytotoxic T cells; CD4T, T helper cells; NK, natural killer T cells; Mono, monocytes; Gran, granulocytes.

3.3 Differentially methylated region analysis

We also tested for genomic regions (Figure 1). In contrast to the single CpG site (DMP) analysis, the regional analysis allowed us to identify multiple CpG sites that demonstrated similar DNAm changes between the pre- and post-SV visits across the three study groups (ΔDMRs). We focused on FDR significant regions of ≥4 DNAm probes where the direction of the change in DNAm (between the pre-SV and post-SV visits) was consistent (100% of probes changed in a similar direction) within one or more of the groups. We identified 11 candidate DMRs (Table 2; Figure 3).

We also tested for regions where the average DNAm levels at the pre- and post-SV visits differed across the groups (μDMRs). We identified 22 FDR significant μDMRs of ≥4 DNAm probes where the direction of the pairwise group differences in DNAm was consistent across all CpG sites included in the region (Table 3; Figure 4).

3.4 eQTM candidate prioritization

We tested the correlation between DNAm and cis- gene expression levels in a subset of overlapping samples. The availability of individual level DNAm data allowed us to look at the entire DMR together. Based on the ΔDMR candidates, we identified two FDR significant cis eQTMs representing one DMR and two gene transcripts, GNAS and ATP5E (ΔDMR1, region on chromosome 20, see Table 4). Within this region, increased DNAm post- vs pre-SV was positively associated with expression of GNAS and ATP5E (see Table 4).

3.5 Metabolite quantitative trait methylation analysis candidate prioritization in overlapping samples

We tested whether the single DMP candidate, cg16066195, as well as the candidate DNAm regions identified in our primary

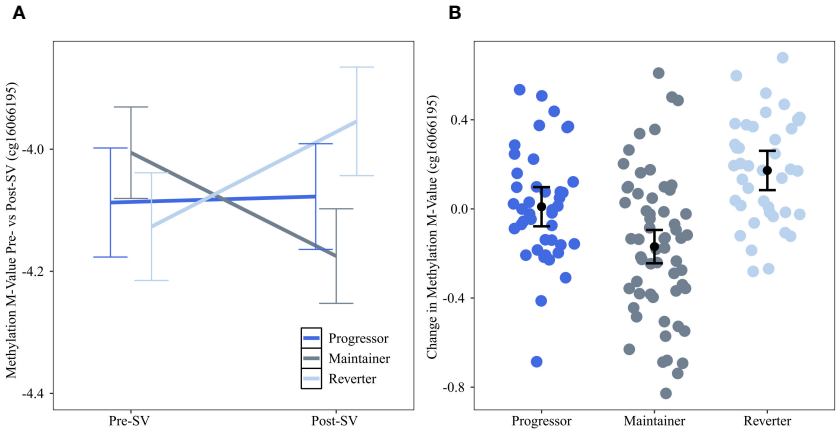


FIGURE 2
Changes in DNAm between the pre- and post-SV visits at cg16066195 across the three IA progression phenotypes. Description: **(A)** provides the average methylation M-values and corresponding 95% confidence intervals within the three IA progression phenotypes pre- and post-SV. **(B)** describes the individual level changes in methylation m-values (y-axis) between the post-SV visit relative to the pre-SV visit in the three IA progression phenotypes (x-axis). Positive values represent increasing DNAm whereas negative values represent decreasing methylation between visits. All DNAm values in **(A, B)** have been adjusted for age, sex, and cell proportions.

analysis were associated with metabolite levels. Consistent with the eQTM analysis, we regressed out age and sex from annotated metabolites and then tested the correlation between annotated metabolites versus DNAm regional PCs. Based on the ΔDMR candidates, we identified 26 annotated metabolites from the Lipid panel that were correlated with 4 DMRs (see Table 5; Figure 5). ΔDMR 8 was correlated with multiple lipids, primarily PCs, ΔDMR 5 was also correlated with multiple lipids, primarily correlated with TGs (fats). ΔDMR 9 and ΔDMR 2 were correlated with a single lipid, an ether lipid, and a TG, respectively. Metabolite candidates primarily consisted of odd-chain fatty acid containing lipid species (OCFA). Furthermore, the majority of the metabolites (29/30) were positively correlated with increasing DNAm levels. The μDMR candidate regions as well as the single DMP candidate were not significantly associated with metabolite levels at our FDR adjusted cutoff of 0.10.

4 Discussion

Epigenetic biomarkers are appealing in the study of complex diseases such as T1D based on their heritability, role in gene expression, and responsiveness to external stimuli. Epigenetic effects in observational studies are challenging to interpret because it is often not possible to determine whether DNA methylation (DNAm) is causative or secondary to the disease process. A strength of our study is the longitudinal analysis of DNAm levels both before and after the onset of IA. We identified a single CpG site as well as genomic regions where changes in DNAm between the post-SV and pre-SV visits were significantly different across the IA progression phenotypes. We also identified regions where average DNAm levels pre- and post-SV differed across the progression phenotypes. Together, the DNAm regions have potential biological relevance to T1D etiology based on their potential role in immune and beta cell function.

We identified a DNAm site, cg16066195, on chromosome 7 where DNAm levels increased between the pre- and post-SV visits among individuals who reverted to an IA negative state (reverters) compared to progressors (who showed no change in DNAm) and maintainers (who showed decreasing DNAm, Figure 2). This island CpG is located near the transcription start site for the protein coding gene *CLIP2*. In a mouse model of diet induced changes in beta cell expression, *CLIP2* gene expression was significantly downregulated among mice fed a carbohydrate containing diabetogenic high-fat diet relative to mice fed a diabetes-protective carbohydrate free high-fat diet (39). Furthermore, SNPs within *CLIP2* (rs2528994 and rs512023) have demonstrated modest associations with T2D in both the Diabetes Genetics Initiative (40) and the Wellcome Trust Case Control Consortium (41).

Our methylation analysis also identified numerous regions where average methylation post- and pre-SV differed across the autoimmune phenotypes in areas of the genome potentially relevant to T1D etiology. We identified a DMR on chromosome 12, μDMR4, characterized by hypermethylation in the reverter group relative to the progressor and maintainer groups (Figure 4). This includes 4 probes that, based on the ENCODE Project Consortium (42), are located in a known enhancer region. Three of the four probes within this region are located within the transcription start site for *NRIP2*, predicted to act upstream or within the notch signaling pathway (43). This pathway is relevant to T1D (44) based on its role in immune cell differentiation and function (45) as well as pancreas development (46), islet cell function (47), and islet cell survival (48). All four probes within μDMR4 are also located within the 5'UTR region for *ITFG2*, a gene expressed in numerous tissues including immune cells. Mouse and *in vitro* models have demonstrated that *ITFG2* deficiency alters B cell maturation and migration (49). In a lupus mouse model, MRL/lpr, autoimmunity development occurred earlier and was more severe in *ITFG2* deficient mice (49). Together, these findings suggest a potential role for *ITFG2*

TABLE 2 Regions where DNAm changes between the post- and pre-SV visits were consistently different across groups (group*visit interaction).

DMR ID	Chr.	Start	Stop	Gene	N CpG Sites	Sidak Adj. Region P	Leading CpG Site	Slope % R*	Slope % P*	Slope % M*	Median Slope R†	Median Slope P†	Median Slope M†
ΔDMR 1	chr20	57426538	57427974	GNAS; GNASAS; GNAS-AS1	29	8.33E-05	cg26496204	69%	100%	100%	0.01	-0.03	-0.05
ΔDMR 2	chr20	36148604	36149751	BLCAP; NNAT	30	1.37E-04	cg24675557	100%	80%	80%	0.05	-0.02	-0.02
ΔDMR 3	chr1	75198582	75199118	TYW3; CRYZ; RP11-17E13.3	8	3.40E-03	cg00121533	100%	88%	100%	0.06	0.04	-0.08
ΔDMR 4	chr14	101291068	101293727	MEG3	25	6.74E-03	cg14034270	96%	85%	100%	0.02	-0.02	-0.02
ΔDMR 5	chr11	1296469	1297386	TOLLIP	7	1.81E-02	cg11095027	86%	57%	100%	0.03	0.03	-0.07
ΔDMR 6	chr15	91473059	91473570	UNC45A	8	2.00E-02	cg03291024	75%	100%	100%	0.01	0.09	-0.09
ΔDMR 7	chr5	1245669	1246292	SLC6A18	4	3.38E-02	cg09075844	100%	100%	100%	-0.03	0.03	-0.06
ΔDMR 8	chr6	170597377	170597899	DLL1	4	3.66E-02	cg05228964	50%	100%	100%	<0.01	0.10	-0.04
ΔDMR 9	chr6	28945322	28945493	RN7SL471P‡	4	6.09E-02	cg10919664	100%	100%	100%	0.07	0.06	-0.16
ΔDMR 10	chr6	27647713	27648355	RP1-15D7.1‡	4	7.14E-02	cg25106913	75%	75%	100%	<0.01	0.06	-0.05
ΔDMR 11	chr5	1867978	1868694	IRX4‡	6	8.71E-02	cg14773178	83%	100%	100%	0.04	0.08	-0.08

DMRs limited to regions with a minimum of 4 probes and 100% of within group slopes in the same direction for one or more groups.
Chr., chromosome.
Start/Stop, DMR start and stop position.
Gene, Gene annotation from the Illumina manifest file, based on UCSC reference genes mapped to CpG sites within DMR and/or genes mapped to CpG sites within known regulatory regions, if gene was not annotated within the Illumina manifest file, noted with ‡, gene name based on closest transcription start site.
Leading CpG site, most significant DMP within the region.
Sidak Adj. Region P, regional p value corrected for multiple testing based on number similarly sized regions possible based on genomic coverage in the DMR analysis.
*R, reverters; P, progressors; M, maintainers, Percent of within group slopes (Pre-SV vs Post-SV) in the same direction (hypo (–) or hyper (+) methylation) across all the probes included in the DMR.
†Median slope (Pre-SV vs Post-SV) across all probes included in the DMR for each group, (+) values indicate increasing DNAm (–), indicate decreasing DNAm.

10.3389/fimmu.2024.1345494

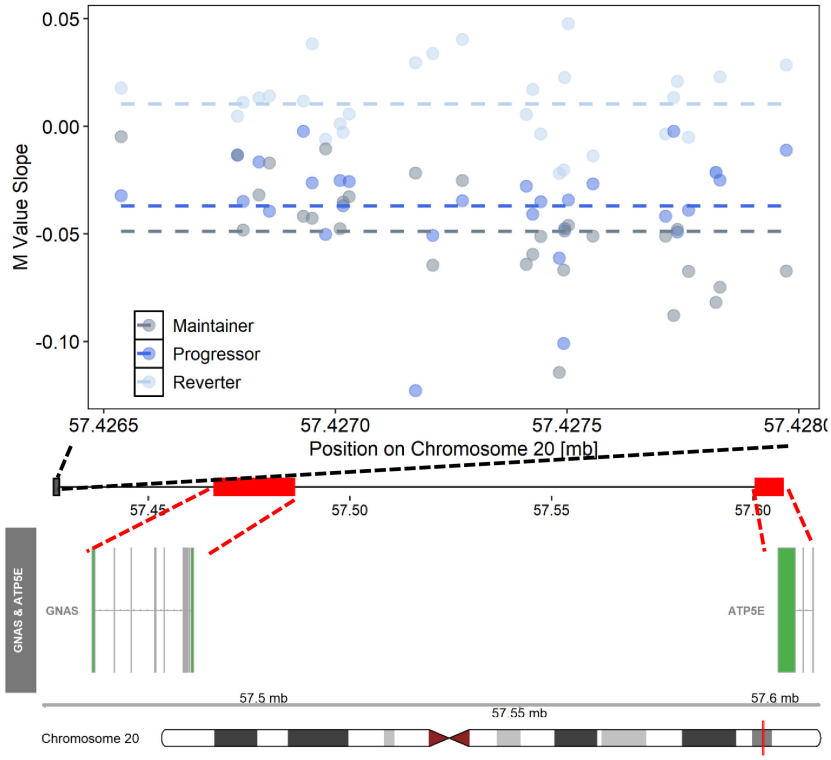


FIGURE 3
Differentially changing methylation region on chromosome 20 where changes in DNAm (pre- vs post-SV) differed across the three IA progression phenotypes. Description. Region on chromosome 20 loc 57426538 to 57427974 (ΔDMR1) where the change in DNA methylation (DNAm) post- vs pre-SV differed across groups. In the top panel, each dot represents the within group slopes (y-axis) or changes in DNAm m-values between the post-SV and pre-SV visit at each of the CpG sites included ΔDMR 1. The x-axis represents the position (mb) of the CpGs within the region. All slope values were adjusted for age, sex, and cell proportions. Positive values indicate methylation values increased following IA seroconversion whereas negative values indicate methylation decreased following IA seroconversion. The dashed lines represent the average slope value within each group across the entire region. The middle panel represents the location of the region (black solid square) relative to the closest genes, *GNAS* and *ATP5E* (red solid boxes). There are multiple known isoforms for *GNAS* and *ATP5E*, the bottom panel displays the most biologically relevant or consensus transcript based on the Ensembl database. The red line on the ideogram, bottom of the figure, represents the location of *GNAS* and *ATP5E* on chromosome 20.

TABLE 3 Regions where average of post- and pre-SV DNAm levels were consistently different across groups (group main effect).

DMR ID	Chr.	Start	Stop	Gene	N Probes	Sidak Adj. Region P	Leading CpG Site	Median PvR‡	Median RvM‡	Median PvM‡
µDMR 1	chr1	180922636	180923341	RP11-46A10.4; RP11-46A10.5	4	1.38E-05	cg00579423	0.09	0.37	0.46
µDMR 2	chr10	99338056	99338241	ANKRD2	4	1.75E-04	cg27469738	-0.11	0.26	0.17
µDMR 3	chr10	52008360	52008906	ASAH2	4	6.45E-03	cg24123634	-0.07	-0.02	-0.11
µDMR 4	chr12	2943902	2944481	NRIP2; ITFG2	4	7.06E-03	cg02852959	-0.15	0.19	0.04
µDMR 5	chr12	75784855	75785098	GLIPR1L2; CAPS2	6	7.59E-03	cg12351126	0.10	0.24	0.34
µDMR 6	chr12	51566379	51567113	TFCP2	7	1.24E-02	cg19016289	0.05	0.15	0.2
µDMR 7	chr1	1289835	1290713	MXRA8	6	1.61E-02	cg07284273	-0.16	0.33	0.15
µDMR 8	chr15	72766637	72767333	ARIH1; RP11-1007O24.3	4	1.93E-02	cg26880891	0.09	0.02	0.14
µDMR 9	chr19	45206843	45207560	CEACAM16	4	2.78E-02	cg24091949	-0.09	-0.04	-0.13
µDMR 10	chr19	2250901	2251068	AMH	4	2.83E-02	cg23218559	-0.18	0.38	0.21

(Continued)

TABLE 3 Continued

DMR ID	Chr.	Start	Stop	Gene	N Probes	Sidak Adj. Region P	Leading CpG Site	Median PvR‡	Median RvM‡	Median PvM‡
µDMR 11	chr18	7567426	7568266	PTPRM	5	3.44E-02	cg05870479	0.09	0.04	0.11
µDMR 12	chr15	85524778	85525674	PDE8A	4	4.02E-02	cg02839273	0.05	0.05	0.13
µDMR 13	chr2	85765644	85766105	MAT2A	4	4.39E-02	cg06978067	0.08	0.05	0.13
µDMR 14	chr19	48048129	48049234	ZNF541	4	4.90E-02	cg22341310	-0.12	0.17	0.06
µDMR 15	chr4	4861683	4862241	MSX1	4	5.94E-02	cg11930592	0.12	-0.04	0.08
µDMR 16	chr11	598325	599091	PHRF1	5	7.14E-02	cg12921473	-0.06	-0.05	-0.10
µDMR 17	chr5	101119084	101119767	OR7H2P*	4	7.67E-02	cg12197752	0.09	0.18	0.29
µDMR 18	chr13	42031761	42032737	C13orf15; RGCC	4	8.16E-02	cg18495682	0.06	0.02	0.09
µDMR 19	chr3	38206610	38207525	OXSR1	4	8.20E-02	cg19728055	0.07	0.05	0.11
µDMR 20	chr10	14372431	14372914	FRMD4A	5	8.45E-02	cg05755354	-0.16	-0.02	-0.18
µDMR 21	chr8	145550361	145551157	DGAT1	5	8.72E-02	cg11127482	0.06	0.04	0.11
µDMR 22	chr11	128693473	128694916	FLI1*; KCNJ1*	9	9.44E-02	cg15509024	-0.12	-0.09	-0.18

DMRs limited to regions with a minimum of 4 probes and direction of pairwise comparison was consistent across all probes in the region.
Chr., chromosome.
Start/Stop, DMR start and stop position.
Sidak Adj. Region P, regional p value corrected for multiple testing based on number similarly sized regions possible based on genomic coverage in the DMR analysis
Leading CpG site = most significant DMP within the region
Gene, Gene annotation from the Illumina manifest file, based on UCSC reference genes mapped to CpG sites within DMR and/or genes mapped to CpG sites within known regulatory regions, if gene was not annotated within the Illumina manifest file, noted with *gene name based on closest transcription start site.
‡R, reverters; P, progressors; M, maintainers, Median effect size across the region representing difference in methylation M values between groups.

in B cell differentiation and as a potential regulator of autoimmunity. Although, average methylation within DMR4 was not correlated with expression of ITFG2 or NRIP2 in our secondary eQTM analysis, three probes within µDMR4 (cg05194726; cg06997549; cg02852959) were correlated with expression of both ITFG2 and NRIP2 in whole blood based on the BIOS QTL browser (50), an online resource that provides a searchable database of FDR significant associations between DNAm and gene expression (eQTM). Additional work is needed to understand the connections between methylation within this region on chr 12, ITFG2 expression, NRIP2 expression, and T1D etiology.

We also identified several regions of differentially changing DNAm that are potentially relevant to T1D etiology based on known associations between DNAm in these regions and relevant environmental risk factors. We identified a region on chr 20 near the *GNAS/GNASAS* loci, ΔDMR 1, that was characterized by decreasing DNAm pre- vs post-SV in maintainers and progressors relative to reverters (Table 2; Figure 3). Based on the ENCODE Project Consortium (42), 25 of the 29 probes in ΔDMR 1 are located within a DNAase hypersensitivity region and 4 probes are known to interact with transcription factor binding. DNAm in this region is responsive to environmental stressors. Umbilical cord blood DNAm near *GNAS* was altered among infants born to a mother affected by gestational diabetes (GDM), a disorder characterized by glucose intolerance during pregnancy (51). Based on the Dutch Hunger Winter Families Study (52), siblings exposed to the war-time Dutch Hunger Winter famine were associated with persistent changes in DNAm in a region near the *GNASAS* locus relative to their unexposed siblings (53). The direction and magnitude of effect depended on timing of exposure and sex of the exposed individual (53). DNAm among exposed siblings was also altered near another gene implicated in metabolic disease *MEG3* (53), a gene that mapped to ΔDMR4 which was also characterized by decreasing methylation among progressors and maintainers relative to reverters (Table 2). Interestingly, both the GNAS (54) and MEG3 (55) genes are maternally imprinted. Loss of maternal imprinting should be investigated as a potential mechanism in the etiology of T1D using whole-genome bisulfite sequencing in order to provide a higher density representation of DNAm changes within imprinted areas of the genome.

The secondary eQTM analysis in a subset of overlapping samples confirmed that changes in methylation within ΔDMR1 were associated with expression of GNAS. Increased methylation post- versus pre-SV was associated with higher levels of GNAS expression at the post-SV visit in a subset of overlapping samples. GNAS is an important regulator of insulin secretion in beta cells (56). GNAS silencing results in decreased insulin secretion and insulin content (56). GNAS encodes the G protein subunit alpha which also plays a role in the interaction between antigen presenting cells and T helper cell differentiation (57). Mice with dendritic cells deficient for GNAS result in a phenotype characterized by preferential Th2 differentiation, Th2 type inflammation, and subsequent development of allergic asthma (57). Overlap between autoimmunity and atopic conditions have long been hypothesized based on disruptions in similar immune pathways (58). Positive associations between childhood asthma and subsequent T1D development have been observed in several countries (59–61). Overall, our results suggest that maintenance of DNAm levels

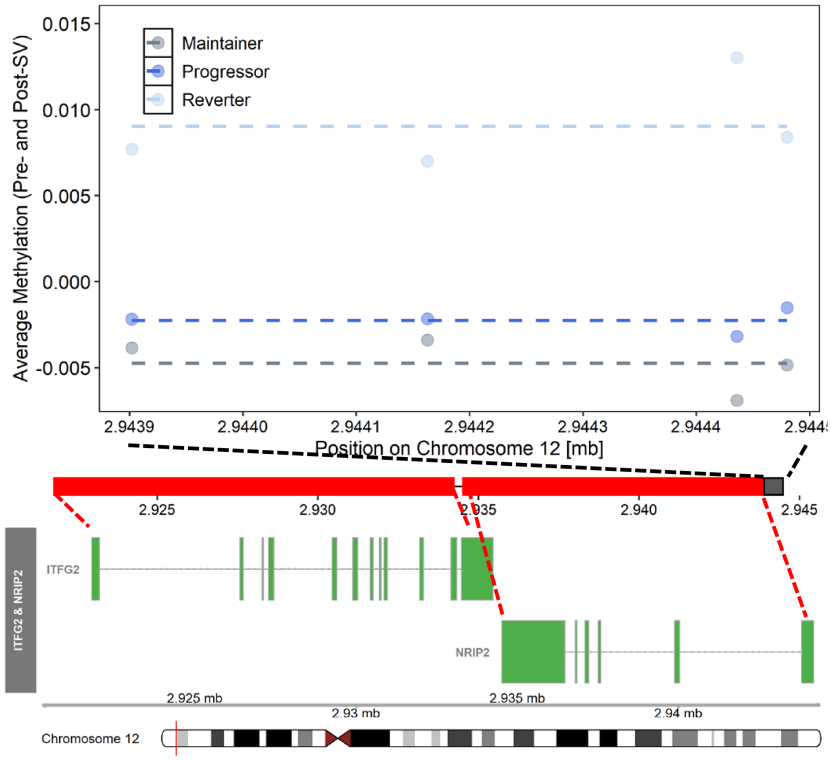


FIGURE 4
Differentially methylated region on chromosome 12 where average (pre- and post-SV) methylation levels differed across the three IA progression phenotypes. Description. Region on chromosome 12 loc 2943902 to 2944481 (µDMR4) where average DNA methylation (DNAm) levels, post- and pre-SV, differed across groups. In the top panel, each dot represents the average DNAm value (y-axis) at each of the CpG sites included µDMR4. The x-axis represents the position (mb) of the CpGs within the region. All DNAm values were adjusted for age, sex, cell proportions, and genetic ancestry. The dashed lines represent the average methylation value within each group across the entire region. The middle panel represents the location of the region (black solid square) relative to the closest genes, *ITFG2* and *NRIP2* (red solid squares). There are multiple known isoforms for *ITFG2* and *NRIP2*, the figure displays the most biologically relevant or consensus transcript based on the Ensembl database. The red line on the ideogram, bottom of the figure, represents the location of *ITFG2* and *NRIP2* on chromosome 12.

near GNAS during IA may represent a unique protective mechanism in reverters.

In order to further characterize the DNAm regions identified in the primary analysis, we tested the correlation between changes in DNAm and changes in annotated metabolites (metQTM). Four differentially changing DMRs were correlated with changes in 26 unique lipid metabolites (Table 5). ΔDMR 8, characterized by increasing methylation in progressors (Figure 5), was correlated with 18 of the 26 lipid metabolites. This region of differentially changing methylation is notable based on its location in an open chromatin region within the body of the *DLL1* gene on chr. 6. As a notch signaling ligand, DLL1 controls the differentiation of pancreatic progenitor cells into exocrine versus endocrine cells (46). The loss of DLL1 results in early progenitor cell differentiation and an

TABLE 4 Summary of FDR significant cis-eQTMs representing correlation between differentially changing methylation regions and gene expression post- SV.

Methylation DMR Information					Cis-Gene Expression Information						
DMR ID	Chr.	DMR Start	DMR Stop	N Probes	Gene Symbol	Ensembl ID	Strand	Gene Start	Gene End	Corr*	FDR
ΔDMR 1	20	57426538	57427974	29	GNAS	ENSG00000087460	1	57414773	57486247	0.559	0.0667
ΔDMR 1	20	57426538	57427974	29	ATP5E	ENSG00000124172	-1	57600522	57607437	0.557	0.0667

DNAm levels for all probes identified in the DMR analysis (Tables 1, 2) were included in a PCA. We then tested the association between the 1st PC and RNA seq data from overlapping visit at the post-SV visit. Only significant cis (TSS +/- 500KB of midpoint of DMR) expression quantitative trait methylation (cis-eQTM) associations are presented.
*Spearman correlation coefficient.
Chr., chromosome.
DMR Start/End, DMR start and end position.
Gene Start/End, Gene start and end positions (based on annotation file for GEO, GSE50244).
Beta, beta coefficient from linear regression model (adjusted for age and sex) representing association between 1st PC from DNAm probes in each DMR and islet cell pancreas expression.
FDR, Benjamini-Hochberg FDR adjusted p value.

TABLE 5 Secondary metQTM analysis of the association between pre- versus post-SV change in methylation across the DMRs and pre- versus post-SV change in metabolite levels.

DMR ID	Chr.	DMR Start	DMR Stop	Metabolite Name†	Standardized Beta	FDR Adj. P Value
ΔDMR 2	chr20	36148604	36149751	TG (49:2)	0.320	0.0992
ΔDMR 5	chr11	1296469	1297386	TG (53:2)	0.411	0.0121
				Phosphatidylcholine (33:1)	0.361	0.0469
				TG (53:3)	0.353	0.0627
				PE (38:4)	0.339	0.0826
				TG (49:2)	0.330	0.0948
				TG (47:0)	0.329	0.0952
				TG (51:3)	0.327	0.0954
				PC (33:1)	0.327	0.0954
				Phosphatidylcholines (35:1)	0.325	0.0954
				TG (53:1)	0.320	0.0992
ΔDMR 8	chr6	170597377	170597899	Phosphatidylcholine (35:4)	0.438	0.0078
				Phosphatidylcholines (33:1)	0.404	0.0121
				Phosphatidylcholines (33:0)	0.403	0.0121
				Phosphatidylcholines (33:1)	0.402	0.0121
				Phosphatidylcholines (35:3)	0.396	0.0138
				LPC (15:0)	0.393	0.0139
				Phosphatidylcholines (38:5)	0.375	0.0527
				Phosphatidylcholines (33:2)	0.366	0.0445
				Phosphatidylcholines (35:4)	0.365	0.0445
				Phosphatidylcholines (31:0)	0.350	0.0647
				Phosphatidylcholines (35:1)	0.347	0.0647
				Phosphatidylcholines (36:3)	0.347	0.0647
				Phosphatidylcholines (p-34:0) or Phosphatidylcholines (o-34:1)	-0.334	0.0940
				TG (49:3)	0.332	0.0940
				Phosphatidylcholines (33:2)	0.332	0.0940
				Phosphatidylcholines (36:3) B	0.325	0.0954
				Phosphatidylcholines (37:6)	0.324	0.0954
				Phosphatidylcholines (35:1)	0.323	0.0975
ΔDMR 9	chr6	28945322	28945493	*Phosphatidylcholine* (p-38:2) or *Phosphatidylcholine*(o-38:3)	0.345	0.0662

DNAm levels for all probes identified in the DMR analysis (Tables 1, 2) were included in a PCA. We then tested the association between the 1st PC changes in metabolites between the pre- and post-SV visits.
Chr., chromosome.
DMR Start/End, DMR start and end position.
Gene Start/End, Gene start and end positions (based on annotation file for GEO, GSE50244).
Standardized Beta, beta coefficient from linear regression model testing the association between change in DNAm and change in metabolites pre-SV vs post-SV. The slopes have been standardized to represent a 1 stdev change in metabolite per 1 standard deviation change in DNAm regional PC levels.
FDR Adj. P value, Benjamini-Hochberg FDR adjusted p value.
†See Appendix 5 (Data Sheet 5) (Tables D, E) for complete annotation for all metabolites included in Table 5.

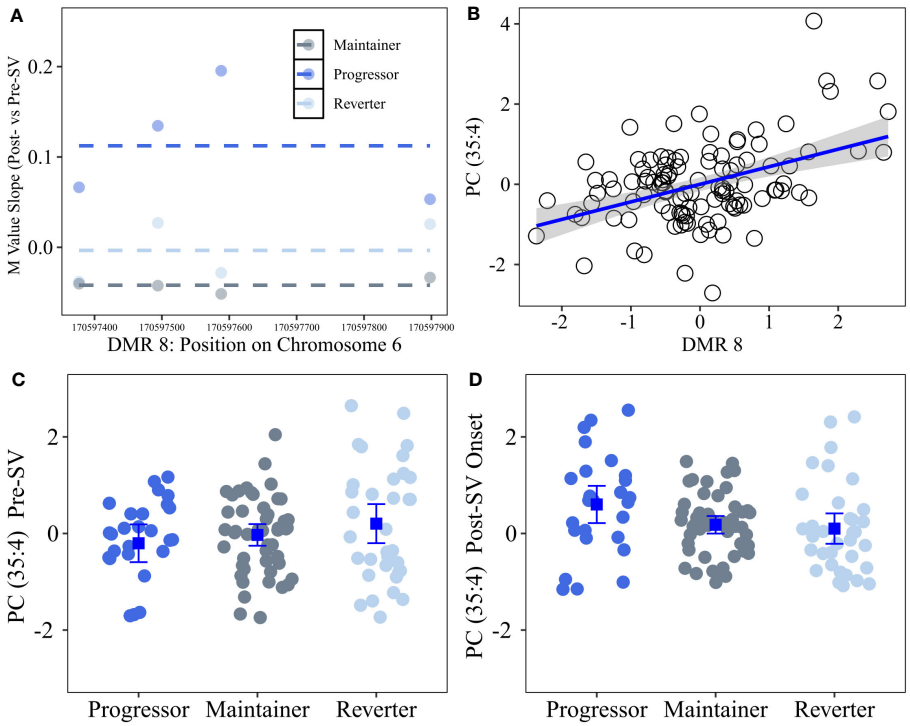


FIGURE 5
Differentially changing region on chromosome 6 (post- vs pre-SV) that was positively correlated with changes in lipid metabolites (post- vs pre-SV). Description: Region on chromosome 6 loc 170597377 to 170597899 (ΔDMR8) where the change in DNA methylation (DNAm) post- vs pre-SV differed across groups. In the top left **(A)**, each dot represents the within group slopes (y-axis) or changes in methylation m-values between the post-SV and pre-SV visit at each of the CpG sites included ΔDMR 8. The x-axis represents the position (mb) of the CpGs within the region. Positive values indicate DNAm values increased following IA seroconversion whereas negative values indicate DNAm decreased following IA seroconversion. The dashed lines represent the average slope value within each group across the entire region. The top right **(B)** represents the association between DMR wide DNAm captured by the 1st PC (x-axis) and changes in metabolite values (y-axis) between the post- and pre-SV visits. DNAm and metabolite expression values have been standardized to facilitate the interpretation of the slope as a 1 standard deviation increase in the change in metabolite levels between the post- and pre-SV visits per 1 standard deviation increase in the change in methylation between post- and pre-SV visits. The bottom panels **(C, D)** represent the average metabolite levels and corresponding 95% confidence intervals within the three groups pre- and post-SV. All DNAm and metabolite values were adjusted for age, sex, and cell proportions.

overabundance of endocrine cells (46). A recent mouse model confirmed DLL1 is also relevant to islet cell function in the mature pancreas based on its high level of expression in beta cells and corresponding role in insulin secretion (47). Furthermore, DLL1 plays an important role in differentiation of B cells and the development of antigen secreting cells; the presence of DLL1 influences AB titer levels and isotype switching (45). Additional work is needed to understand the connection between a concordant increase in lipid levels and DNAm within the *DLL1* gene following seroconversion.

Our secondary metQTM was unique in that DNAm and metabolite levels were available pre- and post-SV in a subset of overlapping samples. This analysis revealed a consistent positive association between increasing lipid metabolite levels, post- vs pre-SV, and increasing DNAm levels across several regions (25 of the 26 unique lipid metabolites were positively correlated with DNAm changes, see Table 5). Numerous studies (62–68) have reported associations between dysregulation in lipid levels and T1D. Although lipid levels have been shown to be influenced by age at sample collection/timing of sample collection relative to onset of IA and type of first appearing autoantibody, prior research suggests lower lipid levels, including sphingomyelins and phosphatidylcholines, are generally associated with increased risk of T1D and/or IA (62–68). In our study, increasing lipid levels, in particular phosphocholines, following the onset of IA were strongly correlated with increasing methylation within ΔDMR8. This region was characterized by increasing methylation within the progressor group. However, as demonstrated in Figure 5, the lipid metabolite most strongly correlated with DNAm changes in this region, *Phosphatidylcholine* (35:4), was lower in the progressor group prior to SV and then subsequently increased following the onset of IA, suggesting higher levels of lipids within the progressor group may be unique to changes that occur following seroconversion.

There was a high prevalence of odd-chain fatty acid (OCFA) containing lipid species among the metabolites correlated with DNAm changes. Recently, there has been increased recognition of OCFA in plasma and their potential biological relevance (69). OCFA levels have been associated with glucose homeostasis, insulin resistance, T2D, and BMI (69, 70). Pfleuger et al (71) observed higher levels of odd-chain triglycerides among autoantibody positive versus negative children in BABYDIAB. This parallels the concordant post-seroconversion increase in OCFA levels and DNAm near the DLL1 gene (ΔDMR 8)

among progressors (Figure 5) in the current study. OCFA have been proposed a marker of dairy intake which has been positively correlated with progression to T1D in prior work in DAISY (72). However, dairy intake contributes modestly to OCFA levels. These lipids primarily originate endogenously from adipocytes as well as from dietary intake of numerous foods including dairy, poultry, and fiber (70, 73, 74). Additional work is needed to understand connections between increasing methylation and increasing OCFA as well as the source of these lipid species.

A major strength of our study was the inclusion of DNAm measurements prior to T1D as well as the multi-omics work used to identify correlations between DNAm and gene expression as well metabolite levels. We measured DNAm before and after SV (ie, the appearance of IA) which builds on prior studies that have included DNAm measures after T1D and/or after IA onset only (14–16). A novel feature of our longitudinal methodology was our group*visit interaction modelling strategy that allowed us to identify changes in DNAm before and after the onset of IA, a critical window in T1D pathogenesis. These within individual effects are essential to understanding the etiology of T1D as they are robust to individual level confounders such as sex, genetic predisposition, and/or family history. Johnson et al (17) also used a longitudinal case-control analysis of T1D cases vs. unaffected controls in DAISY. In contrast, the current study design focused on individuals who developed IA and furthermore, tested for differences in DNAm post- vs pre-SV (group*visit interaction) rather than testing for differences in methylation by age (group*age interaction). Comparing the DMRs identified by this study versus Johnson et al (17), only two regions were located within 1 MB of each other–one on chr 6 ΔDMR 9 (28945322–28945493) in the current study vs chr 6 28973328-28973521 in Johnson et al (17), and one on chr 20 ΔDMR 2 (36148604–36149751) in the current study vs chr 20 36148954-36149232 in Johnson et al (17). Consistent with prior work, ΔDMR 9 and ΔDMR 2 were both associated with differential changes in DNAm in progressors relative to maintainers and/or reverters.

4.1 Limitations

We obtained DNAm from whole blood, which means we were unable to identify cell subtype specific effects. Similarly, our study focused on blood tissue only. DNAm changes within the blood may not reflect DNAm changes within other tissues that contribute to T1D, such as the pancreas. Due to advancements in technology during the study, DNAm was measured on two platforms. Individuals were randomly assigned to the platforms to minimize bias. We looked for cis-eQTMs. Given that it is possible that regions act over larger areas of the genome, we may have missed larger effects that occurred outside of our 500 KB window. Due to the small sample size, the eQTM was underpowered to identify FDR significant DMR vs gene transcript pairs. This limitation may explain lack of concordance between eQTM results and BIOS QTL results (μDMR4). Furthermore, among the two gene transcripts that were correlated with changes in methylation within ΔDMR1, gene expression data were only available at the post-SV visit. Therefore, it was not possible to determine whether gene expression also changed pre- versus post-SV. Finally, metabolite levels are influenced by age and dietary patterns. Although we adjusted for age, the large differences in age between the progressor group and the reverter and maintainer groups creates challenges in interpreting the metabolite vs methylation correlations. Additional work is need to replicate the metabolite vs DNAm regional effects.

5 Conclusion

T1D is an autoimmune disease characterized by immune mediated destruction of beta cells. Beta cell stress has been proposed as a mechanism connecting environmental perturbations such as infection, inflammation, diet, and increased insulin secretion to disease progression (75). Our EWAS identified DNAm candidates known to be modified by diabetes relevant environmental factors including diet and glucose levels (*CLIP2*, *GNAS/GNAS-AS*, *MEG3*). Our results also implicated genes (*DLL1* and *GNAS*) with functional roles in both beta and immune cells. Our results build upon prior work by identifying specific areas of the genome where DNAm changes pre- and post-SV visits differentiated between reversion versus progression of IA. The correlation between changes in DNAm and changes in lipid levels reveal common connections between DMRs in different areas of the genome that may be related to disruptions in lipid metabolic pathways. Additional work is needed to replicate these findings, test for cell-specific changes in DNAm pre- vs post-seroconversion, and to identify modifiable factors that lead to these DNAm changes; ideally, the first step in the development of preventative strategies that delay or prevent progression of IA.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA597238.

Ethics statement

The studies involving humans were approved by Colorado Multi-institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

PC: Writing – review & editing, Writing – original draft, Project administration, Methodology, Formal analysis, Conceptualization. LV: Writing – review & editing, Software, Methodology, Data curation. RJ: Writing – review & editing, Methodology, Data curation. TB: Writing – review & editing. AS: Writing – review & editing, Supervision. KK: Writing – review & editing, Supervision, Methodology. IY: Writing –

review & editing, Supervision. TF: Writing – review & editing, Supervision. OF: Writing – review & editing, Resources, Data curation. MR: Writing – review & editing, Supervision, Resources, Funding acquisition. JN: Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by National Institutes R01-DK104351, R01-DK32493, R21-AI142483, and P30-DK116073.

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1345494/full#supplementary-material

References

1. Ziegler AG, Rewers M, Simell O, Simell T, Lempainen J, Steck A, et al. Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. *JAMA*. (2013) 309:2473–9. doi: 10.1001/jama.2013.6285

2. Vehik K, Lynch KF, Schatz DA, Akolkar B, Hagopian W, Rewers M, et al. Reversion of beta-cell autoimmunity changes risk of type 1 diabetes: TEDDY study. *Diabetes Care*. (2016) 39:1535–42. doi: 10.2337/dc16-0181

3. Spencer KM, Tarn A, Dean BM, Lister J, Bottazzo GF. Fluctuating islet-cell autoimmunity in unaffected relatives of patients with insulin-dependent diabetes. *Lancet*. (1984) 1:764–6. doi: 10.1016/S0140-6736(84)91278-9

4. Yu J, Yu L, Bugawan TL, Erlich HA, Barriga K, Hoffman M, et al. Transient antiislet autoantibodies: infrequent occurrence and lack of association with "genetic" risk factors. *J Clin Endocrinol Metab*. (2000) 85:2421–8. doi: 10.1210/jc.85.7.2421

5. Kulmala P, Rahko J, Savola K, Vahasalo P, Veijola R, Sjoroos M, et al. Stability of autoantibodies and their relation to genetic and metabolic markers of Type I diabetes in initially unaffected schoolchildren. *Diabetologia*. (2000) 43:457–64. doi: 10.1007/s001250051329

6. Kimpimaki T, Kulmala P, Savola K, Kupila A, Korhonen S, Simell T, et al. Natural history of beta-cell autoimmunity in young children with increased genetic susceptibility to type 1 diabetes recruited from the general population. *J Clin Endocrinol Metab*. (2002) 87:4572–9. doi: 10.1210/jc.2002-020018

7. Robertson CC, Inshaw JRJ, Onengut-Gumuscu S, Chen WM, Santa Cruz DF, Yang H, et al. Fine-mapping, trans-ancestral and genomic analyses identify causal variants, cells, genes and drug targets for type 1 diabetes. *Nat Genet*. (2021) 53:962–71. doi: 10.1038/s41588-021-00880-5

8. Jerram ST, Leslie RD. The genetic architecture of type 1 diabetes. *Genes (Basel)*. (2017) 8. doi: 10.3390/genes8080209

9. Patterson CC, Dahlquist GG, Gyurus E, Green A, Soltesz G, Group ES. Incidence trends for childhood type 1 diabetes in Europe during 1989-2003 and predicted new cases 2005-20: a multicentre prospective registration study. *Lancet*. (2009) 373:2027–33. doi: 10.1016/S0140-6736(09)60568-7

10. Ziegler AG, Pflueger M, Winkler C, Achenbach P, Akolkar B, Krischer JP, et al. Accelerated progression from islet autoimmunity to diabetes is causing the escalating incidence of type 1 diabetes in young children. *J Autoimmun*. (2011) 37:3–7. doi: 10.1016/j.jaut.2011.02.004

11. Lappalainen T, Greally JM. Associating cellular epigenetic models with human phenotypes. *Nat Rev Genet*. (2017) 18:441–51. doi: 10.1038/nrg.2017.32

12. Sawalha AH. Epigenetics and T-cell immunity. *Autoimmunity*. (2008) 41:245–52. doi: 10.1080/08916930802024145

13. Allan RS, Nutt SL. Deciphering the epigenetic code of T lymphocytes. *Immunol Rev*. (2014) 261:50–61. doi: 10.1111/imr.12207

14. Rakyan VK, Beyan H, Down TA, Hawa MI, Maslau S, Aden D, et al. Identification of type 1 diabetes-associated DNA methylation variable positions that precede disease diagnosis. *PloS Genet*. (2011) 7:e1002300. doi: 10.1371/journal.pgen.1002300

15. Paul DS, Teschendorff AE, Dang MA, Lowe R, Hawa MI, Ecker S, et al. Increased DNA methylation variability in type 1 diabetes across three immune effector cell types. *Nat Commun*. (2016) 7:13555. doi: 10.1038/ncomms13555

16. Elboudwarej E, Cole M, Briggs FB, Fouts A, Fain PR, Quach H, et al. Hypomethylation within gene promoter regions and type 1 diabetes in discordant monozygotic twins. *J Autoimmun*. (2016) 68:23–9. doi: 10.1016/j.jaut.2015.12.003

17. Johnson RK, Vanderlinden LA, Dong F, Carry PM, Seifert J, Waugh K, et al. Longitudinal DNA methylation differences precede type 1 diabetes. *Sci Rep*. (2020) 10:3721. doi: 10.1038/s41598-020-60758-0

18. Yu L, Robles DT, Abiru N, Kaur P, Rewers M, Kelemen K, et al. Early expression of antiinsulin autoantibodies of humans and the NOD mouse: evidence for early determination of subsequent diabetes. *Proc Natl Acad Sci USA*. (2000) 97:1701–6. doi: 10.1073/pnas.040556697

19. Yu L, Rewers M, Gianani R, Kawasaki E, Zhang Y, Verge C, et al. Antiislet autoantibodies usually develop sequentially rather than simultaneously. *J Clin Endocrinol Metab*. (1996) 81:4264–7. doi: 10.1210/jcem.81.12.8954025

20. Bonifacio E, Yu L, Williams AK, Eisenbarth GS, Bingley PJ, Marcovina SM, et al. Harmonization of glutamic acid decarboxylase and islet antigen-2 autoantibody assays for national institute of diabetes and digestive and kidney diseases consortia. *J Clin Endocrinol Metab*. (2010) 95:3360–7. doi: 10.1210/jc.2010-0293

21. Wenzlau JM, Juhl K, Yu L, Moua O, Sarkar SA, Gottlieb P, et al. The cation efflux transporter ZnT8 (Slc30A8) is a major autoantigen in human type 1 diabetes. *Proc Natl Acad Sci USA*. (2007) 104:17040–5. doi: 10.1073/pnas.0705894104

22. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. (2014) 30:1363–9. doi: 10.1093/bioinformatics/btu049

23. Vanderlinden LA, Johnson RK, Carry PM, Dong F, DeMeo DL, Yang IV, et al. An effective processing pipeline for harmonizing DNA methylation data from Illumina's 450K and EPIC platforms for epidemiological studies. *BMC Res Notes*. (2021) 14:352. doi: 10.1186/s13104-021-05741-2

24. Carry PM, Waugh K, Vanderlinden LA, Johnson RK, Buckner T, Rewers M, et al. Changes in the co-expression of innate immunity genes during persistent islet autoimmunity are associated with progression of islet autoimmunity: the diabetes autoimmunity study in the young (DAISY). *Diabetes*. (2022) 71(9):2048–57. doi: 10.2337/figshare.20060480

25. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf*. (2011) 12:323. doi: 10.1186/1471-2105-12-323

26. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. (2014) 15:550. doi: 10.1186/s13059-014-0550-8

27. Risso D, Ngai J, Speed TP, Dudoit S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat Biotechnol*. (2014) 32:896–902. doi: 10.1038/nbt.2931

28. Carry PM, Vanderlinden LA, Johnson RK, Buckner T, Fiehn O, Steck AK, et al. Phospholipid levels at seroconversion are associated with resolution of persistent islet autoimmunity: the diabetes autoimmunity study in the young. *Diabetes*. (2021) 70:1592–601. doi: 10.2337/db20-1251

29. Showalter MR, Nonnecke EB, Linderholm AL, et al. Obesogenic diets alter metabolism in mice. *PLoS One*. (2018) 13:e0190632.

30. Fiehn O, Wohlgemuth G, Scholz M, et al. Quality control for plant metabolomics: reporting MSI-compliant studies. *Plant J*. (2008) 53:691–704.

31. Cajka T, Smilowitz JT, Fiehn O. Validating quantitative untargeted lipidomics across nine liquid chromatography-high-resolution mass spectrometry platforms. *Anal Chem*. (2017) 89:12360–80.

32. Skogerson K, Wohlgemuth G, Barupal DK, Fiehn O. The volatile compound BinBase mass spectral database. *BMC Bioinformatics*. (2011) 12:321.

33. Tsugawa H, Cajka T, Kind T, et al. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat Methods*. (2015) 12:523–6.

34. Kind T, Liu KH, Lee DY, DeFelice B, Meissen JK, Fiehn O. LipidBlast in silico tandem mass spectrometry database for lipid identification. *Nat Methods*. (2013) 10:755–8.

35. Fan S, Kind T, Cajka T, et al. Systematic error removal using random forest for normalizing large-scale untargeted lipidomics data. *Anal Chem*. (2019) 91:3590–6.

36. Buckner T, Johnson RK, Vanderlinden LA, Carry PM, Romero A, Onengut-Gumuscu S, et al. Genome-wide analysis of oxylipins and oxylipin profiles in a pediatric population. *Front Nutr*. (2023) 10:1040993. doi: 10.3389/fnut.2023.1040993

37. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B (Methodological)*. (1995) 57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

38. Pedersen BS, Schwartz DA, Yang IV, Kechris KJ. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. *Bioinformatics*. (2012) 28:2986–8. doi: 10.1093/bioinformatics/bts545

39. Dreja T, Jovanovic Z, Rasche A, Kluge R, Herwig R, Tung YC, et al. Diet-induced gene expression of isolated pancreatic islets from a polygenic mouse model of the metabolic syndrome. *Diabetologia*. (2010) 53:309–20. doi: 10.1007/s00125-009-1576-4

40. Diabetes Genetics Initiative of Broad Institute of H, Mit LU and Novartis Institutes of BioMedical R, , Saxena R, Voight BF, Lyssenko V, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*. (2007) 316:1331–6. doi: 10.1126/science.1142358

41. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science*. (2007) 316:1336–41. doi: 10.1126/science.1142364

42. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature*. (2012) 489:57–74. doi: 10.1038/nature11247

43. Alliance of Genome Resources C. Alliance of Genome Resources Portal: unified model organism research platform. *Nucleic Acids Res*. (2020) 48:D650–D8. doi: 10.1093/nar/gkz813

44. Kim W, Shin YK, Kim BJ, Egan JM. Notch signaling in pancreatic endocrine cell and diabetes. *Biochem Biophys Res Commun*. (2010) 392:247–51. doi: 10.1016/j.bbrc.2009.12.115

45. Garis M, Garrett-Sinha LA. Notch signaling in B cell immune responses. *Front Immunol*. (2020) 11:609324. doi: 10.3389/fimmu.2020.609324

46. Apelqvist A, Li H, Sommer L, Beatus P, Anderson DJ, Honjo T, et al. Notch signalling controls pancreatic cell differentiation. *Nature*. (1999) 400:877–81. doi: 10.1038/23716

47. Rubey M, Chhabra NF, Gradinger D, Sanz-Moreno A, Lickert H, Przemeck GKH, et al. DLL1- and DLL4-mediated notch signaling is essential for adult pancreatic islet homeostasis. *Diabetes*. (2020) 69:915–26. doi: 10.2337/db19-0795

48. Dror V, Nguyen V, Walia P, Kalynyak TB, Hill JA, Johnson JD. Notch signalling suppresses apoptosis in adult human and mouse pancreatic islet cells. *Diabetologia*. (2007) 50:2504–15. doi: 10.1007/s00125-007-0835-5

49. Al-Shami A, Crisostomo J, Wilkins C, Xu N, Humphries J, Chang WC, et al. Integrin-alpha FG-GAP repeat-containing protein 2 is critical for normal B cell differentiation and controls disease development in a lupus model. *J Immunol*. (2013) 191:3789–98. doi: 10.4049/jimmunol.1203534

50. Bonder MJ, Luijk R, Zhernakova DV, Moed M, Deelen P, Vermaat M, et al. Disease variants alter transcription factor levels and methylation of their binding sites. *Nat Genet*. (2017) 49:131–8. doi: 10.1038/ng.3721

51. Chen D, Zhang A, Fang M, Fang R, Ge J, Jiang Y, et al. Increased methylation at differentially methylated region of GNAS in infants born to gestational diabetes. *BMC Med Genet*. (2014) 15:108. doi: 10.1186/s12881-014-0108-3

52. Lumey LH, Stein AD, Kahn HS, van der Pal-de Bruin KM, Blauw GJ, Zybert PA, et al. Cohort profile: the Dutch Hunger Winter families study. *Int J Epidemiol*. (2007) 36:1196–204. doi: 10.1093/ije/dym126

53. Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD, et al. DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum Mol Genet*. (2009) 18:4046–53. doi: 10.1093/hmg/ddp353

54. Weinstein LS, Xie T, Zhang QH, Chen M. Studies of the regulation and function of the Gs alpha gene Gnas using gene targeting technology. *Pharmacol Ther*. (2007) 115:271–91. doi: 10.1016/j.pharmthera.2007.03.013

55. Miyoshi N, Wagatsuma H, Wakana S, Shiroishi T, Nomura M, Aisaka K, et al. Identification of an imprinted gene, Meg3/Gtl2 and its human homologue MEG3, first mapped on mouse distal chromosome 12 and human chromosome 14q. *Genes Cells*. (2000) 5:211–20. doi: 10.1046/j.1365-2443.2000.00320.x

56. Taneera J, Dhaiban S, Mohammed AK, Mukhopadhyay D, Aljaibeji H, Sulaiman N, et al. GNAS gene is an important regulator of insulin secretory capacity in pancreatic beta-cells. *Gene*. (2019) 715:144028. doi: 10.1016/j.gene.2019.144028

57. Lee J, Kim TH, Murray F, Li X, Choi SS, Broide DH, et al. Cyclic AMP concentrations in dendritic cells induce and regulate Th2 immunity and allergic asthma. *Proc Natl Acad Sci USA*. (2015) 112:1529–34. doi: 10.1073/pnas.1417972112

58. Shah A. The pathologic and clinical intersection of atopic and autoimmune disease. *Curr Allergy Asthma Rep*. (2012) 12:520–9. doi: 10.1007/s11882-012-0293-0

59. Smew AI, Lundholm C, Savendahl L, Lichtenstein P, Almqvist C. Familial coaggregation of asthma and type 1 diabetes in children. *JAMA Netw Open*. (2020) 3:e200834. doi: 10.1001/jamanetworkopen.2020.0834

60. Metsala J, Lundqvist A, Virta LJ, Kaila M, Gissler M, Virtanen SM, et al. The association between asthma and type 1 diabetes: a paediatric case-cohort study in Finland, years 1981-2009. *Int J Epidemiol*. (2018) 47:409–16. doi: 10.1093/ije/dyx245

61. Yun HD, Knoebel E, Fenta Y, Gabriel SE, Leibson CL, Loftus EV Jr., et al. Asthma and proinflammatory conditions: a population-based retrospective matched cohort study. *Mayo Clin Proc*. (2012) 87:953–60. doi: 10.1016/j.mayocp.2012.05.020

62. La Torre D, Seppanen-Laakso T, Larsson HE, Hyotylainen T, Ivarsson SA, Lernmark A, et al. Decreased cord-blood phospholipids in young age-at-onset type 1 diabetes. *Diabetes*. (2013) 62:3951–6. doi: 10.2337/db13-0215

63. Lamichhane S, Ahonen L, Dyrlund TS, Kemppainen E, Siljander H, Hyoty H, et al. Dynamics of plasma lipidome in progression to islet autoimmunity and type 1 diabetes - type 1 diabetes prediction and prevention study (DIPP). *Sci Rep*. (2018) 8:10635. doi: 10.1038/s41598-018-28907-8

64. Oresic M, Gopalacharyulu P, Mykkanen J, Lietzen N, Makinen M, Nygren H, et al. Cord serum lipidome in prediction of islet autoimmunity and type 1 diabetes. *Diabetes*. (2013) 62:3268–74. doi: 10.2337/db13-0159

65. Li Q, Parikh H, Butterworth MD, Lernmark A, Hagopian W, Rewers M, et al. Longitudinal metabolome-wide signals prior to the appearance of a first islet autoantibody in children participating in the TEDDY study. *Diabetes*. (2020) 69:465–76. doi: 10.2337/db19-0756

66. Johnson RK, Vanderlinden L, DeFelice BC, Kechris K, Uusitalo U, Fiehn O, et al. Metabolite-related dietary patterns and the development of islet autoimmunity. *Sci Rep*. (2019) 9:14819.

67. Oresic M, Simell S, Sysi-Aho M, Nanto-Salonen K, Seppanen-Laakso T, Parikka V, et al. Dysregulation of lipid and amino acid metabolism precedes islet autoimmunity in children who later progress to type 1 diabetes. *J Exp Med*. (2008) 205:2975–84. doi: 10.1084/jem.20081800

68. Lamichhane S, Ahonen L, Dyrlund TS, Dickens AM, Siljander H, Hyoty H, et al. Cord-blood lipidome in progression to islet autoimmunity and type 1 diabetes. *Biomolecules*. (2019) 9. doi: 10.3390/biom9010033

69. Huynh K, Barlow CK, Jayawardana KS, Weir JM, Mellett NA, Cinel M, et al. High-throughput plasma lipidomics: detailed mapping of the associations with cardiometabolic risk factors. *Cell Chem Biol*. (2019) 26:71–84 e4. doi: 10.1016/j.chembiol.2018.10.008

70. Prada M, Wittenbecher C, Eichelmann F, Wernitz A, Drouin-Chartier JP, Schulze MB. Association of the odd-chain fatty acid content in lipid groups with type 2 diabetes risk: A targeted analysis of lipidomics data in the EPIC-Potsdam cohort. *Clin Nutr*. (2021) 40:4988–99. doi: 10.1016/j.clnu.2021.06.006

71. Pflueger M, Seppanen-Laakso T, Suortti T, Hyotylainen T, Achenbach P, Bonifacio E, et al. Age- and islet autoimmunity-associated differences in amino acid and lipid metabolites in children at risk for type 1 diabetes. *Diabetes*. (2011) 60:2740–7. doi: 10.2337/db10-1652

72. Lamb MM, Miller M, Seifert JA, Frederiksen B, Kroehl M, Rewers M, et al. The effect of childhood cow's milk intake and HLA-DR genotype on risk of islet autoimmunity and type 1 diabetes: the Diabetes Autoimmunity Study in the Young. *Pediatr Diabetes*. (2015) 16:31–8. doi: 10.1111/pedi.2015.16.issue-1

73. Green CR, Wallace M, Divakaruni AS, Phillips SA, Murphy AN, Ciaraldi TP, et al. Branched-chain amino acid catabolism fuels adipocyte differentiation and lipogenesis. *Nat Chem Biol*. (2016) 12:15–21. doi: 10.1038/nchembio.1961

74. Dąbrowski G, Konopka I. Update on food sources and biological activity of odd-chain, branched and cyclic fatty acids—-A review. *Trends Food Sci Technol*. (2021) 119:514–29. doi: 10.1016/j.tifs.2021.12.019

75. Rewers M, Ludvigsson J. Environmental risk factors for type 1 diabetes. *Lancet*. (2016) 387:2340–8. doi: 10.1016/S0140-6736(16)30507-4

frontiers | Frontiers in Immunology

TYPE Review
PUBLISHED 10 September 2024
DOI 10.3389/fimmu.2024.1456145

OPEN ACCESS

EDITED BY
Frederick Miller,
National Institute of Environmental Health Sciences (NIH), United States

REVIEWED BY
Shepherd Schurman,
National Institutes of Health (NIH), United States
Kenneth Michael Pollard,
The Scripps Research Institute, United States
Adam Schiffenbauer,
National Institutes of Health (NIH), United States

*CORRESPONDENCE
May Y. Choi
may.choi@ucalgary.ca

RECEIVED 28 June 2024
ACCEPTED 16 August 2024
PUBLISHED 10 September 2024

CITATION
Choi MY, Costenbader KH and Fritzler MJ (2024) Environment and systemic autoimmune rheumatic diseases: an overview and future directions.
Front. Immunol. 15:1456145.
doi: 10.3389/fimmu.2024.1456145

Environment and systemic autoimmune rheumatic diseases: an overview and future directions

May Y. Choi[1,2*], Karen H. Costenbader[3,4] and Marvin J. Fritzler[1]

[1]Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada, [2]McCaig Institute for Bone and Joint Health, Calgary, AB, Canada, [3]Department of Medicine, Div of Rheumatology, Inflammation and Immunity, Brigham and Women's Hospital, Boston, MA, United States, [4]Medicine, Harvard Medical School, Boston, MA, United States

Introduction: Despite progress in our understanding of disease pathogenesis for systemic autoimmune rheumatic diseases (SARD), these diseases are still associated with high morbidity, disability, and mortality. Much of the strongest evidence to date implicating environmental factors in the development of autoimmunity has been based on well-established, large, longitudinal prospective cohort studies.

Methods: Herein, we review the current state of knowledge on known environmental factors associated with the development of SARD and potential areas for future research.

Results: The risk attributable to any particular environmental factor ranges from 10-200%, but exposures are likely synergistic in altering the immune system in a complex interplay of epigenetics, hormonal factors, and the microbiome leading to systemic inflammation and eventual organ damage. To reduce or forestall the progression of autoimmunity, a better understanding of disease pathogenesis is still needed.

Abbreviations: aHR, adjusted hazards ratio; AI, artificial intelligence; ANA, antinuclear antibody; BWHS, Black Women's Health Study; CI, confidence interval; COVID-19; coronavirus disease 2019; BlyS, B-cell lymphocyte stimulator; CCP, cyclic citrullinated peptide; CRP, C-reactive protein; dsDNA, anti-double-stranded DNA; DNAm, DNA methylation; EBV, Epstein-Barr virus; GRS, genetic risk score; HCQ, hydroxychloroquine; HLA, human lymphocyte antigen; HR, hazard ratio; ML, machine learning; IFN, interferon; IL, interleukin; IIM, idiopathic inflammatory myopathies; IU, international units; NHS, Nurses' Health Study; OR, odds ratio; NHSII, PTSD, post-traumatic stress disorder; RA, rheumatoid arthritis; rRNA, ribosomal RNA; SARD, systemic autoimmune rheumatic diseases; SARS-CoV2, severe acute respiratory syndrome coronavirus 2; SjD, Sjögren disease; SLE, systemic lupus erythematosus; SSc, systemic sclerosis; TNF, tumor necrosis factor; UV, ultraviolet; VH3 BCR, VH3 B Cell Repertoire.

Conclusion: Owing to the complexity and multifactorial nature of autoimmune disease, machine learning, a type of artificial intelligence, is increasingly utilized as an approach to analyzing large datasets. Future studies that identify patients who are at high risk of developing autoimmune diseases for prevention trials are needed.

KEYWORDS

autoimmunity, autoimmune diseases, environment, autoantibodies, epigenetics, microbiome, machine learning, artificial intelligence

Introduction

Environmental factors operating on the background of hormonal factors and genetic vulnerability may be accelerating factors included in a long-held paradigm that helps explain the etiology of systemic autoimmune rheumatic disease (SARD), including systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), systemic sclerosis (SSc), Sjögren's disease (SjD), idiopathic inflammatory myopathies (IIM) and others (1). On the backdrop of an increasing prevalence of SARD and other autoimmune diseases (2–6), potential accelerating factors include several environmental and socioeconomic factors that include alterations of foods, increasing exposure to xenobiotics due to water and air pollution, heat and other extreme weather events (i.e., climate change), biodiversity loss, ultraviolet (UV) light exposure, pandemics and infections, and socioeconomic factors such as changes in personal lifestyles and psychological stress.

Extensive research over the past three to four decades has elucidated the environmental factors associated with SLE (7) and other SARD. In general, the environmental factors can be classified as airborne, waterborne, workplace/occupational, social, and behavioral (8). While it has not been possible to identify a universal environmental "pathogen" for all SARD, there is compelling evidence that some environmental exposures clearly serve as risk factors for disease onset. The central importance of identifying these factors is that many of these factors are actionable and modifiable through intervention and remediation. Expanding the use of machine learning (ML), a form of artificial intelligence (AI), to analyze large datasets including environmental exposures may lead to the identification of other modifiable environmental risk factors, and allow the development of new disease-specific remediation programs (2).

Environmental factors and autoimmunity

The development of SARD has been associated with several lifestyle behaviors. For instance, cigarette smoke (9–11), obesity (12), alcohol use (moderate consumption being protective) (10, 13–15), poor nutrition and intake of ultra-processed foods (16), psychosocial factors (e.g., major depression (17), sleep deprivation (18), child abuse, personal trauma, post-traumatic stress disorder [PTSD]) (19, 20), and reproductive factors (21–23) have been associated with SLE development. Environmental exposures such as air pollution (24), occupational hazards (25), residential proximity to hazardous waste sites or pesticide exposure (26, 27), UV light (28–33), vitamin D deficiency (34), and exposure to viruses (35, 36) have also been linked to increased SLE risk. Similar lifestyle factors have been reported for increased risk of developing RA (moderate alcohol consumption decreases RA risk), SSc, IIM, other SARD, and autoinflammatory conditions (Tables 1, 2).

Precisely how and the extent to which these lifestyle factors contribute to individual risk of autoimmune disease likely varies (57, 58). This has been particularly well-studied using large cohort studies including cohorts enrolled in the Nurses' Health Study (NHS) and Black Women's Health Study (BWHS). In SLE, each factor independently increases the risk of disease development by 10-200%, but they likely interact with each other and with genetic risk, potentially synergistically, to accelerate brewing autoimmunity in SLE [reviewed in (57–60)]. Using SLE as an example below, we discuss several potential biologic pathways involving epigenomics, the microbiome, and immune dysregulation that lead to inflammation and organ damage, mechanisms that may also apply to the development of other SARD (Figure 1).

Common pathways of pathogenesis: immune dysregulation, epigenomics, the microbiome

Immune dysregulation

Inflammation is an adaptive response to stressors that involves multiple physiological processes that include the innate and adaptive immune systems. In turn, inflammation regulates – and is regulated by – several highly interconnected systems including the epigenome and microbiome (64). Unhealthy lifestyle behaviors (i.e., smoking, sedentary lifestyle, and consumption of ultra-

10.3389/fimmu.2024.1456145

TABLE 1 Environmental factors that increase risk for systemic autoimmune rheumatic diseases.

Lifestyle Exposure	Disease Association	Reported Risk from Select Key References (Citation)
Air Pollution	RA	• HR 1.31 (95%CI: 0.98–1.74) living near traffic pollution (road) vs. not (24)
	SLE	• Increases in air pollutants nitrogen dioxide (NO_2), carbon monoxide (CO), and fine particles ($PM_{2.5}$) (HR 1.21 [95% CI: 1.08–1.36], HR 1.44 [95% CI: 1.31–1.59], and HR 1.12 [95% CI: 1.02–1.23], respectively) (37)
	SARD[1]	• OR 1.13 (95%CI: 1.02-1.25) for lowest vs. highest satellite fine particulate air pollution level (38)
Cigarette Smoke	RA	• RR 3.8 (95%CI: 2.0-6.9) in current smokers vs. never smokers (39) • OR 1.65 (95%CI: 1.03–2.64) for >20 versus 0 pack-years) for anti-CCP-positive RA (40)
	SLE	• OR 1.50 (95%CI: 1.09–2.08) for current smokers compared with non-smokers (11) • HR 1.86 (95%CI: 1.14–3.04) for current vs. never smokers for dsDNA+ SLE risk (9)
Diet	SLE	• Women in the highest tertile of cumulatively updated dietary ultra-processed food (UPF) intake/day were at almost 50% greater risk of developing SLE vs. women in the lowest tertile of UPF daily intake (16)
Hazardous Waste Sites	SLE	• Exposure to volatile organic compounds ($P < 0.05$) (26)
Obesity	RA	• History of obesity (OR 1.24 [95%CI: 1.01–1.53]) (41)
	SLE	• An 85% (HR 1.85 [95%CI: 1.17-2.91]) significantly increased risk of SLE among obese compared to normal BMI women in the more recent NHSII cohort (12), but not NHS
Organic Solvents, Pesticides and Heavy Metal	RA	• Application of chemical fertilizers (adjusted OR 1.7 [95%CI: 1.1-2.7]) and cleaning with solvents (OR 1.6 [95%CI: 1.1-2.4]) (42)
	SLE	• Pesticide exposure (adjusted OR 2.24 [95%CI: 1.28–3.93]) (27) • Association with SLE risk seen with mercury (OR 3.6 [95%CI: 1.3-10.0]) and mixing pesticides for agricultural work (OR 7.4 [95%CI: 1.4-40.0]) (43)
	SSc	• OR 2.9 (95%CI: 1.1-7.6) for solvent organic solvent exposure (male SSc vs controls) (44)
Periodontitis	RA	• OR 1.16 (95%CI: 1.13-1.21) history of periodontitis (45)
Psychosocial	SLE	• Probable PTSD (HR 2.94 [95%CI: 1.19–7.26]) and trauma exposure (HR 2.83 [95%CI: 1.29–6.21]) (19) • Women with a history of depression vs. no depression (HR 2.67 [95%CI: 1.91-3.75]) (17) • Adverse childhood experiences (abuse, neglect, and household challenges) associated with increased risk of SLE. Exposure to the highest vs. lowest physical and emotional abuse was associated with 2.57 times greater risk of SLE (95%CI: 1.30–5.12) (46). HR for ≥2 episodes of severe sexual abuse compared to no abuse was 2.51 (95%CI: 1.29–4.85) and ≥5 episodes of severe physical abuse was 2.37 (95%CI: 1.13–4.99) among Black women) (20).
Reproductive/ Hormonal Factors	SLE	• Pooled RR 1.5 (95%CI: 1.1-2.1) oral contraceptive use and use of postmenopausal hormones RR 1.9 (95%CI: 1.2-3.1) (21)
Silica	RA	• Silica exposed men OR 2.2 (95%CI: 1.2-3.9) among men aged 18 to 70 years and 2.7 (95%CI: 1.2-5.8) among those aged 50 to 70 years (47)
	SLE	• Medium silica exposure was OR 2.1 (95%CI: 1.1–4.0), high exposure OR 4.6 (95%CI: 1.4–15.4) (25)
	Vasculitis	• Overall significant summary effect estimate of silica "ever exposure" with development of AAV (OR 2.56 (95%CI: 1.51–4.36) (48)
	SSc	• The combined estimator of relative risk for studies in females was 1.03 (95%CI: 0.74–1.44) and was 3.02 (95%CI: 1.24–7.35) for males (49).
Sleep Deprivation	SLE	• HR 2.47 (95%CI: 1.29-4.75) for chronic low sleep duration (≤5 hours/night versus >7–8 hours) (18)
UV Radiation	SLE	• History of more than one serious sunburn before the age of 20 years (OR 2.2, 95%CI: 1.2–4.1) and sunburn-susceptible skin type (OR 2.9, 95%CI: 1.6–5.1) (32)
Viruses	SLE	• Epstein-Barr virus serologic reactivation among unaffected SLE relatives (viral capsid antigen IgG OR 1.28 [95%CI: 1.07-1.53], p=0.007 and early antigen IgG OR 1.43 [95%CI: 1.06-1.93], p=0.02) (36)
	SARD	• Higher risk of RA (adjusted HR (aHR) 2.98 [95%CI: 2.78–3.20]), SLE (aHR 2.99 [95%CI: 2.68–3.34]), dermatopolymyositis (aHR 1.96 [95%CI: 1.47–2.61]), SSc (aHR 2.58 [95%CI: 2.02–3.28]), SjD (aHR 2.62 [95%CI: 2.29–3.00]), mixed connective tissue disease (aHR 3.14 [95%CI: 2.26–4.36]), Behçet's disease (aHR 2.32 [95%CI: 1.38–3.89]), polymyalgia rheumatica (aHR 2.90 [95%CI: 2.36–3.57]), and vasculitis (aHR 1.96 [95%CI: 1.74–2.20]) among COVID-19 vs. non-COVID-19 exposed unvaccinated individuals (50).

AAV, anti-neutrophil cytoplasmic autoantibody (ANCA)-associated vasculitis; CI, confidence interval; CCP, cyclic citrullinated peptide; HR, hazard ratio; NHSII, Nurses' Health Study Cohort 2; OR, odds ratio; RA, rheumatoid arthritis; RR, relative risk; SARD, systemic autoimmune rheumatic diseases; SjD, Sjögren disease; SLE, systemic lupus erythematosus; SSc, systemic sclerosis; UV, ultraviolet.
1. SARD included systemic lupus erythematosus, Sjögren's disease, scleroderma, polymyositis, dermatomyositis, or undifferentiated connective tissue disease.

TABLE 2 Environmental factors that decrease risk for systemic autoimmune rheumatic diseases.

Lifestyle Exposure	Disease Association	Reported Risk from Select Key References (Citation)
Alcohol	RA	• HR 0.78 (95%CI: 0.61–1.00) for alcohol use of 5.0–9.9 gm/day (51)
	SLE	• HR 0.65 [95%CI: 0.45–0.96] among women who drank 2 or more servings of wine had significantly decreased SLE risk compared to women who did not drink wine (13)
Diet	RA	• HR 0.67 (95%CI: 0.51-0.88) among women aged ≤55 years, better quality diet was associated with lower RA risk, particularly seropositive RA (52)
Exercise	SLE	• Regular exercise (performing at least 19 metabolic equivalent hours of exercise per week) assessed with other healthy behaviors (never or past smoker, healthy diet, moderate alcohol consumption, healthy body weight) was associated with a 19% reduction in SLE risk per additional healthy behavior, such that women with four or more healthy lifestyle factors had the lowest risk (HR 0.42 [95%CI: 0.25-0.70]) (53).
	RA	• Similar to the SLE study above, a lower risk of RA was also observed with a healthier lifestyle including regular exercise, i.e., women with five healthy lifestyle factors had the lowest risk (HR 0.42 [95% CI: 0.22-0.80]) (54).
Reproductive/ Hormonal Factors	RA	• RR 0.8 (95%CI: 0.6–1.0) for breastfeeding for 2–23 total months (55)
Vitamin D	SARD[1]	• Vitamin D 2000IU daily supplementation was associated with a 22% reduction in the development of autoimmune disease (HR 0.78 [95% CI: 0.61, 0.99], P=0.05) (56).

CI, confidence interval; HR, hazard ratio; RA, rheumatoid arthritis; RR, relative risk; SARD, systemic autoimmune rheumatic diseases; SLE, systemic lupus erythematosus; 1. This included RA, polymyalgia rheumatic, autoimmune thyroid disease, psoriasis, inflammatory bowel disease, and many others (e.g., SLE, systemic sclerosis).

processed foods) promote systemic inflammation leading to chronic inflammatory diseases, including SARD. Before developing overt clinical manifestations, individuals developing SARD have a period of asymptomatic autoimmunity and inflammation of variable intensity and duration, characterized by increasing oxidative stress, loss of immune tolerance, autoantibody formation, immune complex deposition and complement activation, epigenetic modifications, and upregulation and/or downregulation of cytokine expression [reviewed in (65)].

In SLE, both obesity and exposure to the toxic components of cigarette smoke induce oxidative stress (66). This, in turn, raises intracellular levels of reactive oxygen species that damage DNA producing immunogenic DNA adducts that can lead to the production of 'pathogenic' anti-double-stranded DNA antibodies (dsDNA) (67–69). In the NHS and NHSII cohorts, smokers were at higher risk of developing anti-dsDNA positive SLE compared to never-smokers (hazard ratio [HR] 1.86 [95% confidence interval (CI): 1.14-13.04]), while there were no significant associations between smoking status or pack-years and overall SLE or anti-dsDNA negative SLE (9). In addition to elevated oxidative stress, the byproducts of smoking could also augment native autoreactive B cells (11) and induce pulmonary antinuclear antibody (ANA) as demonstrated in the lungs of exposed mice (70). Smoking may also influence specific genes in the pathogenesis of SLE (57). An individual with a high SLE genetic risk score or GRS (score based on 86 single-nucleotide polymorphisms and 10 classic HLA alleles previously associated with SLE) and a status of current/recent smoking was strongly associated with SLE risk (odds ratio [OR] 1.5, p=0.0003 versus more distant past/never smoking) and even stronger in the presence of anti-dsDNA antibodies. Not surprisingly, smoking also affects circulating cytokines. Elevated SARD-related cytokines including the B-cell lymphocyte stimulator (BlyS) (70), tumor necrosis factor-alpha (TNF-α), and interleukin (IL)-6 (71, 72), but lower IL-10 (an anti-inflammatory cytokine) have been detected in smokers (73). These cytokines affect the function of T cells and $CD4^+$ regulatory T cells, which are important in maintaining self-tolerance. Similarly, adipose tissue, in particular visceral fat, secretes pro-inflammatory adipocyte-derived cytokines and exhibits higher levels of C-reactive protein (CRP), TNF-α receptor 2, and IL-6 than non-obese individuals (74).

The association between SLE risk and diet is less clear in humans (75–77) compared to other autoimmune diseases such as RA [reviewed in (78)]. There is evidence from SLE-prone mice models that low dietary fiber intake and a Western-type diet (i.e., high in sugar, fat, refined grains, and red meat) are associated with increased autoantibody production (79, 80). In the BWHS, a diet high in carbohydrates and low in fats was associated with an increased risk of developing SLE in African American women (HR 1.88 [95%CI: 1.06-3.35]) (75). Consumption of ultra-processed foods, in particular sugar and artificially sweetened beverages, has been associated with an increased risk of developing SLE among women (16). Low to moderate alcohol consumption (approximately 1/2 drink a day), on the other hand, has been shown to reduce the risk of SLE development among women (10, 13–15). Alcohol (e.g., ethanol) and antioxidants may counteract the changes induced by smoking and obesity, i.e., inhibiting key enzymes in DNA synthesis and suppressing TNF-α, IL-6, IL-8, and interferon (IFN)-γ that lower systemic inflammation (81, 82).

Several studies have reported an association between lack of sleep and SLE risk in humans (18, 83, 84). In the NHS and NHSII cohorts, chronic low sleep duration (</=5 hours/night versus the recommended >7-8 hours) was associated with increased SLE risk (adjusted HR 2.47 [95% CI: 1.29, 4.75]), with stronger effects among those with body pain and depression. In sleep-deprived individuals, increased levels of IL-6 and TNF-α have been reported (85–89).

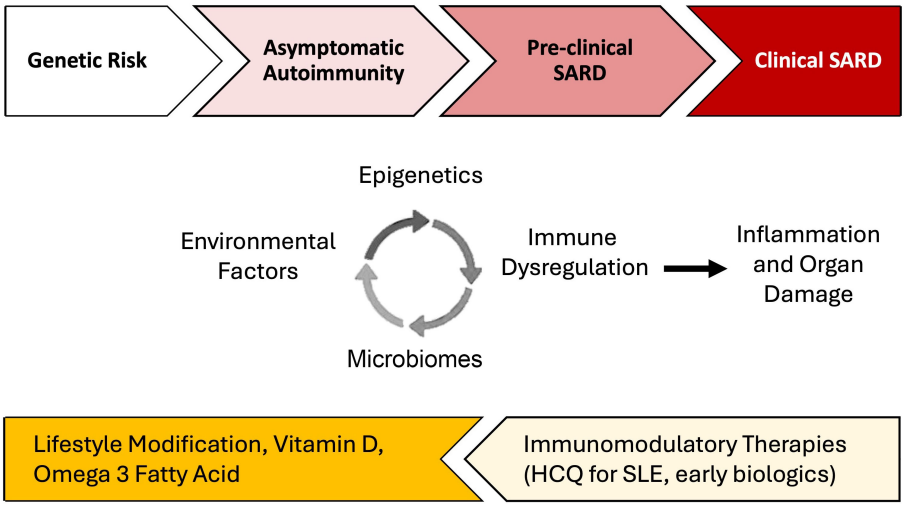


FIGURE 1
Environmental factor-associated pathogenesis and personalized preventative vs. treatment interventions for systemic autoimmune rheumatic diseases (SARD). Among individuals genetically predisposed to SARD development, unhealthy lifestyle behaviors and other environmental factors can trigger dysregulation in the microbiome, epigenetic changes, and immune dysregulation which, together, drive inflammation. In turn, inflammation can drive further derangements in the microbiome, cause distinct epigenetic changes, and lead to additional immune dysregulation. During the periods of asymptomatic autoimmunity and pre-clinical SARD, this positive feedback leads to a process wherein inflammation becomes chronic and self-sustaining, ultimately driving autoimmunity and eventually leading to organ damage and clinical disease. Effective lifestyle interventions, supplementation, and early introduction of immunomodulatory therapies may help prevent disease progression. There may be a potential role for treatments such as hydroxychloroquine for pre-SLE [SMILE trial underway (61)] and Abatacept, a T-cell co-stimulation inhibitor, for pre-RA (62, 63).

In SLE-prone mice, sleep deprivation was associated with accelerated production of autoantibodies and earlier disease onset (90). Sleep disturbances arising in individuals who have had childhood or adult trauma, PTSD, or occupational stress from working night or rotating shifts, may also explain why these factors have also been linked to SLE onset (17, 19, 20, 43, 91, 92). In the NHSII, PTSD, a condition arising after exposure to trauma and marked by severe psychological stress, was associated with increased SLE risk (HR 2.94 [95% CI: 1.19-7.26], p<0.05) compared to women with no trauma, even after adjusting for other SLE risk factors smoking, body mass index (BMI), and oral contraceptive use (19). In the NHS and NHSII, women with a history of depression had a higher risk of SLE (HR 2.67 [95:CI: 1.91-3.75] p<0.001) compared to women with no depression (17). Systemic inflammation, denoted by elevated TNF, IL-6, and CRP levels, has been repeatedly reported in individuals with emotional stress and distress (91, 93–102).

There is also evidence that sex hormones are important in SLE development (21, 22), a disease, like some other SARD, that predominantly affects females. In SLE, a population-based nested case-control study using the UK's General Practice Research Database demonstrated that there was a dose-response in oral contraceptive pill (ethinyl estradiol) and SLE risk (adjusted rate ratio [aRR] 1.42, 1.63, and 2.92 for < or =30 microgram, 31-49 microgram, and 50 microgram, respectively) (22). They also reported that the rate was particularly increased among females who recently started taking oral contraceptive pills (aRR 2.52 [95% CI: 1.14-5.57]) compared with longer-term current users. Estrogen prevents B cell receptor-mediated apoptosis and upregulates several genes that contribute to B cell activation and survival (cd22, shp-1, bcl-2, and vcam-1) (103).

Chemical and physical exposures have also been historically linked to SLE onset, including crystalline silica dust (25, 33, 104, 105), heavy metals such as mercury (43), air pollution and other respiratory particulates (38, 106), residential proximity to hazardous waste sites (26), agricultural pesticides (27, 43, 107), and organic solvents (42, 44). Proposed mechanisms of pathogenesis include stimulation of cellular necrosis and release of intracellular antigens resulting in systemic inflammation and IFN upregulation. These environmental exposures have also been described as important risk factors in the development of RA (42), SSc (44), and vasculitis (48). A comprehensive review of the literature (~1980-2010) on environmental factors and SARD development concluded that among these chemical factors, crystalline silica exposure, solvent exposure, and smoking had the strongest level of evidence (108). Since then, however, multiple studies have been published. The evidence for metal exposure and SARD development including mercury at that time was felt to be insufficient, although there is renewed interest in mercury-induced autoimmunity in more recent studies (109, 110). Mercury exposure has been associated with autoimmune features that are more consistent with pre- or sub-clinical autoimmunity in humans, and in animal studies, acts independently of type I IFN to induce milder disease (111).

UVB radiation can exacerbate pre-existing SLE, however, whether it contributes to SLE disease onset or pathogenesis is less

clear. While UVB radiation can up-regulate Th2 cells and down-regulate Th1 cells, induce IL-10 production, increase type I IFN expression, and prolong T cell activation to increase SLE risk (29–31), another subset of UV radiation, UVA, is used as a phototherapy modality to treat cutaneous forms of lupus (112). UVB also has an important role in vitamin D3 synthesis in the skin, which has been hypothesized to *lower* SLE risk (28, 113). Vitamin D deficiency is reportedly common among SLE patients (34) and is important in the regulatory pathways of numerous genes involved in inflammation and immunity including IL-2 inhibition, antibody production, and lymphocyte proliferation (114, 115). We will later discuss a large, randomized, double-blind, placebo-controlled clinical trial called the vitamin D and omega 3 trial (VITAL) trial, where vitamin D 2000 IU daily supplementation was associated with a 22% reduction in the development of autoimmune disease (HR 0.78 [95% CI: 0.61, 0.99], p=0.05) (56).

Viral triggers, particularly Epstein-Barr Virus (EBV), have also been associated with SLE development (35). In a recent study of 436 unaffected SLE patient relatives who were followed for 6.3 ± 3.9 years and evaluated for interim transitioning to SLE, increased serological reactivation of EBV was associated with higher risk of transitioning to SLE (viral capsid antigen IgG OR 1.28 [95%CI: 1.07-1.53], p=0.007 and expression of EBV early antigen IgG (OR 1.43 [95%CI: 1.06-1.93], p=0.02) (36). Proposed mechanisms include molecular mimicry and the release of EBV-encoded small RNAs from infected cells resulting in the induction of type-1 IFN and proinflammatory cytokines via activating toll-like receptor (TLR)-3 signaling (116). The interest in triggering of autoimmune conditions by viral infections was renewed during the coronavirus disease 2019 (COVID-19) pandemic when there were outbreaks of pediatric inflammatory multisystemic syndrome [PIMS also referred to as multisystem inflammatory syndrome in children (MIS-C)] that reportedly followed severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in children. These reports included cases of Kawasaki-like disease, Kawasaki disease shock syndrome, toxic shock syndrome, myocarditis and macrophage activation syndrome (117–119). In adults, SARS-CoV-2 infection has also been linked to a higher risk of developing a diverse spectrum of new-onset autoimmune diseases as highlighted by two large retrospective studies (50, 120). Chang et al. used data from the TriNetX network and propensity score matching (two cohorts [COVID-19 and non-COVID-19] of 887,455 SARS-CoV-2 unvaccinated individuals) to identify the incidence of autoimmune conditions during the study period (1 January 2020 to 31 December 2021) (50). Unlike EBV, there was a wider spectrum of SARD seen including higher risk of RA (adjusted hazard ratio (aHR) 2.98 [95%CI: 2.78–3.20]), SLE (aHR 2.99 [95%CI: 2.68–3.34]), dermato/polymyositis (aHR 1.96 [95%CI: 1.47–2.61]), SSc (aHR 2.58 [95%CI: 2.02–3.28]), SjD (aHR 2.62 [95% CI: 2.29–3.00]), and other autoimmune diseases. Future studies that elucidate how viruses, such as SARS-CoV-2, increase the risk of SARD development may help implement preventive measures and early treatment in individuals who have had these infections to prevent morbidity and mortality.

A key pathway involved in both anti-viral response and the pathogenesis of SLE and other SARD including IIM and SSc is the type I IFN pathway (121). Approximately 50-70% of adult and pediatric SLE patients have an upregulated IFN signature, a cluster of IFN-stimulated genes, that correlates with disease activity and severity (122). A recent study demonstrated that type-1 IFN inhibits the aryl hydrocarbon receptor (AHR) pathway. Suppressed AHR signaling promotes T cell production of CXC ligand 13 (CXCL13), a chemokine that regulates B cell recruitment and lymphoid aggregation in inflamed tissues (123). AHR is important for sensing changes in the cellular milieu provided by the environment, diet, commensal flora, and host metabolism (124). In response to these environmental ligands, AHR has a protective role against inflammation by downregulating pro-inflammatory pathways (124). In the gut, AHR is expressed in epithelial cells and immune cells in the lamina propria to also stabilize the gut epithelial barrier (124). In the central nervous system, AHR is upregulated in astrocytes and microglia in response to ligands that cross the blood-brain barrier (124). Lower AHR expression has been described as a potential mechanism of pathogenesis for several autoimmune conditions including inflammatory bowel disease (125), multiple sclerosis (126), and psoriasis (127). In SLE, deficits in the AHR-driven immunoregulation exacerbated by the type-1 IFN may explain how alterations in the environment lead to the development of autoimmunity and uncontrolled inflammation. Moreover, polycyclic aromatic hydrocarbons, smoking, air pollution, and other environmental exposures cause DNA methylation changes in the AHR repressor genes, potentially linking these exposures to the development of autoimmunity (128–130). Future studies are warranted to elucidate the pathways by which regulation of the AHR pathway is related to lymphocyte activation status in the pathogenesis of autoimmunity.

Epigenetic changes

The currently accepted etiologic model for SARD implicates an interaction of inherited genetic factors and environmental exposures over time. DNA methylation (DNAm), an epigenetic change controlling gene expression, is influenced by both genetics and environmental exposures and therefore, may provide a critical link between them [reviewed in (131–133)]. For instance, UV light exposure, infections, silica, heavy metals and pesticide exposures, cigarette smoking, and air pollution are all thought to inhibit DNAm by oxidative stress, which could promote SARD onset specifically or non-specifically (134). In addition to DNAm, cigarette smoking is linked to the activation of enzymes that regulate other types of epigenetic modifications (i.e., post-translational modifications of histones via methylation, acetylation, phosphorylation, ubiquitination, and regulation of non-coding RNA sequences) to mediate the expression of multiple inflammatory genes, thereby participating in the onset development of autoinflammatory diseases (135).

DNAm occurs when a methyl group is added to a cytosine base in a cytosine-phosphate-guanine dinucleotide (CpG) which, in general, silences nearby gene expression. By comparison, demethylation activates gene expression. These changes, mainly

demethylation and in particular IFN gene hypomethylation, have been observed in various cell subsets, including CD4 T cells in patients affected by SLE (136–145). Upregulation of type I IFN in SLE is thought to induce an "IFN epigenomic signature", activating latent enhancers and "bookmarking" chromatin, reprogramming genes to be hyper-responsive, amplifying the inflammatory cascade (146–148). Emerging data reveal that some of these epigenetic changes are correlated with SLE disease manifestations (malar and discoid rash, dsDNA autoantibodies, lupus nephritis) and disease severity (137, 139, 144, 149), and are highly specific to SLE such that they distinguish individuals with existing SLE from controls and other SARD (141, 150). Well-designed epidemiologic studies are still needed to determine whether other epigenetic changes precede the development of SARD and whether such changes could be modified to abrogate disease.

Microbiome influences

There is mounting evidence that imbalances in the microbiota contribute to metabolic and immune regulatory dysfunction, which may contribute to the pathogenesis of chronic inflammatory diseases such as SARD [reviewed in (151)]. Several independent reported studies of 16S rRNA libraries have identified characteristic patterns of gut dysbiosis in SLE, in which there is an inverse relationship between disease activity and overall biodiversity of the intestinal microbiota (152–154). In studies of 61 female SLE patients, there was an eight-fold increase in *Ruminococcus gnavus* abundance compared to the healthy subjects, and most patients with high R. *gnavus* abundance had active nephritis (152). Increases in *R. gnavus* abundance have also been observed in other diseases including allergies and spondyloarthropathies with inflammatory bowel disease (155–157). Importantly, many strains of *R. gnavus* express a VH3 B cell repertoire (BCR) targeted B cell superantigen, particularly relevant to SLE given the importance of B cell activation in disease pathogenesis (158).

Evidence suggests that SLE patients may suffer chronic microbial translocation through impaired gut barrier integrity contributing to immunologic dysregulation (159). Oral microbiome studies confirm that SLE patients have a distinct microbiome signature compared to healthy controls, with evidence of translocation of bacteria, e.g., *Veillonella* species, from the oral cavity to the intestine (160, 161).

In healthy adults, the microbiome, even at the level of strains, is relatively stable over many years (162). However, the microbiome can be altered by diet, sleep, exercise, stress, medications (antibiotics and non-antibiotics), and the environment (163). Perturbations in the gut microbiome composition have been suggested to trigger SLE onset or disease flares and *vice versa* (164). In-depth studies examining the impact of lifestyle and environmental factors on changes to the microbiome and subsequent risk of autoimmune diseases are needed.

Other host barriers should also be considered as potential targets for prevention including the oral cavity and lung mucosa as these have been identified as sites of pathogenic autoreactive immune responses that contribute to autoimmune disease. The initiation of RA by inflammation characterized by an aberrant Th-17-dominated immune response, neutrophil activation, antigen citrullination, and anti-cyclic citrullinated peptide (CCP) production is exacerbated by microbial dysbiosis, the presence of oral pathobionts (e.g., *Porphyromonas gingivalis* and *Aggregatibacter actinomycetemcomitans*), and periodontitis has been described (45, 165–167).

The lung mucosa is another site of protein citrullination leading to RA development, promoted by microbial infection or dysbiosis and the inhalation of pollutants such as tobacco smoke or other pollutants (168, 169). This anti-CCP production and translocation into the systemic circulation has been proposed to accelerate the development of RA with interstitial lung disease for individuals who are genetically predisposed (e.g., gain-of-function MUC5B promoter variant reducing mucociliary function in small airways responsible for clearing inhaled particles in the lungs (170)). It is difficult to be certain that microbiome alterations observed in recent studies of SARD patients are not due to established and treated disease. Additional studies of the microbiome before disease onset are warranted.

Mitigation of environmental factors

Traditional cohort studies

Our current understanding of lifestyle factors and autoimmune diseases has largely depended on large observational epidemiological studies (53, 54, 171). Many of these studies used self-reported data including the use of validated and standardized questionnaires. These studies also relied on the retention of subjects in the long term to enable repeated measurement of lifestyle behaviors. Nevertheless, these studies have filled important knowledge gaps in our understanding of the link between environmental exposures and autoimmunity.

In the NHS and NHSII cohorts, our group demonstrated that adherence to multiple healthy behaviors (healthy diet (highest 40th percentile of the Alternative Healthy Eating Index), regular exercise (performing at least 19 metabolic equivalent hours of exercise per week), never or past smoker, moderate alcohol consumption (drinking ≥5 gm/day alcohol), and maintaining a healthy body weight (body mass index <25 kg/m2) was associated with a 19% reduction in SLE risk per additional healthy behavior, such that women with four or more healthy lifestyle factors had the lowest risk (HR 0.42 [95%CI: 0.25-0.70]) (53). An even greater reduction per healthy behavior (22%) was observed for the risk of anti-dsDNA-positive SLE. Overall, the population-attributable risk, or the proportion of the risk in this population that could be attributed to these five modifiable lifestyle risk factors was 47.7% [95%CI: 23.1-66.6%]. Using the same cohorts and similar modeling, a lower risk of RA was also observed with a healthier lifestyle, i.e., women with five healthy lifestyle factors had the lowest risk (HR 0.42 [95%

CI: 0.22-0.80]) (54). Therefore, a significant proportion of the risks of both SLE and RA may be preventable by adhering to healthy lifestyles.

Intervention and prevention trials

There is a scarcity of clinical trials examining lifestyle and environmental interventions and prevention strategies to reduce the risk of autoimmune disease development. One of the challenges in designing a strong and well-powered prevention study is identifying which at-risk individuals to study. Our group has previously developed SLE risk prediction models having 76% accuracy by combining family history, genetic factors, and lifestyle, medical and behavioral exposures that classify a woman's risk of SLE in the next two years (172). There is also a rapidly growing panel of potential biomarkers of SLE risk or early disease including anti-dense fine speckled 70 (DFS70) as a rule-out SARD test (173), anti-C1q antibodies as a rule-in test (174), cytokines and chemokines (175, 176), IFN signature (177), as well as markers of complement activation (178). Therefore, identifying individuals for screening, risk-stratifying, assessing biomarkers, and testing intervention and prevention strategies before clinical disease onset has recently become possible (65, 179).

In a pivotal randomized, double-blind, placebo-controlled vitamin D and omega 3 trial (VITAL) trial with a two-by-two factorial design (n=25 871 participants followed for a median of 5.3 years), vitamin D (2000IU/day) supplementation for five years [with or without omega 3 fatty acid (1000 mg/day)] had a significant reduction in the risk of confirmed autoimmune disease of 22% (HR 0.78 [95% CI: 0.61, 0.99], p=0.05) (56). This included RA, polymyalgia rheumatica, autoimmune thyroid disease, psoriasis, inflammatory bowel disease, and others (e.g., SLE, SSc). Individuals who received an omega-3 fatty acid supplementation (with or without vitamin D supplementation) had a reduced rate of incident autoimmune disease by 15% but this was not statistically significant. However, the two-year post-intervention observation study where participants were no longer provided with any supplements but were invited to continue being observed while off assigned supplements, demonstrated that the protective effects of the 5.3 years of randomized exposure to 2000 IU/day of vitamin D dissipated, but the randomized supplementation with 1,000 mg/day of omega-3 fatty acids for the 5.3 years was seen to have a sustained effect in reducing autoimmune disease incidence (180). The results suggest that vitamin D supplementation of 2000 IU/day should be given continuously for long-term prevention of autoimmune disease, while the beneficial effects of omega-3 fatty acids may be more sustained.

The only SLE-specific prevention trial to date is the "Study of Anti-Malarials in Incomplete Lupus Erythematosus (SMILE)" (61), which was set to determine whether SLE progression can be abrogated by using hydroxychloroquine (HCQ) among patients with a positive ANA test and at least one (but not three or more) additional clinical or laboratory criterion from the 2012 Systemic Lupus Inception Collaborating Clinics (SLICC) classification criteria (181). This highly anticipated, multicenter, randomized, double-blind, placebo-controlled, 24-month trial is expected to be completed soon.

A similar HCQ prevention trial in RA ("Strategy to Prevent the Onset of Clinically-Apparent Rheumatoid Arthritis" or STOP-RA) was halted early due to the futility of the treatment (182). In the interim analysis it was observed that in individuals who were anti-CCP positive but without inflammatory arthritis at baseline, one year of HCQ was not superior to placebo in preventing or delaying the development of inflammatory arthritis, and the classification of individuals as having RA at 3 years (probabilities of RA development were 34% in the HCQ arm and 36% in the placebo; p=0.844). Therefore, in RA, HCQ did not help prevent or delay the onset of clinical disease compared to placebo. The study did suggest however that anti-CCP at levels of ≥40 units will be an important enrolment criterion in future RA prevention studies. Therefore, as we strive towards a future of prevention over cure in any SARD, a better and more standardized approach to identifying the timing of intervention and which patients are at the highest risk is urgently needed to ensure the success of prevention trials.

Other RA prevention trials such as the "TREAT Early Arthralgia to Reverse or Limit Impending Exacerbation to Rheumatoid arthritis" (TREAT EARLIER) trial examining one year of methotrexate also did not meet its endpoint of development of clinical arthritis among individuals with arthralgia clinically suspected of progressing to RA and magnetic resonance imaging (MRI)-detected subclinical joint inflammation (183). The T-cell co-stimulation inhibitor abatacept has shown greater promise in delaying RA development in two different at-risk populations. In the "Abatacept inhibits inflammation and onset of rheumatoid arthritis in individuals at high risk" or ARIAA trial, abatacept treatment for six months among RA-at-risk individuals (anti-CCP positive and showing MRI signs of inflammation) reduced subclinical joint inflammation and delays the development of RA (62). In the "Arthritis Prevention In the Pre-clinical Phase of RA with Abatacept" (APIPPRA) trial, at-risk individuals were defined as individuals with arthralgia, anti-CCP plus rheumatoid factor (RF) positive or high anti-CCP titers ≥3 x upper limit of normal plus RF negative, without synovitis at baseline (63). In this randomized, double-blind, multicenter, parallel, placebo-controlled, phase 2b clinical trial, 52 weeks of abatacept treatment reduced RA development over two years compared to placebo. However, by 24 months, the effect of abatacept treatment on symptom burden and subclinical inflammation as determined by ultrasound was not sustained. Therefore, longer treatment with abatacept beyond 12 months might be required. These studies again highlight the need for criteria that identify at-risk individuals from patients with early RA and the most appropriate time to target preventative interventions (184).

Future technologies for research on environmental exposures and SARD

In the last decade, there has been an exponential uptake of AI technologies to study diseases including SARD [reviewed in (185–

187)]. Much of this is due to greater access to a variety of data sources, e.g., images, efficient data collection tools, and supercomputer and analytic methods to rapidly compute. ML is a type of AI that refers to utilizing computers to perform specific tasks by learning from the data rather than being explicitly programmed with instructions such as traditional statistical tests. Within ML, different algorithms are generally categorized into supervised, unsupervised, reinforcement, and deep learning.

In the study of SARD, ML has proven useful in developing prediction models for diagnosis and disease outcomes and in elucidating pathogenesis [reviewed in (185)]. As SARD are highly complex, multifactorial, and heterogeneous diseases, ML is an ideal approach because it can reveal patterns and interactions between variables in large and complex datasets more accurately and efficiently than traditional statistical methods. As we enter an era of 'multi-omics', information on our patients is becoming increasingly 'layered' and challenging to interpret and ML holds promise for new insights and interpretations.

Utilizing ML, we recently demonstrated that there are four unique SLE clusters defined by longitudinal autoantibody profiles alone (188). While these clusters are predictive of disease activity, treatment requirements, complications, and mortality, it also points to autoantibodies as being a fundamental underlying mechanism of immune dysregulation and disease pathogenesis of SLE. This approach can be adopted to study pathogenesis for other SARD and inform more personalized monitoring and treatment plans. The focus of current SLE ML models is on the identification of patients with established disease or the prediction of specific SLE manifestations, e.g., nephritis, neuropsychiatric disease. This includes a validated diagnostic algorithm called the SLE Risk Probability Index (SLERPI) where a SLERPI score of greater than 7 was highly accurate (94.2%) and sensitive for detecting early disease (93.8%) and severe manifestations including kidney (97.9%) and neuropsychiatric involvement (91.8%) (189). Future studies to develop ML models that predict the development of new-onset SLE utilizing datasets that include environmental exposures are needed.

Conclusions

Our examination of risk and protective factors for SARD development, including adherence to multiple healthy lifestyle behaviors, has helped our understanding of the pathogenesis of autoimmunity that involves immune dysregulation, epigenetics, and an altered microbiome. Multiple environmental exposures, including social and behavioral factors throughout our lifespan are likely synergistic and interactive with each other and with genetic factors, influencing the immune system in a complex interplay of epigenetic, hormonal, and microbiome influences, leading to systemic inflammation and eventual organ damage in some. While a major focus has been placed on identifying new targets for disease treatment, shifting the care paradigm to disease prevention is an attractive proposition, especially as our ability to identify high-risk individuals improves. In the few prevention trials that have been conducted, the importance of identifying patients at the highest risk and the likelihood of benefiting from preventative treatment has been highlighted, and thus far, biomarkers have played a critical role in risk stratification. Given the complexity and vast clinical heterogeneity of SARD, ML approaches will become increasingly relied upon to study SARD pathogenesis and prevention.

Author contributions

MC: Conceptualization, Formal analysis, Investigation, Validation, Visualization, Writing – original draft, Writing – review & editing. KC: Conceptualization, Investigation, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing. MF: Conceptualization, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. KC is supported by NIH K24 AR066109 and R01 AR057327. MC is supported by CIHR, McCaig Institute for Bone and Joint Health, Lupus Foundation of America, Lupus Canada, and the Arthritis Society.

Conflict of interest

MC has received consulting fees from AstraZeneca, Mallinckrodt Pharmaceuticals, MitogenDx, Werfen, Celltrion, Organon, and GlaxoSmithKline. MF is the Medical Director of Mitogen Diagnostics Corp and a consultant to Werfen.

The remaining author declares 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

References

1. Pollard KM. Environment, autoantibodies, and autoimmunity. *Front Immunol.* (2015) 6:60. doi: 10.3389/fimmu.2015.00060

2. Miller FW. The increasing prevalence of autoimmunity and autoimmune diseases: an urgent call to action for improved understanding, diagnosis, treatment, and prevention. *Curr Opin Immunol.* (2023) 80:102266. doi: 10.1016/j.coi.2022.102266

3. Heidari H, Lawrence DA. Climate stressors and physiological dysregulations: mechanistic connections to pathologies. *Int J Environ Res Public Health.* (2023) 21. doi: 10.3390/ijerph21010028

4. Skevaki C, Nadeau KC, Rothenberg ME, Alahmad B, Mmbaga BT, Masenga GG, et al. Impact of climate change on immune responses and barrier defense. *J Allergy Clin Immunol.* (2024) 153(5):1194–205. doi: 10.1016/j.jaci.2024.01.016

5. Lee AS, Aguilera J, Efobi JA, Jung YS, Seastedt H, Shah MM, et al. Climate change and public health: The effects of global warming on the risk of allergies and autoimmune diseases: The effects of global warming on the risk of allergies and autoimmune diseases. *EMBO Rep.* (2023) 24:e56821. doi: 10.15252/embr.202356821

6. Pollard KM, Christy JM, Cauvi DM, Kono DH. Environmental xenobiotic exposure and autoimmunity. *Curr Opin Toxicol.* (2018) 10:15–22. doi: 10.1016/j.cotox.2017.11.009

7. Cardelli C, Zucchi D, Elefante E, Signorini V, Menchini M, Stagnaro C, et al. Environment and systemic lupus erythematosus. *Clin Exp Rheumatol.* (2024) 42:1104–14. doi: 10.55563/clinexprheumatol/17vmqc

8. Silman AJ, Hochberg MC. Occupational and environmental influences on scleroderma. *Rheum Dis Clin North Am.* (1996) 22:737–49. doi: 10.1016/S0889-857X(05)70298-2

9. Barbhaiya M, Tedeschi SK, Lu B, Malspeis S, Kreps D, Sparks JA, et al. Cigarette smoking and the risk of systemic lupus erythematosus, overall and by anti-double stranded DNA antibody subtype, in the Nurses' Health Study cohorts. *Ann Rheum Dis.* (2018) 77:196–202. doi: 10.1136/annrheumdis-2017-211675

10. Cozier YC, Barbhaiya M, Castro-Webb N, Conte C, Tedeschi SK, Leatherwood C, et al. Relationship of cigarette smoking and alcohol consumption to incidence of systemic lupus erythematosus in a prospective cohort study of black women. *Arthritis Care Res (Hoboken).* (2019) 71:671–7. doi: 10.1002/acr.23703

11. Costenbader KH, Kim DJ, Peerzada J, Lockman S, Nobles-Knight D, Petri M, et al. Cigarette smoking and the risk of systemic lupus erythematosus: a meta-analysis. *Arthritis Rheumatism.* (2004) 50:849–57. doi: 10.1002/art.20049

12. Tedeschi SK, Barbhaiya M, Malspeis S, Lu B, Sparks JA, Karlson EW, et al. Obesity and the risk of systemic lupus erythematosus among women in the Nurses' Health Studies. *Semin Arthritis Rheumatol.* (2017) 47:376–83. doi: 10.1016/j.semarthrit.2017.05.011

13. Barbhaiya M, Lu B, Sparks JA, Malspeis S, Chang SC, Karlson EW, et al. Influence of alcohol consumption on the risk of systemic lupus erythematosus among women in the nurses' Health study cohorts. *Arthritis Care Res (Hoboken).* (2017) 69:384–92. doi: 10.1002/acr.22945

14. Bengtsson AA, Rylander L, Hagmar L, Nived O, Sturfelt G. Risk factors for developing systemic lupus erythematosus: a case-control study in southern Sweden. *Rheumatol (Oxford).* (2002) 41:563–71. doi: 10.1093/rheumatology/41.5.563

15. Wang J, Pan HF, Ye DQ, Su H, Li XP. Moderate alcohol drinking might be protective for systemic lupus erythematosus: a systematic review and meta-analysis. *Clin Rheumatol.* (2008) 27:1557–63. doi: 10.1007/s10067-008-1004-z

16. Rossato S, Oakes EG, Barbhaiya M, Sparks JA, Malspeis S, Willett WC, et al. Ultra-processed food intake and risk of systemic lupus erythematosus among women followed in the nurses' health study cohorts. *Arthritis Care Res.* (2024), 1–11. doi: 10.1002/acr.25395

17. Roberts AL, Kubzansky LD, Malspeis S, Feldman CH, Costenbader KH. Association of depression with risk of incident systemic lupus erythematosus in women assessed across 2 decades. *JAMA Psychiatry.* (2018) 75:1225–33. doi: 10.1001/jamapsychiatry.2018.2462

18. Choi MY, Malspeis S, Sparks JA, Cui J, Yoshida K, Costenbader KH. Association of sleep deprivation and the risk of developing systemic lupus erythematosus among women. *Arthritis Care Res (Hoboken).* (2023) 75:1206–12. doi: 10.1002/acr.25017

19. Roberts AL, Malspeis S, Kubzansky LD, Feldman CH, Chang SC, Koenen KC, et al. Association of trauma and posttraumatic stress disorder with incident systemic lupus erythematosus in a longitudinal cohort of women. *Arthritis Rheumatol.* (2017) 69:2162–9. doi: 10.1002/art.40222

20. Cozier YC, Barbhaiya M, Castro-Webb N, Conte C, Tedeschi S, Leatherwood C, et al. Association of child abuse and systemic lupus erythematosus in Black women during adulthood. *Arthritis Care Res.* (2021) 73:833–40. doi: 10.1002/acr.24188

21. Costenbader KH, Feskanich D, Stampfer MJ, Karlson EW. Reproductive and menopausal factors and risk of systemic lupus erythematosus in women. *Arthritis Rheumatol.* (2007) 56:1251–62. doi: 10.1002/art.22510

22. Bernier MO, Mikaeloff Y, Hudson M, Suissa S. Combined oral contraceptive use and the risk of systemic lupus erythematosus. *Arthritis Rheumatol.* (2009) 61:476–81. doi: 10.1002/art.24398

23. Lateef A, Petri M. Hormone replacement and contraceptive therapy in autoimmune diseases. *J Autoimmun.* (2012) 38:J170–6. doi: 10.1016/j.jaut.2011.11.002

24. Hart JE, Laden F, Puett RC, Costenbader KH, Karlson EW. Exposure to traffic pollution and increased risk of rheumatoid arthritis. *Environ Health Perspectives.* (2009) 117:1065–9. doi: 10.1289/ehp.0800503

25. Parks CG, Cooper GS, Nylander-French LA, Sanderson WT, Dement JM, Cohen PL, et al. Occupational exposure to crystalline silica and risk of systemic lupus erythematosus: a population-based, case-control study in the southeastern United States. *Arthritis Rheumatol.* (2002) 46:1840–50. doi: 10.1002/art.10368

26. Karlson EW, Watts J, Signorovitch J, Bonetti M, Wright E, Cooper GS, et al. Effect of glutathione S-transferase polymorphisms and proximity to hazardous waste sites on time to systemic lupus erythematosus diagnosis: results from the Roxbury lupus project. *Arthritis Rheumatol.* (2007) 56:244–54. doi: 10.1002/art.22308

27. Williams JN, Chang SC, Sinnette C, Malspeis S, Parks CG, Karlson EW, et al. Pesticide exposure and risk of systemic lupus erythematosus in an urban population of predominantly African-American women. *Lupus.* (2018) 27:2129–34. doi: 10.1177/0961203318805844

28. Barbhaiya M, Costenbader KH. Ultraviolet radiation and systemic lupus erythematosus. *Lupus.* (2014) 23:588–95. doi: 10.1177/0961203314530488

29. Sontheimer C, Liggitt D, Elkon KB. Ultraviolet B irradiation causes stimulator of interferon genes-dependent production of protective type I interferon in mouse skin by recruited inflammatory monocytes. *Arthritis Rheumatol.* (2017) 69:826–36. doi: 10.1002/art.39987

30. Yin Q, Xu X, Lin Y, Lv J, Zhao L, He R. Ultraviolet B irradiation induces skin accumulation of plasmacytoid dendritic cells: a possible role for chemerin. *Autoimmunity.* (2014) 47:185–92. doi: 10.3109/08916934.2013.866105

31. Wolf SJ, Estadt SN, Theros J, Moore T, Ellis J, Liu J, et al. Ultraviolet light induces increased T cell activation in lupus-prone mice via type I IFN-dependent inhibition of T regulatory cells. *J Autoimmun.* (2019) 103:102291. doi: 10.1016/j.jaut.2019.06.002

32. Bengtsson A, Rylander L, Hagmar L, Nived O, Sturfelt G. Risk factors for developing systemic lupus erythematosus: a case–control study in southern Sweden. *Rheumatology.* (2002) 41:563–71. doi: 10.1093/rheumatology/41.5.563

33. Cooper GS, Wither J, Bernatsky S, Claudio JO, Clarke A, Rioux JD, et al. Occupational and environmental exposures and risk of systemic lupus erythematosus: silica, sunlight, solvents. *Rheumatology.* (2010) 49:2172–80. doi: 10.1093/rheumatology/keq214

34. Ritterhouse LL, Crowe SR, Niewold TB, Kamen DL, Macwana SR, Roberts VC, et al. Vitamin D deficiency is associated with an increased autoimmune response in healthy individuals and in patients with systemic lupus erythematosus. *Ann rheumatic diseases.* (2011) 70:1569–74. doi: 10.1136/ard.2010.148494

35. Jog NR, James JA. Epstein barr virus and autoimmune responses in systemic lupus erythematosus. *Front Immunol.* (2020) 11:623944. doi: 10.3389/fimmu.2020.623944

36. Jog NR, Young KA, Munroe ME, Harmon MT, Guthridge JM, Kelly JA, et al. Association of Epstein-Barr virus serological reactivation with transitioning to systemic lupus erythematosus in at-risk individuals. *Ann Rheumatic Diseases.* (2019) 78:1235–41. doi: 10.1136/annrheumdis-2019-215361

37. Jung C-R, Chung W-T, Chen W-T, Lee R-Y, Hwang B-F. Long-term exposure to traffic-related air pollution and systemic lupus erythematosus in Taiwan: A cohort study. *Sci Total Environment.* (2019) 668:342–9. doi: 10.1016/j.scitotenv.2019.03.018

38. Bernatsky S, Smargiassi A, Barnabe C, Svenson LW, Brand A, Martin RV, et al. Fine particulate air pollution and systemic autoimmune rheumatic disease in two Canadian provinces. *Environ Res.* (2016) 146:85–91. doi: 10.1016/j.envres.2015.12.021

39. Heliovaara M, Aho K, Aromaa A, Knekt P, Reunanen A. Smoking and risk of rheumatoid arthritis. *J Rheumatol.* (1993) 20:1830–5.

40. Pedersen M, Jacobsen S, Klarlund M, Pedersen BV, Wiik A, Wohlfahrt J, et al. Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. *Arthritis Res Ther.* (2006) 8:1–15. doi: 10.1186/ar2022

41. Crowson CS, Matteson EL, Davis JM III, Gabriel SE. Contribution of obesity to the rise in incidence of rheumatoid arthritis. *Arthritis Care Res.* (2013) 65:71–7. doi: 10.1002/acr.21660

42. Parks CG, Hoppin JA, De Roos AJ, Costenbader KH, Alavanja MC, Sandler DP. Rheumatoid arthritis in agricultural health study spouses: associations with pesticides and other farm exposures. *Environ Health perspectives.* (2016) 124:1728–34. doi: 10.1289/EHP129

43. Cooper GS, Parks CG, Treadwell EL, St Clair EW, Gilkeson GS, Dooley MA. Occupational risk factors for the development of systemic lupus erythematosus. *J Rheumatol.* (2004) 31:1928–33.

44. Nietert PJ, Sutherland SE, Silver RM, Pandey JP, Knapp RG, Hoel DG, et al. Is occupational organic solvent exposure a risk factor for scleroderma? *Arthritis Rheumatism: Off J Am Coll Rheumatol.* (1998) 41:1111–8.

45. Chen H-H, Huang N, Chen Y-M, Chen T-J, Chou P, Lee Y-L, et al. Association between a history of periodontitis and the risk of rheumatoid arthritis: a nationwide,

population-based, case–control study. *Ann Rheumatic Diseases*. (2013) 72:1206–11. doi: 10.1136/annrheumdis-2012-201593

46. Feldman CH, Malspeis S, Leatherwood C, Kubzansky L, Costenbader KH, Roberts AL. Association of childhood abuse with incident systemic lupus erythematosus in adulthood in a longitudinal cohort of women. *J Rheumatol*. (2019) 46:1589–96. doi: 10.3899/jrheum.190009

47. Stolt P, Källberg H, Lundberg I, Sjögren B, Klareskog L, Alfredsson L. Silica exposure is associated with increased risk of developing rheumatoid arthritis: results from the Swedish EIRA study. *Ann rheumatic diseases*. (2005) 64:582–6. doi: 10.1136/ard.2004.022053

48. Gómez-Puerta JA, Gedmintas L, Costenbader KH. The association between silica exposure and development of ANCA-associated vasculitis: systematic review and meta-analysis. *Autoimmun Rev*. (2013) 12:1129–35. doi: 10.1016/j.autrev.2013.06.016

49. McCormic ZD, Khuder SS, Aryal BK, Ames AL, Khuder SA. Occupational silica exposure as a risk factor for scleroderma: a meta-analysis. *Int Arch Occup Environ Health*. (2010) 83:763–9. doi: 10.1007/s00420-009-0505-7

50. Chang R, Chen TY, Wang SI, Hung YM, Chen HY, Wei CC. Risk of autoimmune diseases in patients with COVID-19: A retrospective cohort study. *EClinicalMedicine*. (2023) 56. doi: 10.1016/j.eclinm.2022.101783

51. Lu B, Solomon DH, Costenbader KH, Karlson EW. Alcohol consumption and risk of incident rheumatoid arthritis in women: a prospective study. *Arthritis Rheumatol*. (2014) 66:1998–2005. doi: 10.1002/art.38634

52. Hu Y, Sparks JA, Malspeis S, Costenbader KH, Hu FB, Karlson EW, et al. Long-term dietary quality and risk of developing rheumatoid arthritis in women. *Ann rheumatic diseases*. (2017) 76:1357–64. doi: 10.1136/annrheumdis-2016-210431

53. Choi MY, Hahn J, Malspeis S, Stevens EF, Karlson EW, Sparks JA, et al. Association of a combination of healthy lifestyle behaviors with reduced risk of incident systemic lupus erythematosus. *Arthritis Rheumatol*. (2022) 74:274–83. doi: 10.1002/art.41935

54. Hahn J, Malspeis S, Choi MY, Stevens E, Karlson EW, Lu B, et al. Association of healthy lifestyle behaviors and the risk of developing rheumatoid arthritis among women. *Arthritis Care Res (Hoboken)*. (2023) 75:272–6. doi: 10.1002/acr.24862

55. Karlson EW, Mandl LA, Hankinson SE, Grodstein F. Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis?: Results from the Nurses' Health Study. *Arthritis Rheumatism*. (2004) 50:3458–67. doi: 10.1002/art.20621

56. Hahn J, Cook NR, Alexander EK, Friedman S, Walter J, Bubes V, et al. Vitamin D and marine omega 3 fatty acid supplementation and incident autoimmune disease: VITAL randomized controlled trial. *BMJ*. (2022) 376:e066452. doi: 10.1136/bmj-2021-066452

57. Cui J, Raychaudhuri S, Karlson EW, Speyer C, Malspeis S, Guan H, et al. Interactions between genome-wide genetic factors and smoking influencing risk of systemic lupus erythematosus. *Arthritis Rheumatol*. (2020) 72:1863–71. doi: 10.1002/art.41414

58. Young KA, Munroe ME, Guthridge JM, Kamen DL, Niewold TB, Gilkeson GS, et al. Combined role of vitamin D status and CYP24A1 in the transition to systemic lupus erythematosus. *Ann Rheum Dis*. (2017) 76:153–8. doi: 10.1136/annrheumdis-2016-209157

59. Barbhaiya M, Costenbader KH. Environmental exposures and the development of systemic lupus erythematosus. *Curr Opin Rheumatol*. (2016) 28:497–505. doi: 10.1097/BOR.0000000000000318

60. Woo JM, Parks CG, Jacobsen S, Costenbader KH, Bernatsky S. The role of environmental exposures and gene-environment interactions in the etiology of systemic lupus erythematous. *J Intern Med*. (2022). doi: 10.1111/joim.13448

61. Olsen NJ, James JA, Arriens C, Ishimori ML, Wallace DJ, Kamen DL, et al. Study of Anti-Malarials in Incomplete Lupus Erythematosus (SMILE): study protocol for a randomized controlled trial. *Trials*. (2018) 19:694. doi: 10.1186/s13063-018-3076-7

62. Rech J, Tascilar K, Hagen M, Kleyer A, Manger B, Schoenau V, et al. Abatacept inhibits inflammation and onset of rheumatoid arthritis in individuals at high risk (ARIAA): a randomised, international, multicentre, double-blind, placebo-controlled trial. *Lancet*. (2024) 403(10429):850–9. doi: 10.1016/S0140-6736(23)02650-8

63. Cope AP, Jasenecova M, Vasconcelos JC, Filer A, Raza K, Qureshi S, et al. Abatacept in individuals at high risk of rheumatoid arthritis (APIPPRA): a randomised, double-blind, multicentre, parallel, placebo-controlled, phase 2b clinical trial. *Lancet*. (2024) 403(10429):838–49. doi: 10.1016/S0140-6736(23)02649-1

64. Vodovotz Y, Barnard N, Hu FB, Jakicic J, Lianov L, Loveland D, et al. Prioritized research for the prevention, treatment, and reversal of chronic disease: recommendations from the lifestyle medicine research summit. *Front Med (Lausanne)*. (2020) 7:585744. doi: 10.3389/fmed.2020.585744

65. Choi MY, Costenbader KH. Understanding the concept of pre-clinical autoimmunity: prediction and prevention of systemic lupus erythematosus: identifying risk factors and developing strategies against disease development. *Front Immunol*. (2022) 13:890522. doi: 10.3389/fimmu.2022.890522

66. Manna P, Jain SK. Obesity, oxidative stress, adipose tissue dysfunction, and the associated health risks: causes and therapeutic strategies. *Metab Syndr Relat Disord*. (2015) 13:423–44. doi: 10.1089/met.2015.0095

67. Włodarczyk M, Nowicka G. Obesity DNA. Damage, and development of obesity-related diseases. *Int J Mol Sci*. (2019) 20(5):1146. doi: 10.3390/ijms20051146

68. Petruzzelli S, Celi A, Pulerà N, Baliva F, Viegi G, Carrozzi L, et al. Serum antibodies to benzo (a) pyrene diol epoxide-DNA adducts in the general population: effects of air pollution, tobacco smoking, and family history of lung diseases. *Cancer Res*. (1998) 58:4122–6.

69. Mooney LA, Perera FP, Van Bennekum AM, Blaner WS, Karkoszka J, Covey L, et al. Gender differences in autoantibodies to oxidative DNA base damage in cigarette smokers. *Cancer Epidemiol Prev Biomarkers*. (2001) 10:641–8.

70. Morissette MC, Gao Y, Shen P, Thayaparan D, Bérubé JC, Paré PD, et al. Role of BAFF in pulmonary autoantibody responses induced by chronic cigarette smoke exposure in mice. *Physiol Rep*. (2016) 4(24):e1. doi: 10.14814/phy2.13057

71. Bermudez EA, Rifai N, Buring JE, Manson JE, Ridker PM. Relation between markers of systemic vascular inflammation and smoking in women. *Am J Cardiol*. (2002) 89:1117–9. doi: 10.1016/S0002-9149(02)02284-1

72. Tracy RP, Psaty BM, Macy E, Bovill EG, Cushman M, Cornell ES, et al. Lifetime smoking exposure affects the association of C-reactive protein with cardiovascular disease risk factors and subclinical disease in healthy elderly subjects. *Arteriosclerosis thrombosis Vasc Biol*. (1997) 17:2167–76. doi: 10.1161/01.ATV.17.10.2167

73. Hahn J, Leatherwood C, Malspeis S, Liu X, Lu B, Roberts AL, et al. Associations between smoking and systemic lupus erythematosus-related cytokines and chemokines among US female nurses. *Arthritis Care Res*. (2021) 73:1583–9. doi: 10.1002/acr.24370

74. Panagiotakos DB, Pitsavos C, Yannakoulia M, Chrysohoou C, Stefanadis C. The implication of obesity and central fat on markers of chronic inflammation: The ATTICA study. *Atherosclerosis*. (2005) 183:308–15. doi: 10.1016/j.atherosclerosis.2005.03.010

75. Castro-Webb N, Cozier YC, Barbhaiya M, Ruiz-Narvaez EA, Li S, Costenbader KH, et al. Association of macronutrients and dietary patterns with risk of systemic lupus erythematosus in the Black Women's Health Study. *Am J Clin Nutr*. (2021) 114:1486–94. doi: 10.1093/ajcn/nqab224

76. Tedeschi SK, Barbhaiya M, Sparks JA, Karlson EW, Kubzansky LD, Roberts AL, et al. Dietary patterns and risk of systemic lupus erythematosus in women. *Lupus*. (2020) 29:67–73. doi: 10.1177/0961203319888791

77. Barbhaiya M, Tedeschi S, Sparks JA, Leatherwood C, Karlson EW, Willett WC, et al. Association of dietary quality with risk of incident systemic lupus erythematosus in the nurses' Health studies. *Arthritis Care Res*. (2020) 73(9):1250–8. doi: 10.1002/acr.24443

78. Tedeschi SK, Costenbader KH. Is there a role for diet in the therapy of rheumatoid arthritis? *Curr Rheumatol Rep*. (2016) 18(5):23. doi: 10.1007/s11926-016-0575-y

79. Schafer AL, Eichhorst A, Hentze C, Kraemer AN, Amend A, Sprenger DTL, et al. Low dietary fiber intake links development of obesity and lupus pathogenesis. *Front Immunol*. (2021) 12:696810. doi: 10.3389/fimmu.2021.696810

80. Pan Y, Ke H, Yan Z, Geng Y, Asner N, Palani S, et al. The western-type diet induces anti-HMGB1 autoimmunity in Apoe(-/-) mice. *Atherosclerosis*. (2016) 251:31–8. doi: 10.1016/j.atherosclerosis.2016.05.027

81. Waldschmidt TJ, Cook RT, Kovacs EJ. Alcohol and inflammation and immune responses: summary of the 2006 Alcohol and Immunology Research Interest Group (AIRIG) meeting. *Alcohol*. (2008) 42:137–42. doi: 10.1016/j.alcohol.2007.11.003

82. Wirleitner B, Schroecksnadel K, Winkler C, Schennach H, Fuchs D. Resveratrol suppresses interferon-gamma-induced biochemical pathways in human peripheral blood mononuclear cells in vitro. *Immunol Lett*. (2005) 100:159–63. doi: 10.1016/j.imlet.2005.03.008

83. Young KA, Munroe ME, Harley JB, Guthridge JM, Kamen DL, Gilkensen GS, et al. Less than 7 hours of sleep per night is associated with transitioning to systemic lupus erythematosus. *Lupus*. (2018) 27:1524–31. doi: 10.1177/0961203318778368

84. Hsiao YH, Chen YT, Tseng CM, Wu LA, Lin WC, Su VY, et al. Sleep disorders and increased risk of autoimmune diseases in individuals without sleep apnea. *Sleep*. (2015) 38:581–6. doi: 10.5665/sleep.4574

85. Irwin MR, Wang M, Campomayor CO, Collado-Hidalgo A, Cole S. Sleep deprivation and activation of morning levels of cellular and genomic markers of inflammation. *Arch Intern Med*. (2006) 166:1756–62. doi: 10.1001/archinte.166.16.1756

86. Vgontzas AN, Zoumakis E, Bixler EO, Lin HM, Follett H, Kales A, et al. Adverse effects of modest sleep restriction on sleepiness, performance, and inflammatory cytokines. *J Clin Endocrinol Metab*. (2004) 89:2119–26. doi: 10.1210/jc.2003-031562

87. Bollinger T, Bollinger A, Skrum L, Dimitrov S, Lange T, Solbach W. Sleep-dependent activity of T cells and regulatory T cells. *Clin Exp Immunol*. (2009) 155:231–8. doi: 10.1111/j.1365-2249.2008.03822.x

88. Clinton JM, Davis CJ, Zielinski MR, Jewett KA, Krueger JM. Biochemical regulation of sleep and sleep biomarkers. *J Clin Sleep Med*. (2011) 7:S38–42. doi: 10.5664/JCSM.1360

89. Wilder-Smith A, Mustafa FB, Earnest A, Gen L, Macary PA. Impact of partial sleep deprivation on immune markers. *Sleep Med*. (2013) 14:1031–4. doi: 10.1016/j.sleep.2013.07.001

90. Palma BD, Gabriel A Jr., Colugnati FA, Tufik S. Effects of sleep deprivation on the development of autoimmune disease in an experimental model of systemic lupus erythematosus. *Am J Physiol Regul Integr Comp Physiol*. (2006) 291:R1527–32. doi: 10.1152/ajpregu.00186.2006

91. Calcagni E, Elenkov I. Stress system activity, innate and T helper cytokines, and susceptibility to immune-related diseases. *Ann N Y Acad Sci*. (2006) 1069:62–76. doi: 10.1196/annals.1351.006

92. Case SM, Feldman CH, Guan H, Stevens E, Kubzansky LD, Koenen KC, et al. Posttraumatic stress disorder (PTSD) and risk of systemic lupus erythematosus (SLE) among medicaid recipients. *Arthritis Care Res (Hoboken)*. (2021) 75(1):174–9. doi: 10.1002/acr.24758

93. Sumner JA, Chen Q, Roberts AL, Winning A, Rimm EB, Gilsanz P, et al. Posttraumatic stress disorder onset and inflammatory and endothelial function biomarkers in women. *Brain Behav Immun*. (2018) 69:203–9. doi: 10.1016/j.bbi.2017.11.013

94. Passos IC, Vasconcelos-Moreno MP, Costa LG, Kunz M, Brietzke E, Quevedo J, et al. Inflammatory markers in post-traumatic stress disorder: a systematic review, meta-analysis, and meta-regression. *Lancet Psychiatry*. (2015) 2:1002–12. doi: 10.1016/S2215-0366(15)00309-0

95. Gill J, Vythilingam M, Page GG. Low cortisol, high DHEA, and high levels of stimulated TNF-alpha, and IL-6 in women with PTSD. *J Trauma Stress*. (2008) 21:530–9. doi: 10.1002/jts.20372

96. Pace TW, Wingenfeld K, Schmidt I, Meinlschmidt G, Hellhammer DH, Heim CM. Increased peripheral NF-kappaB pathway activity in women with childhood abuse-related posttraumatic stress disorder. *Brain Behav Immun*. (2012) 26:13–7. doi: 10.1016/j.bbi.2011.07.232

97. Lindqvist D, Wolkowitz OM, Mellon S, Yehuda R, Flory JD, Henn-Haase C, et al. Proinflammatory milieu in combat-related PTSD is independent of depression and early life stress. *Brain Behav Immun*. (2014) 42:81–8. doi: 10.1016/j.bbi.2014.06.003

98. Slopen N, Kubzansky LD, McLaughlin KA, Koenen KC. Childhood adversity and inflammatory processes in youth: a prospective study. *Psychoneuroendocrinology*. (2013) 38:188–200. doi: 10.1016/j.psyneuen.2012.05.013

99. Gola H, Engler H, Sommershof A, Adenauer H, Kolassa S, Schedlowski M, et al. Posttraumatic stress disorder is associated with an enhanced spontaneous production of pro-inflammatory cytokines by peripheral blood mononuclear cells. *BMC Psychiatry*. (2013) 13:40. doi: 10.1186/1471-244X-13-40

100. Hartwell KJ, Moran-Santa Maria MM, Twal WO, Shaftman S, DeSantis SM, McRae-Clark AL, et al. Association of elevated cytokines with childhood adversity in a sample of healthy adults. *J Psychiatr Res*. (2013) 47:604–10. doi: 10.1016/j.jpsychires.2013.01.008

101. Danese A, Moffitt TE, Pariante CM, Ambler A, Poulton R, Caspi A. Elevated inflammation levels in depressed adults with a history of childhood maltreatment. *Arch Gen Psychiatry*. (2008) 65:409–15. doi: 10.1001/archpsyc.65.4.409

102. Stewart JC, Rand KL, Muldoon MF, Kamarck TW. A prospective evaluation of the directionality of the depression-inflammation relationship. *Brain Behav Immun*. (2009) 23:936–44. doi: 10.1016/j.bbi.2009.04.011

103. Grimaldi CM, Cleary J, Dagtas AS, Moussai D, Diamond B. Estrogen alters thresholds for B cell apoptosis and activation. *J Clin Invest*. (2002) 109:1625–33. doi: 10.1172/JCI0214873

104. Boudigaard SH, Schlunssen V, Vestergaard JM, Sondergaard K, Toren K, Peters S, et al. Occupational exposure to respirable crystalline silica and risk of autoimmune rheumatic diseases: a nationwide cohort study. *Int J Epidemiol*. (2021) 50:1213–26. doi: 10.1093/ije/dyaa287

105. Finckh A, Cooper GS, Chibnik LB, Costenbader KH, Watts J, Pankey H, et al. Occupational silica and solvent exposures and risk of systemic lupus erythematosus in urban women. *Arthritis Rheumatol*. (2006) 54:3648–54. doi: 10.1002/art.22210

106. Bernatsky S, Smargiassi A, Johnson M, Kaplan GG, Barnabe C, Svenson L, et al. Fine particulate air pollution, nitrogen dioxide, and systemic autoimmune rheumatic disease in Calgary, Alberta. *Environ Res*. (2015) 140:474–8. doi: 10.1016/j.envres.2015.05.007

107. Parks CG, Walitt BT, Pettinger M, Chen JC, de Roos AJ, Hunt J, et al. Insecticide use and risk of rheumatoid arthritis and systemic lupus erythematosus in the Women's Health Initiative Observational Study. *Arthritis Care Res (Hoboken)*. (2011) 63:184–94. doi: 10.1002/acr.20335

108. Miller FW, Alfredsson L, Costenbader KH, Kamen DL, Nelson LM, Norris JM, et al. Epidemiology of environmental exposures and human autoimmune diseases: findings from a National Institute of Environmental Health Sciences Expert Panel Workshop. *J Autoimmun*. (2012) 39:259–71. doi: 10.1016/j.jaut.2012.05.002

109. Crowe W, Allsopp PJ, Watson GE, Magee PJ, Strain JJ, Armstrong DJ, et al. Mercury as an environmental stimulus in the development of autoimmunity - A systematic review. *Autoimmun Rev*. (2017) 16:72–80. doi: 10.1016/j.autrev.2016.09.020

110. Pollard KM, Cauvi DM, Toomey CB, Hultman P, Kono DH. Mercury-induced inflammation and autoimmunity. *Biochim Biophys Acta Gen Subj*. (2019) 1863:129299. doi: 10.1016/j.bbagen.2019.02.001

111. Pollard KM, Escalante GM, Huang H, Haraldsson KM, Hultman P, Christy JM, et al. Induction of systemic autoimmunity by a xenobiotic requires endosomal TLR trafficking and signaling from the late endosome and endolysosome but not type I IFN. *J Immunol*. (2017) 199:3739–47. doi: 10.4049/jimmunol.1700332

112. McGrath H. Ultraviolet-A1 irradiation decreases clinical disease activity and autoantibodies in patients with systemic lupus erythematosus. *Clin Exp Rheumatol*. (1994) 12:129–35.

113. Costenbader KH, Feskanich D, Holmes M, Karlson EW, Benito-Garcia E. Vitamin D intake and risks of systemic lupus erythematosus and rheumatoid arthritis in women. *Ann Rheum Dis*. (2008) 67:530–5. doi: 10.1136/ard.2007.072736

114. Cutolo M, Otsa K, Paolino S, Yprus M, Veldi T, Seriolo B. Vitamin D involvement in rheumatoid arthritis and systemic lupus erythaematosus. *Ann Rheum Dis*. (2009) 68:446–7. doi: 10.1136/ard.2008.093476

115. Iruretagoyena M, Hirigoyen D, Naves R, Burgos PI. Immune response modulation by vitamin D: role in systemic lupus erythematosus. *Front Immunol*. (2015) 6:513. doi: 10.3389/fimmu.2015.00513

116. Iwakiri D, Zhou L, Samanta M, Matsumoto M, Ebihara T, Seya T, et al. Epstein-Barr virus (EBV)-encoded small RNA is released from EBV-infected cells and activates signaling from Toll-like receptor 3. *J Exp Med*. (2009) 206:2091–9. doi: 10.1084/jem.20081761

117. Verdoni L, Mazza A, Gervasoni A, Martelli L, Ruggeri M, Ciuffreda M, et al. An outbreak of severe Kawasaki-like disease at the Italian epicentre of the SARS-CoV-2 epidemic: an observational cohort study. *Lancet*. (2020) 395:1771–8. doi: 10.1016/S0140-6736(20)31103-X

118. Riphagen S, Gomez X, Gonzalez-Martinez C, Wilkinson N, Theocharis P. Hyperinflammatory shock in children during COVID-19 pandemic. *Lancet*. (2020) 395:1607–8. doi: 10.1016/S0140-6736(20)31094-1

119. Belhadjer Z, Méot M, Bajolle F, Khraiche D, Legendre A, Abakka S, et al. Acute heart failure in multisystem inflammatory syndrome in children in the context of global SARS-CoV-2 pandemic. *Circulation*. (2020) 142:429–36. doi: 10.1161/CIRCULATIONAHA.120.048360

120. Tesch F, Ehm F, Vivirito A, Wende D, Batram M, Loser F, et al. Incident autoimmune diseases in association with SARS-CoV-2 infection: a matched cohort study. *Clin Rheumatol*. (2023) 42:2905–14. doi: 10.1007/s10067-023-06670-0

121. Banchereau J, Pascual V. Type I interferon in systemic lupus erythematosus and other autoimmune diseases. *Immunity*. (2006) 25:383–92. doi: 10.1016/j.immuni.2006.08.010

122. Wahadat MJ, Bodewes ILA, Maria NI, van Helden-Meeuwsen CG, van Dijk-Hummelman A, Steenwijk EC, et al. Type I IFN signature in childhood-onset systemic lupus erythematosus: a conspiracy of DNA- and RNA-sensing receptors? *Arthritis Res Ther*. (2018) 20:4. doi: 10.1186/s13075-017-1501-z

123. Law C, Wacleche VS, Cao Y, Pillai A, Sowerby J, Hancock B, et al. Interferon subverts an AHR–JUN axis to promote CXCL13+ T cells in lupus. *Nature*. (2024) 631, 857–866. doi: 10.1038/s41586-024-07627-2

124. Rothhammer V, Quintana FJ. The aryl hydrocarbon receptor: an environmental sensor integrating immune responses in health and disease. *Nat Rev Immunol*. (2019) 19:184–97. doi: 10.1038/s41577-019-0125-8

125. Zhao Y, Ma T, Chen W, Chen Y, Li M, Ren L, et al. MicroRNA-124 promotes intestinal inflammation by targeting aryl hydrocarbon receptor in Crohn's disease. *J Crohn's Colitis*. (2016) 10:703–12. doi: 10.1093/ecco-jcc/jjw010

126. Quintana FJ, Basso AS, Iglesias AH, Korn T, Farez MF, Bettelli E, et al. Control of Treg and TH17 cell differentiation by the aryl hydrocarbon receptor. *Nature*. (2008) 453:65–71. doi: 10.1038/nature06880

127. Lebwohl MG, Stein Gold L, Strober B, Papp KA, Armstrong AW, Bagel J, et al. Phase 3 trials of tapinarof cream for plaque psoriasis. *New Engl J Med*. (2021) 385:2219–29. doi: 10.1056/NEJMoa2103629

128. Andersen AM, Lei MK, Beach SRH, Philibert RA, Sinha S, Colgan JD. Cigarette and cannabis smoking effects on GPR15+ Helper T cell levels in peripheral blood: relationships with epigenetic biomarkers. *Genes (Basel)*. (2020) 11. doi: 10.3390/genes11020149

129. Talbot J, Peres RS, Pinto LG, Oliveira RDR, Lima KA, Donate PB, et al. Smoking-induced aggravation of experimental arthritis is dependent of aryl hydrocarbon receptor activation in Th17 cells. *Arthritis Res Ther*. (2018) 20:119. doi: 10.1186/s13075-018-1609-9

130. Tantoh DM, Wu MC, Chuang CC, Chen PH, Tyan YS, Nfor ON, et al. AHRR cg05575921 methylation in relation to smoking and PM(2.5) exposure among Taiwanese men and women. *Clin Epigenet*. (2020) 12:117. doi: 10.1186/s13148-020-00908-3

131. Somers EC, Richardson BC. Environmental exposures, epigenetic changes and the risk of lupus. *Lupus*. (2014) 23:568–76. doi: 10.1177/0961203313499419

132. Teruel M, Sawalha AH. Epigenetic variability in systemic lupus erythematosus: what we learned from genome-wide DNA methylation studies. *Curr Rheumatol Rep*. (2017) 19:32. doi: 10.1007/s11926-017-0657-5

133. Richardson B, Scheinbart L, Strahler J, Gross L, Hanash S, Johnson M. Evidence for impaired T cell DNA methylation in systemic lupus erythematosus and rheumatoid arthritis. *Arthritis Rheumatol*. (1990) 33:1665–73. doi: 10.1002/art.1780331109

134. Gorelik GJ, Yarlagadda S, Patel DR, Richardson BC. Protein kinase Cdelta oxidation contributes to ERK inactivation in lupus T cells. *Arthritis Rheumatol*. (2012) 64:2964–74. doi: 10.1002/art.34503

135. Zong D, Liu X, Li J, Ouyang R, Chen P. The role of cigarette smoke-induced epigenetic alterations in inflammation. *Epigenet Chromatin*. (2019) 12:65. doi: 10.1186/s13072-019-0311-8

136. Lanata CM, Chung SA, Criswell LA. DNA methylation 101: what is important to know about DNA methylation and its role in SLE risk and disease heterogeneity. *Lupus Sci Med*. (2018) 5:e000285. doi: 10.1136/lupus-2018-000285

137. Chung SA, Nititham J, Elboudwarej E, Quach HL, Taylor KE, Barcellos LF, et al. Genome-wide assessment of differential DNA methylation associated with

autoantibody production in systemic lupus erythematosus. *PloS One*. (2015) 10: e0129813. doi: 10.1371/journal.pone.0129813

138. Mok A, Solomon O, Nayak RR, Coit P, Quach HL, Nititham J, et al. Genome-wide profiling identifies associations between lupus nephritis and differential methylation of genes regulating tissue hypoxia and type 1 interferon responses. *Lupus Sci Med*. (2016) 3:e000183. doi: 10.1136/lupus-2016-000183

139. Coit P, Renauer P, Jeffries MA, Merrill JT, McCune WJ, Maksimowicz-McKinnon K, et al. Renal involvement in lupus is characterized by unique DNA methylation changes in naive CD4+ T cells. *J Autoimmun*. (2015) 61:29–35. doi: 10.1016/j.jaut.2015.05.003

140. Absher DM, Li X, Waite LL, Gibson A, Roberts K, Edberg J, et al. Genome-wide DNA methylation analysis of systemic lupus erythematosus reveals persistent hypomethylation of interferon genes and compositional changes to CD4+ T-cell populations. *PloS Genet*. (2013) 9:e1003678. doi: 10.1371/journal.pgen.1003678

141. Zhao M, Zhou Y, Zhu B, Wan M, Jiang T, Tan Q, et al. IFI44L promoter methylation as a blood biomarker for systemic lupus erythematosus. *Ann Rheum Dis*. (2016) 75:1998–2006. doi: 10.1136/annrheumdis-2015-208410

142. Miller S, Tsou PS, Coit P, Gensterblum-Miller E, Renauer P, Rohraff DM, et al. Hypomethylation of STAT1 and HLA-DRB1 is associated with type-I interferon-dependent HLA-DRB1 expression in lupus CD8+ T cells. *Ann Rheum Dis*. (2019) 78:519–28. doi: 10.1136/annrheumdis-2018-214323

143. Coit P, Ortiz-Fernandez L, Lewis EE, McCune WJ, Maksimowicz-McKinnon K, Sawalha AH. A longitudinal and transancestral analysis of DNA methylation patterns and disease activity in lupus patients. *JCI Insight*. (2020) 5. doi: 10.1172/jci.insight.143654

144. Lanata CM, Paranjpe I, Nititham J, Taylor KE, GianFrancesco M, Paranjpe M, et al. A phenotypic and genomics approach in a multi-ethnic cohort to subtype systemic lupus erythematosus. *Nat Commun*. (2019) 10:3902. doi: 10.1038/s41467-019-11845-y

145. Joseph S, George NI, Green-Knox B, Treadwell EL, Word B, Yim S, et al. Epigenome-wide association study of peripheral blood mononuclear cells in systemic lupus erythematosus: Identifying DNA methylation signatures associated with interferon-related genes based on ethnicity and SLEDAI. *J Autoimmun*. (2019) 96:147–57. doi: 10.1016/j.jaut.2018.09.007

146. Barrat FJ, Crow MK, Ivashkiv LB. Interferon target-gene expression and epigenomic signatures in health and disease. *Nat Immunol*. (2019) 20:1574–83. doi: 10.1038/s41590-019-0466-2

147. Tsou PS, Coit P, Kilian NC, Sawalha AH. EZH2 modulates the DNA methylome and controls T cell adhesion through junctional adhesion molecule A in lupus patients. *Arthritis Rheumatol*. (2018) 70:98–108. doi: 10.1002/art.40338

148. Park SH, Kang K, Giannopoulou E, Qiao Y, Kang K, Kim G, et al. Type I interferons and the cytokine TNF cooperatively reprogram the macrophage epigenome to promote inflammatory activation. *Nat Immunol*. (2017) 18:1104–16. doi: 10.1038/ni.3818

149. Renauer P, Coit P, Jeffries MA, Merrill JT, McCune WJ, Maksimowicz-McKinnon K, et al. DNA methylation patterns in naive CD4+ T cells identify epigenetic susceptibility loci for malar rash and discoid rash in systemic lupus erythematosus. *Lupus Sci Med*. (2015) 2:e000101. doi: 10.1136/lupus-2015-000101

150. Breitbach ME, Ramaker RC, Roberts K, Kimberly RP, Absher D. Population-specific patterns of epigenetic defects in the B cell lineage in patients with systemic lupus erythematosus. *Arthritis Rheumatol*. (2020) 72:282–91. doi: 10.1002/art.41083

151. Silverman GJ, Azzouz DF, Alekseyenko AV. Systemic Lupus Erythematosus and dysbiosis in the microbiome: cause or effect or both? *Curr Opin Immunol*. (2019) 61:80–5. doi: 10.1016/j.coi.2019.08.007

152. Azzouz D, Omarbekova A, Heguy A, Schwudke D, Gisch N, Rovin BH, et al. Lupus nephritis is linked to disease-activity associated expansions and immunity to a gut commensal. *Ann Rheum Dis*. (2019) 78:947–56. doi: 10.1136/annrheumdis-2018-214856

153. He Z, Shao T, Li H, Xie Z, Wen C. Alterations of the gut microbiome in Chinese patients with systemic lupus erythematosus. *Gut Pathog*. (2016) 8:64. doi: 10.1186/s13099-016-0146-9

154. Hevia A, Milani C, Lopez P, Cuervo A, Arboleya S, Duranti S, et al. Intestinal dysbiosis associated with systemic lupus erythematosus. *mBio*. (2014) 5:e01548–14. doi: 10.1128/mBio.01548-14

155. Chua H-H, Chou H-C, Tung Y-L, Chiang B-L, Liao C-C, Liu H-H, et al. Intestinal dysbiosis featuring abundance of Ruminococcus gnavus associates with allergic diseases in infants. *Gastroenterology*. (2018) 154:154–67. doi: 10.1053/j.gastro.2017.09.006

156. Breban M, Tap J, Leboime A, Said-Nahal R, Langella P, Chiocchia G, et al. Faecal microbiota study reveals specific dysbiosis in spondyloarthritis. *Ann rheumatic diseases*. (2017) 76:1614–22. doi: 10.1136/annrheumdis-2016-211064

157. Hall AB, Yassour M, Sauk J, Garner A, Jiang X, Arthur T, et al. A novel Ruminococcus gnavus clade enriched in inflammatory bowel disease patients. *Genome Med*. (2017) 9:1–12. doi: 10.1186/s13073-017-0490-5

158. Bunker JJ, Drees C, Watson AR, Plunkett CH, Nagler CR, Schneewind O, et al. B cell superantigens in the human intestinal microbiota. *Sci Transl Med*. (2019) 11 (507):eaau9356. doi: 10.1126/scitranslmed.aau9356

159. Shi L, Zhang Z, Yu AM, Wang W, Wei Z, Akhter E, et al. The SLE transcriptome exhibits evidence of chronic endotoxin exposure and has widespread dysregulation of non-coding and coding RNAs. *PloS One*. (2014) 9:e93846. doi: 10.1371/journal.pone.0093846

160. Li BZ, Zhou HY, Guo B, Chen WJ, Tao JH, Cao NW, et al. Dysbiosis of oral microbiota is associated with systemic lupus erythematosus. *Arch Oral Biol*. (2020) 113:104708. doi: 10.1016/j.archoralbio.2020.104708

161. Correa JD, Calderaro DC, Ferreira GA, Mendonca SM, Fernandes GR, Xiao E, et al. Subgingival microbiota dysbiosis in systemic lupus erythematosus: association with periodontal status. *Microbiome*. (2017) 5:34. doi: 10.1186/s40168-017-0252-z

162. Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, et al. The long-term stability of the human gut microbiota. *Science*. (2013) 341:1237439. doi: 10.1126/science.1237439

163. Maier L, Pruteanu M, Kuhn M, Zeller G, Telzerow A, Anderson EE, et al. Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature*. (2018) 555:623–8. doi: 10.1038/nature25979

164. Azzouz DF, Chen Z, Izmirly PM, Chen LA, Li Z, Zhang C, et al. Longitudinal gut microbiome analyses and blooms of pathogenic strains during lupus disease flares. *Ann rheumatic diseases*. (2023) 82:1315–27. doi: 10.1136/ard-2023-223929

165. Mikuls TR, Payne JB, Yu F, Thiele GM, Reynolds RJ, Cannon GW, et al. Periodontitis and Porphyromonas gingivalis in patients with rheumatoid arthritis. *Arthritis Rheumatol*. (2014) 66:1090–100. doi: 10.1002/art.38348

166. Mikuls TR, Thiele GM, Deane KD, Payne JB, O'Dell JR, Yu F, et al. Porphyromonas gingivalis and disease-related autoantibodies in individuals at increased risk of rheumatoid arthritis. *Arthritis Rheumatism*. (2012) 64:3522–30. doi: 10.1002/art.34595

167. Konig MF, Abusleme L, Reinholdt J, Palmer RJ, Teles RP, Sampson K, et al. Aggregatibacter actinomycetemcomitans–induced hypercitrullination links periodontal infection to autoimmunity in rheumatoid arthritis. *Sci Trans Med*. (2016) 8:369ra176–369ra176. doi: 10.1126/scitranslmed.aaj1921

168. Scher JU, Joshua V, Artacho A, Abdollahi-Roodsaz S, Öckinger J, Kullberg S, et al. The lung microbiota in early rheumatoid arthritis and autoimmunity. *Microbiome*. (2016) 4:1–10. doi: 10.1186/s40168-016-0206-x

169. Reynisdottir G, Karimi R, Joshua V, Olsen H, Hensvold AH, Harju A, et al. Structural changes and antibody enrichment in the lungs are early features of anti–citrullinated protein antibody–positive rheumatoid arthritis. *Arthritis Rheumatol*. (2014) 66:31–9. doi: 10.1002/art.38201

170. Juge P-A, Lee JS, Ebstein E, Furukawa H, Dobrinskikh E, Gazal S, et al. MUC5B promoter variant and rheumatoid arthritis with interstitial lung disease. *New Engl J Med*. (2018) 379:2209–19. doi: 10.1056/NEJMoa1801562

171. Rodriguez Huerta MD, Trujillo-Martin MM, Rua-Figueroa I, Cuellar-Pompa L, Quiros-Lopez R, Serrano-Aguilar P, et al. Healthy lifestyle habits for patients with systemic lupus erythematosus: A systemic review. *Semin Arthritis Rheumatol*. (2016) 45:463–70. doi: 10.1016/j.semarthrit.2015.09.003

172. Cui J, Malspeis S, Choi MY, Lu B, Sparks JA, Yoshida K, et al. Risk prediction models for incident systemic lupus erythematosus among women in the Nurses' health study cohorts using genetics, family history, and lifestyle and environmental factors. *Semin Arthritis Rheumatol*. (2023) 58:152143. doi: 10.1016/j.semarthrit.2022.152143

173. Choi MY, Clarke AE, St Pierre Y, Hanly JG, Urowitz MB, Romero-Diaz J, et al. The prevalence and determinants of anti-DFS70 autoantibodies in an international inception cohort of systemic lupus erythematosus patients. *Lupus*. (2017) 26:1051–9. doi: 10.1177/0961203317692437

174. Olsen NJ, McAloose C, Carter J, Han BK, Raman I, Li QZ, et al. Clinical and immunologic profiles in incomplete lupus erythematosus and improvement with hydroxychloroquine treatment. *Autoimmune Dis*. (2016) 2016:8791629. doi: 10.1155/2016/8791629

175. Lu R, Munroe ME, Guthridge JM, Bean KM, Fife DA, Chen H, et al. Dysregulation of innate and adaptive serum mediators precedes systemic lupus erythematosus classification and improves prognostic accuracy of autoantibodies. *J Autoimmun*. (2016) 74:182–93. doi: 10.1016/j.jaut.2016.06.001

176. Munroe ME, Lu R, Zhao YD, Fife DA, Robertson JM, Guthridge JM, et al. Altered type II interferon precedes autoantibody accrual and elevated type I interferon activity prior to systemic lupus erythematosus classification. *Ann Rheum Dis*. (2016) 75:2014–21. doi: 10.1136/annrheumdis-2015-208140

177. Niewold TB, Hua J, Lehman TJA, Harley JB, Crow MK. High serum IFN-α activity is a heritable risk factor for systemic lupus erythematosus. *Genes Immunity*. (2007) 8:492–502. doi: 10.1038/sj.gene.6364408

178. Ramsey-Goldman R, Alexander RV, Massarotti EM, Wallace DJ, Narain S, Arriens C, et al. Complement activation in patients with probable systemic lupus erythematosus and ability to predict progression to american college of rheumatology-classified systemic lupus erythematosus. *Arthritis Rheumatol*. (2020) 72(1):78–88. doi: 10.1002/art.41093

179. Deane KD. Targeting environmental risks to prevent rheumatic disease. *Rheum Dis Clin North Am*. (2022) 48:931–43. doi: 10.1016/j.rdc.2022.06.011

180. Costenbader KH, Cook NR, Lee IM, Hahn J, Walter J, Bubes V, et al. Vitamin D and marine n-3 fatty acids for autoimmune disease prevention: outcomes two years after completion of a double-blind, placebo-controlled trial. *Arthritis Rheumatol*. (2024) 76(6):973–83. doi: 10.1002/art.42811

181. Petri M, Orbai AM, Alarcon GS, Gordon C, Merrill JT, Fortin PR, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheumatol.* (2012) 64:2677–86. doi: 10.1002/art.34473

182. Deane KD, Striebich C, Feser M, Demoruelle K, Moss L, Bemis E, et al. Hydroxychloroquine does not prevent the future development of rheumatoid arthritis in a population with baseline high levels of antibodies to citrullinated protein antigens and absence of inflammatory arthritis: interim analysis of the StopRA trial. *Arthritis Rheumatol.* (2022) 74:3180–2.

183. Krijbolder DI, Verstappen M, van Dijk BT, Dakkak YJ, Burgers LE, Boer AC, et al. Intervention with methotrexate in patients with arthralgia at risk of rheumatoid arthritis to reduce the development of persistent arthritis and its disease burden (TREAT EARLIER): a randomised, double-blind, placebo-controlled, proof-of-concept trial. *Lancet.* (2022) 400:283–94. doi: 10.1016/S0140-6736(22)01193-X

184. Finckh A, Courvoisier D, Lamacchia C. Recherche clinique en rhumatismes i. Measuring ACPA in the general population or primary care: is it useful? *RMD Open.* (2020) 6(1). doi: 10.1136/rmdopen-2019-001085

185. Zhan K, Buhler KA, Chen IY, Fritzler MJ, Choi MY. Systemic lupus in the era of machine learning medicine. *Lupus Sci Med.* (2024) 11. doi: 10.1136/lupus-2023-001140

186. Filipp FV. Opportunities for artificial intelligence in advancing precision medicine. *Curr Genet Med Rep.* (2019) 7:208–13. doi: 10.1007/s40142-019-00177-4

187. Foulquier N, Redou P, Saraux A. How health information technologies and artificial intelligence may help rheumatologists in routine practice. *Rheumatol Ther.* (2019) 6:135–8. doi: 10.1007/s40744-019-0154-6

188. Choi MY, Chen I, Clarke AE, Fritzler MJ, Buhler KA, Urowitz M, et al. Machine learning identifies clusters of longitudinal autoantibody profiles predictive of systemic lupus erythematosus disease outcomes. *Ann rheumatic diseases.* (2023) 82:927–36. doi: 10.1136/ard-2022-223808

189. Adamichou C, Genitsaridi I, Nikolopoulos D, Nikoloudaki M, Repa A, Bortoluzzi A, et al. Lupus or not? SLE Risk Probability Index (SLERPI): a simple, clinician-friendly machine learning-based model to assist the diagnosis of systemic lupus erythematosus. *Ann Rheum Dis.* (2021) 80:758–66. doi: 10.1136/annrheumdis-2020-219069

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 14 January 2025
DOI 10.3389/fimmu.2024.1523990

OPEN ACCESS

EDITED BY
Esther Erdei,
University of New Mexico Health Sciences Center, United States

REVIEWED BY
Justyna Roszkiewicz,
Medical University of Lodz, Poland
Lillemor Berntson,
Uppsala University, Sweden

*CORRESPONDENCE
Vilde Øverlien Dåstøl
vilde.dastol@gmail.com

RECEIVED 06 November 2024
ACCEPTED 16 December 2024
PUBLISHED 14 January 2025

CITATION
Dåstøl VØ, Haftorn KL, Rudsari HK, Jaholkowski PP, Størdal K, Håberg SE, Weinberg CR, Rider LG, Andreassen OA, Brantsæter AL, Caspersen IH and Sanner H (2025) Maternal seafood intake, dietary contaminant exposure, and risk of juvenile idiopathic arthritis: exploring gene-environment interactions.
Front. Immunol. 15:1523990.
doi: 10.3389/fimmu.2024.1523990

Maternal seafood intake, dietary contaminant exposure, and risk of juvenile idiopathic arthritis: exploring gene-environment interactions

Vilde Øverlien Dåstøl[1,2*], Kristine Løkås Haftorn[1], Hamid Khoshfekr Rudsari[1], Piotr Pawel Jaholkowski[3,4], Ketil Størdal[5,6], Siri Eldevik Håberg[7,8], Clarice R. Weinberg[9], Lisa G. Rider[10], Ole A. Andreassen[2,4,11], Anne Lise Brantsæter[12], Ida Henriette Caspersen[1,7] and Helga Sanner[1,13]

[1]Department of Rheumatology, Oslo University Hospital, Oslo, Norway, [2]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway, [3]Center for Precision Psychiatry, Division of Mental Health and Addiction, Oslo University Hospital, and Institute of Clinical Medicine,University of Oslo, Oslo, Norway, [4]KG Jebsen Centre for Neurodevelopmental Disorders, University of Oslo and Oslo University Hospital, Oslo, Norway, [5]Department of Pediatric Research, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway, [6]Department of Pediatric and Adolescent Medicine, Oslo University Hospital, Oslo, Norway, [7]Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway, [8]Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway, [9]Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC, United States, [10]Environmental Autoimmunity Group, Clinical Research Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Bethesda, MD, United States, [11]Center for Precision Psychiatry, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway, [12]Department of Food Safety and Centre for Sustainable Diets, Norwegian Institute of Public Health, Oslo, Norway, [13]Oslo New University College, Oslo, Norway

Objectives: Juvenile idiopathic arthritis (JIA) originates from a complex interplay between genetic and environmental factors. We investigated the association between seafood intake and dietary contaminant exposure during pregnancy and JIA risk, to identify sex differences and gene-environment interactions.

Methods: We used the Norwegian Mother, Father, and Child Cohort Study (MoBa), a population-based prospective pregnancy cohort (1999–2008). JIA patients were identified through the Norwegian Patient Registry, with remaining mother-child pairs serving as controls. We assessed maternal seafood intake and dietary contaminants typically found in seafood using a food frequency questionnaire completed during pregnancy, mainly comparing high (≥90th percentile, P90) vs low (<P90) intake. Multivariable logistic regression calculated adjusted odds ratios (aOR), including sex-stratification analyses. A polygenic risk score (PRS) for JIA was used in a subsample to assess gene-environment interactions.

Results: We identified 217 JIA patients and 71,884 controls. High vs low maternal intake of lean/semi-oily fish was associated with JIA (aOR 1.51, 95% CI 1.02-2.22), especially among boys (aOR 2.13, 95% CI 1.21-3.75). A significant gene-environment interaction was observed between total fish intake and PRS, with high fish intake associated with JIA primarily in those with low PRS (p<0.03). We

found no associations between high vs low exposure to other types of seafood or environmental contaminants and JIA.

Conclusions: We found a modestly increased risk of JIA associated with high intake of lean/semi-oily fish during pregnancy, not explained by estimated exposure to dietary contaminants. Our data suggest a more pronounced association in children with a lower genetic predisposition for JIA.

KEYWORDS

juvenile idiopathic arthritis (JIA), MoBa, fish, contaminants, heavy metals, polygenic risk score, gene-environment interaction, sex differences

1 Introduction

Juvenile idiopathic arthritis (JIA), the most common inflammatory rheumatic disease of childhood, manifests as arthritis before the age of 16 years which persists more than six weeks, and without an apparent cause. It consists of seven heterogeneous subgroups, reflecting the complex interplay between genetic predisposition and environmental influences that contribute to the diverse clinical manifestations (1). Known genetic variants are estimated to account for 13-25% of the risk for JIA, while the remaining risk is attributed to environmental factors and their interaction with genetic predisposition (1, 2). Limited high-quality data and modest sample sizes have constrained prior attempts to pinpoint environmental risk and protective factors (3). Furthermore, despite JIA being more prevalent in girls than in boys (4), few studies have investigated this sex disparity, which is important for understanding the underlying pathomechanisms of disease development.

Diet is an example of an environmental factor that remains underexplored in relation to JIA risk (3). Results from a Swedish prospective cohort study showed that fish intake more than once per week during pregnancy and the first year of life was associated with increased risk of JIA, which was mainly attributed to high heavy metal exposure (5).

Among the environmental contributors, heavy metals like mercury and cadmium, and persistent organic pollutants (POPs), have emerged as potential triggers of autoimmunity (6–9). Mercury is associated with subclinical autoimmunity in humans through the production of autoantibodies and cytokines (10–13), while in individuals with a genetic predisposition, cadmium may exacerbate autoimmunity (14) and increase the risk of rheumatoid arthritis (RA) (15, 16). Furthermore, exposure to POPs has also been linked to autoimmune diseases, with research suggesting increased risk of celiac disease, especially in girls (8), and of RA (17).

Diet serves as a major source of these contaminants (18), with seafood being a significant contributor to mercury (19) and shellfish contributing to cadmium exposure (20). Individuals consuming high amounts of seafood are also at greater risk of POPs exposure (21, 22). It has been suggested that diseases with a sex disparity should be investigated for environmental risk factors like contaminant exposure, as differences in vulnerability and susceptibility between the sexes may account for the prevalence disparities (23).

Our primary aim was to explore the association between seafood intake and dietary environmental contaminant exposure during pregnancy and JIA risk. Secondary aims included exploring sex disparities and possible interactions between seafood intake and genetic predisposition to JIA.

2 Material and methods

2.1 Study population and design

We used data from the Norwegian Mother, Father, and Child Cohort Study (MoBa), which was linked by national identification (ID) numbers to the individual records in the following population-based health registers: the Norwegian Patient Registry (NPR) and the Medical Birth Registry of Norway (MBRN).

MoBa is a population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health. Participants were recruited from all over Norway from 1999-2008. Of those invited to participate, 41% of women consented. The cohort includes approximately 114,500 children, 95,200 mothers, and 75,200 fathers. The current study is based on version 12 of the quality-assured data files released for research in 2019 (24). Genotype data was available for a subsample of 51,804 children, which is further described under "Genotyping Data, Polygenic Risk Score (PRS) for JIA".

Three questionnaires were sent to the mothers during pregnancy, the second being a semi-quantitative food frequency questionnaire (FFQ). The FFQ was distributed in gestational week 22 and covered the average intake of 255 food items and beverages during the first half of pregnancy (25). The MoBa FFQ has been validated and found to be a reliable tool to estimate intake of

nutrients and foods during pregnancy, including various types of fish and seafood (26, 27). The FFQ was introduced in March 2002 and all pregnancies recruited between 2002 and 2008 are included in our study. Figure 1 outlines the flow of subject for inclusion in our study from the MoBa cohort.

2.2 Outcome

The Norwegian Patient Registry (NPR) contains data with personal ID numbers from all Norwegian public hospitals and specialists with public funding from 2008. We defined a JIA case as having at least two International Classification of Diseases (ICD)-10 codes (≥2 M08, ≥2 M09, or 1 M08 *and* 1 M09). We recently validated this case definition and have found a positive predictive value of 93.4% (28). For cases where the child received their first ICD-10 code in 2021, we accepted a single relevant ICD-10 code (M08 or M09), recognizing that there might have been only one doctor visit before we received our latest updates from NPR in December 2021. Controls were defined as live births that were non-JIA cases. We excluded children with only one registration of ICD-10 code M08 or M09 between 2008-2020 because they might have JIA.

2.3 Exposure variables: intake of seafood and environmental contaminants

We estimated maternal seafood intake and the exposure to dietary contaminants by the FFQ that was developed and validated for pregnant women in MoBa (25).

Food frequencies reported in the FFQ were converted to food amounts (grams/day) using FoodCalc and the Norwegian food table (26). Seafood intake was allocated into five variables, of which three were strictly related to fish intake: 1) oily fish (more than 8% fat, such as herring, mackerel, salmon), 2) lean/semi-oily fish (up to 8% fat, such as cod, haddock, saithe), and 3) total fish (total amount of oily fish and lean/semi-oily fish). In addition, we included 4)

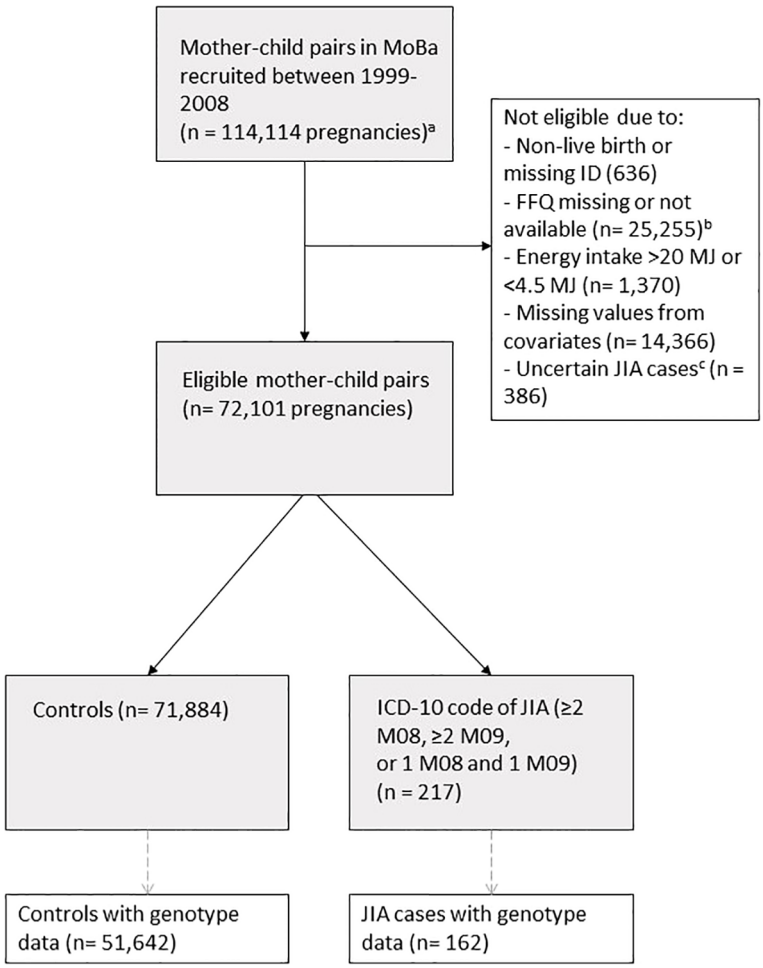


FIGURE 1
Flowchart of study population with exclusion criteria. [a] This number includes siblings. [b] The MoBa FFQ was introduced in 2002. [c] We excluded all controls with a single ICD-10 code (M08 or M09) to rule out potential JIA cases, except for those who received their first code in 2021. For these cases, we accepted a single relevant ICD-10 code, acknowledging that they might have had only one doctor visit before our latest NPR update in 2022.

shellfish intake (capturing crab, shrimp, and mussels), and 5) total seafood (total fish and shellfish, including fish liver, roe, and fish liver/roe spread). We converted these continuous variables (grams/day) into categorical variables in the following way:

We categorized the seafood variables into high intake, defined as equal to or exceeding the 90th percentile (≥P90) of the population, and low intake, defined as less than 90th percentile (<P90). In secondary analyses, the five seafood variables were also divided into quintiles with the lowest group serving as reference. Lastly, because the Norwegian Directorate of Health recommends between 300-450 grams of fish each week (29), an exposure variable was also set at ≥300 grams of fish per week, which was compared to intake <300 grams/week.

The exposure to dietary environmental contaminants was estimated by combining consumption data from the FFQ with concentrations of contaminants in Norwegian food, based on data across various Nordic studies and databases, with the mean or median values from these studies used for the estimation of dietary contaminant exposure. The food contamination data spans several years, corresponding with the period when the FFQ was completed, and is described elsewhere (30, 31). Dietary contaminant exposure was categorized into two main groups: 1) heavy metals and 2) persistent organic pollutants (POPs). Heavy metals included a) mercury, and b) cadmium, while POPs included c) dioxins and dioxin-like (dl) compounds, and d) non-dioxin-like (non-dl) polychlorinated biphenyls (using PCB-153 as a proxy). The exposure to dioxins and dl-compounds is expressed as toxic equivalents (TEQ) when assessing their combined effect (32). The dietary contaminant variables were calculated per kilogram of pre-pregnancy body weight (kg bw), which was self-reported. We analyzed high vs. low intake and across quintiles as defined above.

2.4 Covariates and confounders

Potential confounding factors included maternal education, and parity (categorical variables); maternal age, pre-pregnancy BMI, daily energy intake (continuous variables); and maternal history of inflammatory rheumatic disease (see definition below), parental smoking status, and maternal supplement use during pregnancy (e.g., fish oil, vitamin D, folate) as dichotomous variables (yes/no). Associations with lean/semi-oily and oily fish were mutually adjusted due to their correlated intake.

When analyzing dietary environmental contaminants, we included the child's birth year from the Medical Birth Registry as a possible confounder because contaminant levels in fish may have varied over the years, and because the cumulative risk of JIA increases with the child's age. Information about region of birth was also obtained from the Medical Birth Registry.

Mother's history of inflammatory rheumatic diseases was obtained via linkage to NPR and included following ICD-10 codes: M05, M06, M07, M08, M09, M30, M31, M32, M33, M34, M35, M45, M46, and L94.

2.5 Genotyping data, polygenic risk score for JIA

In MoBa, umbilical cord blood samples were collected at birth and DNA was stored at the Norwegian Institute of Public Health (33). Genotyping was carried out over several years through various research projects (34). MoBaPsychGen genotype quality control (QC) pipeline was developed to manage the complex relationships within the cohort. This pipeline includes steps for pre-imputation QC, phasing, imputation, and post-imputation QC, and it accounts for array and batch effects (35).

We focused on individuals of European ancestry, identified by visually comparing the first seven genetic principal components (PCs) to those from unrelated samples in the 1000 Genomes phase 1 project (35). Related individuals with a kinship coefficient >0.05 had one member excluded, prioritizing the retention of JIA cases, with other exclusions made randomly.

To estimate the genetic risk for JIA, we calculated PRSs using data from a genome-wide association study (GWAS) of JIA (36). The calculation was done using PRSice version 2.3.3 (37), applying different P-value thresholds as 5E-8, 1E-6, 1E-5, 1E-4, 1E-3, 1E-2, 5E-2, 1E-1, 5E-1, and 1. We then extracted the first PC of PRSs across all P-value thresholds, following a widely used method (38). The standardized PRS was then converted into a binary variable with cut-off at 0, of which the PRS <0 was regarded as "low", whereas the PRS ≥0 was regarded as "high".

2.6 Statistical analysis

Stata V.17.0 statistical software (StataCorp) and R version 4.2.3 (39) were used to conduct all statistical analyses. Characteristics of high vs low consumers of fish were reported as mean (SD) or median (IQR), as appropriate for continuous variables and by distribution differences (counts and percentages) for categorical variables. We used multiple logistic regression to examine the associations between seafood intake, dietary environmental contaminant exposure and risk of JIA. All associations are reported as odds ratios (OR) with 95% confidence intervals (CI), and as adjusted ORs (aOR) when adjusted for possible confounding factors listed above. The number of subjects with missing values on covariates was low for both cases (n=40, 18%) and controls (n=14,366, 20%), and all estimates are therefore based on complete case analyses. All analyses were further stratified by sex. In a sensitivity analysis, we included the region of birth (South-East, West, Middle and North), and thus presumably the region where the pregnancy took place, as a possible confounder because research shows a two-fold increased incidence of JIA in northern compared to southern Norway (28) and reports of geographical variations in fish intake (40).

To assess potential interactions between fish intake and genetic predisposition to JIA, we conducted multiple logistic regression analyses with an interaction term between fish intake and PRS. We

10.3389/fimmu.2024.1523990

included the same variables as in the main model except maternal history of inflammatory rheumatic disease to avoid over-adjustment. The Wald test was used to assess statistical significance of an interaction, and a p-value <0.05 was regarded as significant. We further investigated the interaction between fish intake and PRS by calculating the products of fish intake and dummy variables of each PRS group and replacing the interaction term in the multiple logistic regression with those products. This allowed us to estimate the association between fish intake and JIA in the low and high PRS groups separately. We used this model to visualize the relationship between fish intake and JIA in both groups predicting JIA risk in a simulated dataset of n = 200. As an additional test for interactions, we applied a case-only analysis by testing for associations between seafood intake and PRS in the cases only (41).

3 Results

3.1 Study sample characteristics

Our final analytical sample included 72,110 mother-child pairs; 217 children with JIA were identified (Figure 1). Of JIA cases, there were 139 (64.1%) girls and 78 (35.9%) boys. The median weekly maternal fish intake was 218 grams. Baseline characteristics categorized by high (≥P90) vs. low (<P90) total fish intake are shown in Table 1.

TABLE 1 Baseline characteristics categorized by high and low total fish intake in 72,101 MoBa participants 2002-2008.

Characteristics	High total fish intake (P≥90)*	Low total fish intake (<P90) *
Population	7,209 (10.0)	64,892 (90.0)
Maternal age at delivery, years, mean (SD)	31.0 (4.8)	30.3 (4.5)
Maternal education		
High school or less	2,722 (37.8)	21,839 (33.7)
College, up to 4 years	2,675 (37.1)	27,202 (42.0)
College, more than 4 years	1,812 (25.1)	15,851 (24.4)
Maternal pre-pregnancy BMI, mean (SD)	24.0 (4.4)	24.1 (4.3)
Maternal parity		
0	3,097 (43.0)	29,714 (46.0)
1	2,487 (34.5)	23,185 (36)
2 or more	1,624 (22.5)	11,993 (18.5)
Inflammatory rheumatic disease in mother		
Yes	209 (2.9)	1,853 (2.9)
No	7,000 (97.1)	63,039 (97.1)
Maternal daily caloric intake, kcal, median (IQR)	2462 (2053, 2939)	2207 (1866, 2620)
Maternal smoking status during pregnancy		
Yes	642 (8.9)	5,048 (7.8)
No	6,567 (91.1)	59,844 (92.2)
Paternal smoking status		
Yes	1448 (20.1)	12,624 (19.5)
No	5761 (79.9)	52,268 (80.6)
Dietary supplement use during pregnancy		
Yes	6,167 (85.6)	56,032 (86.4)
No	1,042 (14.5)	8,860 (13.7)
Region of birth		
South-East	3,365 (46.7)	35,794 (55.2)
West	1,955 (27.1)	16,132 (24.9)
Middle	1,207 (16.7)	9,296 (14.3)
North	682 (9.5)	3,670 (5.7)

*High is defined as equal to or above 90th percentile, while low is defined as below 90th percentile.
Numbers are n (%), mean (SD) or median (IQR).

3.2 Seafood intake and JIA

High vs low intake of lean/semi-oily fish during pregnancy was associated with JIA (aOR 1.51, 95% CI 1.02-2.22) (Table 2). After adjusting for region of birth, the confidence interval included 1 (aOR 1.45, 95% CI 0.99-2.18) (Supplementary Table 1). Additional results with region of birth as a covariate are presented in Supplementary Table 1-Supplementary Table 2. We found no other evidence of associations between high vs low intake of other seafood variables and JIA risk (Table 2).

After sex-stratification, we found an association with lean/semi-oily fish intake among boys (aOR 2.07, 95% CI 1.17-3.66), but not in girls (Table 2 and Supplementary Table 1). Similarly, high shellfish intake was associated with increased risk among boys (aOR 1.86, 95% CI 1.02-3.38), but not girls (Table 2). Additionally, consuming fish ≥300 vs. <300 grams/week during pregnancy, regardless of fat content, was linked to higher odds of JIA in boys (aOR 1.92, 95% CI: 1.22-3.04), but not in girls (Supplementary Table 3). When analyzing by quintiles, no other convincing evidence of associations were observed (Supplementary Table 4).

3.3 Interactions between fish intake and polygenic risk score

The following results are based on a smaller sample than our main analyses (controls n= 51,642, JIA case n= 162) due to lack of genetic data on all observations. To account for this, we ran the main analyses on the smaller dataset as a sensitivity analysis, with the results provided in Supplementary Table 5.

We found evidence of an interaction between total fish intake and PRS (aOR 0.33, 95% CI 0.12-0.90, p-value 0.03), but not with the other seafood variables (Supplementary Table 6). The association between total fish intake and JIA was only apparent in

10.3389/fimmu.2024.1523990

TABLE 2 Overall and sex-stratified associations between high vs. low seafood intake and JIA.

	All (controls n= 71,884, JIA cases n= 217)		Boys (controls n= 36,784 and JIA cases n= 78)		Girls (controls n= 35,100, JIA cases n= 139)	
	Unadjusted OR (95% CI)	aOR[a] (95% CI)	Unadjusted OR (95% CI)	aOR[a] (95% CI)	Unadjusted OR (95% CI)	aOR[a] (95% CI)
High total fish intake						
<90th percentile	Ref	Ref	Ref	Ref	Ref	Ref
≥90th percentile (≥423.5 grams/week)	1.02 (0.65-1.58)	1.02 (0.65-1.59)	1.78 (0.98-3.23)	1.80 (0.98-3.31)	0.63 (0.32-1.24)	0.63 (0.32-1.24)
High lean/semioily fish intake						
<90th percentile	Ref	Ref	Ref	Ref	Ref	Ref
≥90th percentile (≥249.5 grams/week)	**1.50 (1.03-2.20)**	**1.51 (1.02-2.22)**	**2.13 (1.21-3.75)**	**2.07 (1.17-3.66)**	1.18 (0.70-1.99)	1.21 (0.72-2.06)
High oily fish intake						
<90th percentile	Ref	Ref	Ref	Ref	Ref	Ref
≥90th percentile (≥156 grams/week)	0.81 (0.50-1.32)	0.80 (0.49-1.31)	1.45 (0.76-2.74)	1.36 (0.71-2.62)	0.49 (0.23-1.04)	0.49 (0.23-1.06)
High shellfish intake						
<90th percentile	Ref	Ref	Ref	Ref	Ref	Ref
≥90th percentile (≥65 grams/week)	1.12 (0.73-1.71)	1.14 (0.74-1.74)	**1.83 (1.01-3.33)**	**1.86 (1.02-3.38)**	0.76 (0.41-1.41)	0.78 (0.42-1.44)
High seafood intake						
<90th percentile	Ref	Ref	Ref	Ref	Ref	Ref
≥90th percentile (≥492 grams/week)	0.91 (0.58-1.49)	0.92 (0.57-1.46)	1.62 (0.87-3.00)	1.64 (0.88-3.07)	0.56 (0.27-1.13)	0.55 (0.27-1.14)

[a]Adjusted for maternal age, education, pre-pregnancy BMI, parity, daily caloric intake, history of inflammatory rheumatic disease in mother, parental smoking status during pregnancy and supplement use during pregnancy. When lean/semioily fish is the main exposure, it is also adjusted for oily fish intake, and vice-versa. Bold text indicates statistically significant results.

TABLE 3 Associations between high seafood intake and JIA risk in groups of high or low genetic risk (PRS of JIA).

Exposure	PRS group[a]	aOR[b] (95% CI)
High total fish	Low	**2.26 (1.08-4.71)**
	High	0.75 (0.38-1.49)
High lean/semioily fish	Low	**2.23 (1.06-4.66)**
	High	1.14 (0.63-2.05)
High oily fish	Low	0.65 (0.20-2.11)
	High	0.84 (0.43-1.61)
High shellfish	Low	1.65 (0.74-3.71)
	High	1.38 (0.81-2.36)
High seafood	Low	1.67 (0.74-3.77)
	High	0.82 (0.43-1.58)

[a]The standardized PRS was converted into a binary variable with cut-off at 0, of which the PRS <0 was regarded as "low", whereas the PRS ≥0 was regarded as "high".
[b]Adjusted for: maternal age, education, pre-pregnancy BMI, parity, daily caloric intake, parental smoking status during pregnancy, supplement use during pregnancy, high PRS and PCs 1-10. When lean/semioily fish is the main exposure, it is also adjusted for oily fish intake, and vice-versa. (controls n= 51,642, JIA cases n= 162).
Bold text indicates statistically significant results.

the low PRS group (aOR 2.26, 95% CI 1.08-4.71) (Table 3 and Figure 2). Furthermore, we also found an association between lean/semi-oily fish and JIA in the low PRS group (aOR 2.23, 95% CI 1.06-4.66), but not with the other seafood variables (Table 3 and Supplementary Figure 1-Supplementary Figure 2). A case-only design was used to test the interaction between fish intake and PRS, which further confirmed the findings from the case-control analyses: the high total fish intake was negatively associated with PRS in the cases, whereas none of the other seafood variables reached statistical significance (Supplementary Table 7).

3.4 Estimated environmental contaminants and JIA

We found no evidence of associations between estimated dietary intake of environmental contaminants and risk of JIA, whether analyzed by high vs. low intake (Table 4) or by quintiles (Supplementary Table 8).

After sex-stratification, we found a positive association between non-dl PCBs and JIA in boys (aOR 2.24, 95% CI 1.03-4.86), when comparing a dietary exposure corresponding to the 4th quintile to the 1st quintile (Supplementary Table 8). Among girls, being in the

10.3389/fimmu.2024.1523990

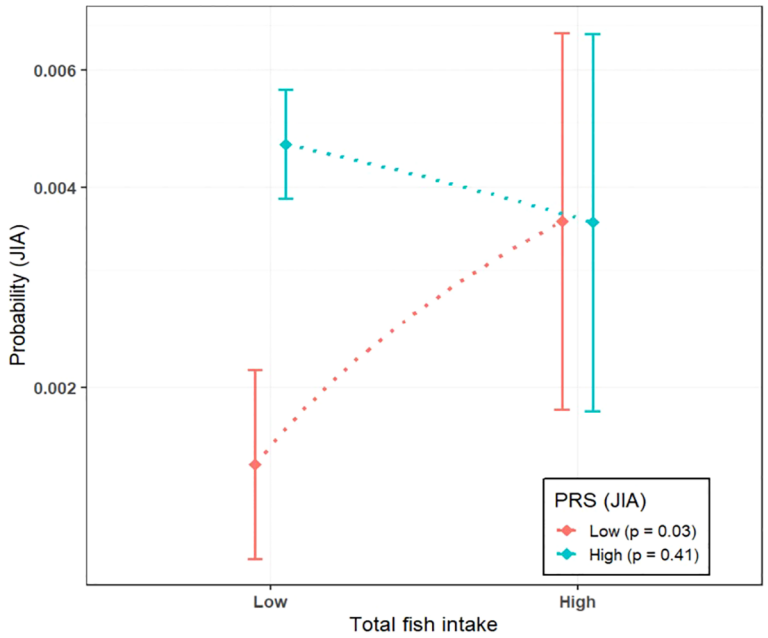


FIGURE 2
Association between total fish intake and JIA risk grouped by high (=>0) and low (<0) polygenic risk score (PRS) for JIA. P-values indicate the significance of the associations between fish intake and JIA risk within each PRS group.

TABLE 4 Overall and sex-stratified associations between high vs. low dietary contaminant exposure[a] and JIA.

	All (controls n= 71,884, JIA cases n= 217)		Boys (controls n= 36,784 and JIA cases n= 78)		Girls (controls n= 35,100, JIA cases n= 139)	
	Unadjusted OR (95% CI)	aOR[b] (95% CI)	Unadjusted OR (95% CI)	aOR[b] (95% CI)	Unadjusted OR (95% CI)	aOR[b] (95% CI)
Mercury						
<90th percentile	Ref	Ref	Ref	Ref	Ref	Ref
≥90th percentile (≥0.3 ug/kg bw/week)	0.91 (0.58-1.45)	0.92 (0.57-1.46)	1.17 (0.58-2.35)	1.22 (0.60-2.48)	0.77 (0.42-1.44)	0.76 (0.41-1.42)
Cadmium						
<90th percentile	Ref	Ref	Ref	Ref	Ref	Ref
≥90th percentile (≥2.1 ug/kg bw/week)	1.39 (0.94-2.06)	1.42 (0.94-2.14)	1.47 (0.78-2.79)	1.51 (0.77-2.99)	1.35 (0.82-2.21)	1.36 (0.80-2.29)
Dioxins and dioxin-like (dl) compounds						
<90th percentile	Ref	Ref	Ref	Ref	Ref	Ref
≥90th percentile (≥7.5 pg TEQ/kg bw/week)	0.86 (0.54-1.38)	0.83 (0.51-1.34)	1.46 (0.77-2.77)	1.45 (0.74-2.82)	0.56 (0.27-1.14)	0.52 (0.25-1.08)
Non-dioxin-like PCBs (PCB-153)						
<90th percentile	Ref	Ref	Ref	Ref	Ref	Ref
≥90th percentile (≥13.2 pg/kg bw/week)	0.91 (0.58-1.45)	0.90 (0.56-1.43)	1.03 (0.49-2.14)	1.02 (0.48-2.14)	0.85 (0.47-1.54)	0.83 (0.46-1.51)

[a]Contaminants were estimated by combining consumption data from the FFQ with concentrations of contaminants in Norwegian food.
[b]Adjusted for maternal age, education, pre-pregnancy BMI, parity, daily caloric intake, history of inflammatory rheumatic disease in mother, parental smoking status during pregnancy, supplement use during pregnancy and the child's birth year.

5th quintile of either dl-compound or non-dl PCB intake, was negatively associated with risk of JIA (aOR 0.40, 95% CI 0.20-0.79 and aOR 0.44, 95% CI 0.23-0.83; Supplementary Table 8).

4 Discussion

In this large population-based study, we found a modestly increased risk of JIA associated with high maternal intake of lean/semi-oily fish (approximately 250 grams or more per week) during pregnancy. No clear associations were found between JIA and overall maternal intake of fish, oily fish, shellfish, or seafood intake. Sex-stratified analyses suggested a stronger positive association between high maternal seafood intake and JIA risk in boys. For instance, an intake of >300 grams of fish per week as recommended by the Norwegian Directorate of Health (29), was linked to increased risk of JIA in boys but not in girls. We observed no clear associations with estimated maternal dietary contaminant exposures. The risk associated with total fish intake depended on genetic predisposition: high fish intake significantly affected JIA risk only in individuals with a low genetic predisposition to JIA.

Our results are partly in line with a Swedish study (5), which found positive associations between fish intake of more than once per week during pregnancy and JIA risk, although our effect sizes were of substantially lower magnitude. The Swedish study did not specify portion sizes, complicating direct comparisons. Furthermore, our study specifically associates lean/semi-oily fish with increased JIA risk, while the Swedish study identified the strongest association with total fish intake without distinguishing between fish varieties (5).

We found no evidence of robust associations between exposure to dietary environmental contaminants and risk of JIA. This differs from the Swedish study which attributed the heightened risk of JIA to increased heavy metal exposure, including mercury, through fish intake (5), and another study showing that prenatal exposure to environmental contaminants can alter the cord serum metabolome, potentially increasing the risk of immune-mediated diseases such as JIA (42). Despite seafood accounting for 88% of total dietary mercury exposure – with lean fish contributing to more than half of this - as well as being a considerable source of other contaminants (20, 43, 44), we found no evidence that it contributed to JIA risk in MoBa. In fact, our sex-stratified analyses show an inverse relationship between exposure to POPs and JIA in girls. Unlike the Swedish study, which measured blood concentrations, our study relies on self-reported dietary data, but includes a much larger sample size (217 vs. 41 JIA cases) (5).

JIA is more prevalent in girls than boys (4), yet our study suggests that high seafood intake is more strongly associated with JIA risk in boys. Sex-stratified analyses showed no indication of increased risk of JIA when comparing high vs low intake of seafood and contaminant exposure (except lean/semi-oily fish and cadmium) in girls, on the contrary, estimates indicated a lower risk of JIA with high intake. In contrast, for boys, all associations indicated an increased risk of JIA.

Most studies on sex disparities in pediatric illnesses do not explore underlying causes (45), making our sex-stratified analyses valuable for addressing this knowledge gap. Although estrogen levels are often suggested as a cause for the higher prevalence of autoimmunity in women, the low and stable levels during childhood suggest other mechanisms (4). The varying patterns of JIA risk between boys and girls with seafood intake may be due to lack of statistical power given the sample size (girls, n = 139, boys, n = 78), and the results should be interpreted cautiously. The inverse relationship between POP exposure and JIA risk in girls observed in our study may not be directly linked to POPs, but could reflect a spurious association with oily fish, which was estimated to have a protective association in girls. This protective association may be related to nutrients in oily fish rather than POPs. A study on diabetes type 1 observed similar findings (46). A separate MoBa study on prenatal exposure to POPs showed immunosuppressive effects (32), which could potentially explain a protective association in girls. Inherent biological differences may also influence these sex-specific trends.

Gene-environment interaction analyses suggest that genetic predisposition modifies the effect of fish intake on JIA risk, and vice versa. Specifically, fish intake had a stronger estimated association with JIA risk in individuals with low genetic predisposition, while its impact was estimated as less pronounced in those with a high genetic risk. Our previous findings show that the PRS is more strongly associated with JIA in girls than in boys, with a higher proportion of female JIA cases having a standardized PRS >0 (submitted for publication)[1]. This might explain why we observe a stronger association between fish intake and JIA risk in boys, as male JIA cases, on average, have a lower genetic risk of JIA.

Our study's strengths include its prospective design, comprehensive data collection with genetic liability, a large study population, and linkage to national registries, ensuring minimal loss to follow-up. A significant and novel strength is the incorporation of a PRS within a subset of our cohort, enabling us to study gene-environment interactions in JIA. By sex-stratification, we discerned variations in risk estimates between boys and girls. To our knowledge, this is the largest population-based prospective cohort study exploring environmental risk factors for JIA, identifying 217 cases.

While including more JIA cases than in previous studies, the sample size remains the main limitation of the study, as it reduces the power to detect small effects, especially in stratified analyses and for the subset with genotype data. We also lack data on JIA subtypes, which is important given the disease's heterogeneity; different subtypes may have distinct pathomechanisms or vulnerabilities. We did not exclude controls with other systemic autoimmune diseases, potentially diluting the observed effects. Additionally, while the recruitment into MoBa was population-based, the cohort is not fully representative of the general population (47). For instance, the homogenous ethnic background of MoBa participants (48) may limit the generalizability of our findings to more diverse populations. The

[1]Haftorn KL, Rudsari HK, Jaholkowski PP, Dåstøl VØ, Hestetun SV, Andreassen OA, et al. *Nonlinearity and sex differences in the performance of a polygenic risk score for juvenile idiopathic arthritis.* (2024).

self-reported dietary data may result in exposure misclassification as the FFQ provide rough estimates, even though it has been validated (26). We cannot study exact dietary intake for the second half of the pregnancy as the FFQ was completed in week 22, however, we assume consistent dietary patterns throughout the pregnancy. Additionally, we lack measured blood concentration of contaminants. Our contamination estimates rely on broader Nordic averages rather than location-specific data, so this approach may not adequately capture exposure differences across Norway, especially in areas of higher contamination, highlighting the need for future research to measure blood concentrations. Although we adjusted for potential confounders, residual confounding cannot be ruled out due to the observational nature of the study. Lastly, since NPR data begins in 2008, JIA cases diagnosed and in remission between 2002-2008 may be missing. Some of the older-diagnosed JIA cases are also missing, because follow up ended in 2021.

In conclusion, we observed an increased risk of JIA in children whose mothers consumed high amounts of lean/semi-oily fish during pregnancy, particularly in boys. Despite lean fish being an important source of dietary mercury exposure, the heightened JIA risk was not explained by contaminant exposure in our study. Our findings also suggest a stronger association between fish intake and JIA in those with a low genetic predisposition to JIA. Further studies are warranted to explore the underlying mechanisms of seafood and JIA, as definitive causation cannot be inferred. This includes more precise assessments of contaminant exposure via blood samples, and the need to clarify the observed sex differences and genetic interactions.

Data availability statement

Data from the Norwegian Mother, Father and Child Cohort Study and the Medical Birth Registry of Norway used in this study are managed by the national health register holders in Norway (Norwegian Institute of public health) and can be made available to researchers, provided approval from the Regional Committees for Medical and Health Research Ethics (REC), compliance with the EU General Data Protection Regulation (GDPR) and approval from the data owners. The consent given by the participants does not open for storage of data on an individual level in repositories or journals. Researchers who want access to data sets for replication should apply through helsedata.no. Access to data sets requires approval from The Regional Committee for Medical and Health Research Ethics in Norway and an agreement with MoBa.

Ethics statement

The establishment of MoBa and initial data collection was based on a license from the Norwegian Data Protection Agency and approval from The Regional Committees for Medical and Health Research Ethics. The MoBa cohort is currently regulated by the Norwegian Health Registry Act. The current study, part of the MoBaRheuma project, was approved by The Regional Committees for Medical and Health Research Ethics (REK), which includes linkages with NPR and MBRN. REK ref. nr. 2019/1222.

Author contributions

VD: Writing – original draft, Writing – review & editing, Conceptualization, Formal analysis, Methodology, Visualization. KH: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – review & editing. HR: Formal analysis, Methodology, Validation, Writing – review & editing. PJ: Formal analysis, Methodology, Supervision, Writing – review & editing. KS: Methodology, Writing – review & editing. SH: Conceptualization, Methodology, Resources, Writing – review & editing. CW: Methodology, Writing – review & editing. LR: Methodology, Writing – review & editing. OA: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing. AB: Conceptualization, Methodology, Resources, Writing – original draft, Writing – review & editing. IC: Conceptualization, Formal analysis, Methodology, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing. HS: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Our work has been financed through several public sources: Foundation DAM (grant number 2023/FO426544), South-Eastern Norway Regional Health Authority (grant number HSO/2023070) and Norwegian Rheumatology Association. LR and CW were supported by the intramural research program of the National Institute of Environmental Health Sciences, National Institutes of Health. The study was partly funded by the Research Council of Norway through its Centres of Excellence funding scheme (project No 262700).

Acknowledgments

The Norwegian Mother, Father and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. We are grateful to all the participating families in Norway who take part in this on-going cohort study and the registries (MBRN and NPR) for making data available. We thank the Norwegian Institute of Public Health (NIPH) for generating high-quality genomic data. This research is part of the HARVEST collaboration, supported by the Research Council of Norway (#229624). We also thank the NORMENT Centre for providing genotype data, funded by the

Research Council of Norway (#223273), South East Norway Versjon 7.03 Health Authorities and Stiftelsen Kristian Gerhard Jebsen. We further thank the Center for Diabetes Research, the University of Bergen for providing genotype data and performing quality control and imputation of the data funded by the ERC AdG project SELECTionPREDISPOSED, Stiftelsen Kristian Gerhard Jebsen, Trond Mohn Foundation, the Research Council of Norway, the Novo Nordisk Foundation, the University of Bergen, and the Western Norway Health Authorities. We would also like to thank Professor Håkon K. Gjessing for valuable discussions and guidance regarding the gene-environment interaction analyses. Part of this work was presented as an oral abstract at the EULAR Annual European Congress of Rheumatology conference in Vienna 2024 (49).

Conflict of interest

OA has received consulting fees from Cortechs.ai and Precision Health AS and declares future stock options with Cortechs.ai and Precision Health AS. OA has also received speaker fees from Lundbeck, Janssen, Otsuka, and Sunovion.

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.

Generative AI statement

The authors declare that Gen AI was used in the creation of this manuscript. ChatGPT was solely used to improve the language clarity and precision, as English is not the primary language of the authors. It was not used for data analysis or any other purpose beyond improving language.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1523990/full#supplementary-material

References

1. Martini A, Lovell DJ, Albani S, Brunner HI, Hyrich KL, Thompson SD, et al. Juvenile idiopathic arthritis. *Nat Rev Dis Primers.* (2022) 8:5. doi: 10.1038/s41572-021-00332-8

2. Horton DB, Shenoi S. Review of environmental factors and juvenile idiopathic arthritis. *Open Access Rheumatol.* (2019) 11:253–67. doi: 10.2147/OARRR.S165916

3. Clarke SLN, Mageean KS, Maccora I, Harrison S, Simonini G, Sharp GC, et al. Moving from nature to nurture: a systematic review and meta-analysis of environmental factors associated with juvenile idiopathic arthritis. *Rheumatol (Oxford).* (2022) 61:514–30. doi: 10.1093/rheumatology/keab627

4. Cattalini M, Soliani M, Caparello MC, Cimaz R. Sex differences in pediatric rheumatology. *Clin Rev Allergy Immunol.* (2019) 56:293–307. doi: 10.1007/s12016-017-8642-3

5. Kindgren E, Guerrero-Bosagna C, Ludvigsson J. Heavy metals in fish and its association with autoimmunity in the development of juvenile idiopathic arthritis: a prospective birth cohort study. *Pediatr Rheumatol Online J.* (2019) 17:33. doi: 10.1186/s12969-019-0344-3

6. Pollard KM, Cauvi DM, Mayeux JM, Toomey CB, Peiss AK, Hultman P, et al. Mechanisms of environment-induced autoimmunity. *Annu Rev Pharmacol Toxicol.* (2021) 61:135–57. doi: 10.1146/annurev-pharmtox-031320-111453

7. Pollard KM, Christy JM, Cauvi DM, Kono DH. Environmental xenobiotic exposure and autoimmunity. *Curr Opin Toxicol.* (2018) 10:15–22. doi: 10.1016/j.cotox.2017.11.009

8. Gaylord A, Trasande L, Kannan K, Thomas KM, Lee S, Liu M, et al. Persistent organic pollutant exposure and celiac disease: A pilot study. *Environ Res.* (2020) 186:109439. doi: 10.1016/j.envres.2020.109439

9. Kharrazian D. Exposure to environmental toxins and autoimmune conditions. *Integr Med (Encinitas).* (2021) 20:20–4.

10. Alves MF, Fraiji NA, Barbosa AC, De Lima DS, Souza JR, Dórea JG, et al. Fish consumption, mercury exposure and serum antinuclear antibody in Amazonians. *Int J Environ Health Res.* (2006) 16:255–62. doi: 10.1080/09603120600734147

11. Nyland JF, Fillion M, Barbosa F Jr., Shirley DL, Chine C, Lemire M, et al. Biomarkers of methylmercury exposure immunotoxicity among fish consumers in Amazonian Brazil. *Environ Health Perspect.* (2011) 119:1733–8. doi: 10.1289/ehp.1103741

12. Crowe W, Allsopp PJ, Watson GE, Magee PJ, Strain JJ, Armstrong DJ, et al. Mercury as an environmental stimulus in the development of autoimmunity - A systematic review. *Autoimmun Rev.* (2017) 16:72–80. doi: 10.1016/j.autrev.2016.09.020

13. Somers EC, Ganser MA, Warren JS, Basu N, Wang L, Zick SM, et al. Mercury exposure and antinuclear antibodies among females of reproductive age in the United States: NHANES. *Environ Health Perspect.* (2015) 123:792–8. doi: 10.1289/ehp.1408751

14. Popov Aleksandrov A, Mirkov I, Tucovic D, Kulas J, Zeljkovic M, Popovic D, et al. Immunomodulation by heavy metals as a contributing factor to inflammatory diseases and autoimmune reactions: Cadmium as an example. *Immunol Lett.* (2021) 240:106–22. doi: 10.1016/j.imlet.2021.10.003

15. Chen L, Sun Q, Peng S, Tan T, Mei G, Chen H, et al. Associations of blood and urinary heavy metals with rheumatoid arthritis risk among adults in NHANES, 1999-2018. *Chemosphere.* (2022) 289:133147. doi: 10.1016/j.chemosphere.2021.133147

16. Joo SH, Lee J, Hutchinson D, Song YW. Prevalence of rheumatoid arthritis in relation to serum cadmium concentrations: cross-sectional study using Korean National Health and Nutrition Examination Survey (KNHANES) data. *BMJ Open.* (2019) 9:e023233. doi: 10.1136/bmjopen-2018-023233

17. Lee DH, Steffes M, Jacobs DR. Positive associations of serum concentration of polychlorinated biphenyls or organochlorine pesticides with self-reported arthritis, especially rheumatoid type, in women. *Environ Health Perspect.* (2007) 115:883–8. doi: 10.1289/ehp.9887

18. Costa LG. Contaminants in fish: risk-benefit considerations. *Arh Hig Rada Toksikol.* (2007) 58:367–74. doi: 10.2478/v10004-007-0025-3

19. Mahaffey KR, Clickner RP, Bodurow CC. Blood organic mercury and dietary mercury intake: National Health and Nutrition Examination Survey, 1999 and 2000. *Environ Health Perspect.* (2004) 112:562–70. doi: 10.1289/ehp.6587

20. Ferrari P, Arcella D, Heraud F, Cappé S, Fabiansson S. Impact of refining the assessment of dietary exposure to cadmium in the European adult population. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess.* (2013) 30:687–97. doi: 10.1080/19440049.2013.777161

21. Kovner K. *Persistent organic pollutants: A global issue, A global response.* Washington, DC, USA: United States ENvironmental Protection Agency (EPA (2002). Available at: https://www.epa.gov/international-cooperation/persistent-organic-pollutants-global-issue-global-response (Accessed March 14, 2024).

22. (WHO) WHO. *Food safety: Persistent organic pollutants (POPs)*. Geneva, Switzerland: Nutrition and Food Safety (NFS)WHO (2020). Available at: https://www.who.int/news-room/questions-and-answers/item/food-safety-persistent-organic-pollutants-(pops) (Accessed March 15, 2024).

23. Butter ME. Are women more vulnerable to environmental pollution? *J Hum Ecol*. (2006) 20:221–6. doi: 10.31901/24566608.2006/20.03.12

24. Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, et al. Cohort profile update: the norwegian mother and child cohort study (MoBa). *Int J Epidemiol*. (2016) 45:382–8. doi: 10.1093/ije/dyw029

25. Meltzer HM, Brantsaeter AL, Ydersbond TA, Alexander J, Haugen M. Methodological challenges when monitoring the diet of pregnant women in a large study: experiences from the Norwegian Mother and Child Cohort Study (MoBa). *Matern Child Nutr*. (2008) 4:14–27. doi: 10.1111/j.1740-8709.2007.00104.x

26. Brantsaeter AL, Haugen M, Alexander J, Meltzer HM. Validity of a new food frequency questionnaire for pregnant women in the Norwegian Mother and Child Cohort Study (MoBa). *Matern Child Nutr*. (2008) 4:28–43. doi: 10.1111/j.1740-8709.2007.00103.x

27. Vejrup K, Brandlistuen RE, Brantsæter AL, Knutsen HK, Caspersen IH, Alexander J, et al. Prenatal mercury exposure, maternal seafood consumption and associations with child language at five years. *Environ Int*. (2018) 110:71–9. doi: 10.1016/j.envint.2017.10.008

28. Hestetun SV, Rudsari HK, Jaholkowski P, Shadrin A, Haftorn KL, Andersen S, et al. Incidence and genetic risk of juvenile idiopathic arthritis in Norway by latitude. *Arthritis & rheumatology* (Hoboken, N.J.) (2024). doi: 10.1002/art.43040

29. *Helsedirektoratet. Kostrådene. Fisk til middag to til tre ganger i uken* (2016). Available online at: https://www.helsedirektoratet.no/faglige-rad/kostradene-og-naeringsstoffer/kostrad-for-befolkningenfisk-til-middag-to-til-tre-ganger-i-uken-praktisk-informasjon (Accessed March 19, 2024).

30. Jenssen MT, Brantsæter AL, Haugen M, Meltzer HM, Larssen T, Kvalem HE, et al. Dietary mercury exposure in a population with a wide range of fish consumption–self-capture of fish and regional differences are important determinants of mercury in blood. *Sci Total Environ*. (2012) 439:220–9. doi: 10.1016/j.scitotenv.2012.09.024

31. Kvalem HE, Knutsen HK, Thomsen C, Haugen M, Stigum H, Brantsaeter AL, et al. Role of dietary patterns for dioxin and PCB exposure. *Mol Nutr Food Res*. (2009) 53:1438–51. doi: 10.1002/mnfr.200800462

32. Stølevik SB, Nygaard UC, Namork E, Haugen M, Meltzer HM, Alexander J, et al. Prenatal exposure to polychlorinated biphenyls and dioxins from the maternal diet may be associated with immunosuppressive effects that persist into early childhood. *Food Chem Toxicol*. (2013) 51:165–72. doi: 10.1016/j.fct.2012.09.027

33. (NIPH) NIoPH. *Protocols for the norwegian mother, father and child cohort study (MoBa)*. Norwegian Institute of Public Health: Norwegian Institute of Public Health (2019) (Oslo, Norway). Available at: https://www.fhi.no/en/publ/2012/protocols-for-moba/ (Accessed May 20, 2024).

34. Høgåsen G. *MoBa genetics* (2024). Available online at: https://github.com/folkehelseinstituttet/mobagen (Accessed May 20, 2024).

35. Corfield E, Shadrin A, Frei O, Rahman Z, Lin A, Athanasiu L, et al. Genetic profile of the norwegian mother, father, and child cohort study (Moba): results from the mobapsychgen pipeline. *Eur Neuropsychopharmacol*. (2022) 63:e292–e3. doi: 10.1016/j.euroneuro.2022.07.517

36. López-Isac E, Smith SL, Marion MC, Wood A, Sudman M, Yarwood A, et al. Combined genetic analysis of juvenile idiopathic arthritis clinical subtypes identifies novel risk loci, target genes and key regulatory mechanisms. *Ann Rheum Dis*. (2021) 80:321–8. doi: 10.1136/annrheumdis-2020-218481

37. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience*. (2019) 8. doi: 10.1093/gigascience/giz082

38. Coombes BJ, Ploner A, Bergen SE, Biernacka JM. A principal component approach to improve association testing with polygenic risk scores. *Genet Epidemiol*. (2020) 44:676–86. doi: 10.1002/gepi.22339

39. R Core Team. *R. A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing (2022).

40. Stabell C. Hvem spiser fisk, in: *Statistisk sentralbyrå: Statistisk sentralbyrå* (2017). Available online at: https://www.ssb.no/helse/artikler-og-publikasjoner/hvem-spiser-fisk (Accessed May 13, 2024).

41. Piegorsch WW, Weinberg CR, Taylor JA. Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case-control studies. *Stat Med*. (1994) 13:153–62. doi: 10.1002/sim.4780130206

42. Karthikeyan BS, Hyötyläinen T, Ghaffarzadegan T, Triplett E, Orešič M, Ludvigsson J. Prenatal exposure to environmental contaminants and cord serum metabolite profiles in future immune-mediated diseases. *J Expo Sci Environ Epidemiol*. (2024) 34:647–58. doi: 10.1038/s41370-024-00680-z

43. Vejrup K, Brantsæter AL, Knutsen HK, Magnus P, Alexander J, Kvalem HE, et al. Prenatal mercury exposure and infant birth weight in the Norwegian Mother and Child Cohort Study. *Public Health Nutr*. (2014) 17:2071–80. doi: 10.1017/S1368980013002619

44. Ho QT, Frantzen S, Nilsen BM, Nøstbakken OJ, Azad AM, Duinker A, et al. Congener-specific accumulation of persistent organic pollutants in marine fish from the Northeast Atlantic Ocean. *J Hazard Mater*. (2023) 457:131758. doi: 10.1016/j.jhazmat.2023.131758

45. Piccini P, Montagnani C, de Martino M. Gender disparity in pediatrics: a review of the current literature. *Ital J Pediatr*. (2018) 44:1. doi: 10.1186/s13052-017-0437-x

46. Rignell-Hydbom A, Elfving M, Ivarsson SA, Lindh C, Jönsson BA, Olofsson P, et al. A nested case-control study of intrauterine exposure to persistent organochlorine pollutants in relation to risk of type 1 diabetes. *PloS One*. (2010) 5:e11281. doi: 10.1371/journal.pone.0011281

47. Nilsen RM, Vollset SE, Gjessing HK, Skjaerven R, Melve KK, Schreuder P, et al. Self-selection and bias in a large prospective pregnancy cohort in Norway. *Paediatr Perinat Epidemiol*. (2009) 23:597–608. doi: 10.1111/j.1365-3016.2009.01062.x

48. Corfield EC, Frei O, Shadrin AA, Rahman Z, Lin A, Athanasiu L, et al. The Norwegian Mother, Father, and Child cohort study (MoBa) genotyping data resource: MoBaPsychGen pipeline v. *1. bioRxiv*. (2022) 2022:06.23.496289. doi: 10.1101/2022.06.23.496289

49. Dåstøl VØ, Caspersen IH, Brantsæter AL, Sanner H. Op0316 fish consumption and dietary mercury exposure during pregnancy and jia risk: A population-based nationwide cohort study. *Ann Rheum Dis*. (2024) 83:231. doi: 10.1136/annrheumdis-2024-eular.1927

frontiers | Frontiers in Immunology

TYPE Perspective
PUBLISHED 20 February 2025
DOI 10.3389/fimmu.2025.1479814

OPEN ACCESS

EDITED BY
Esther Erdei,
University of New Mexico Health Sciences Center, United States

REVIEWED BY
Anne Satterthwaite,
University of Texas Southwestern Medical Center, United States
Sergey V. Kozyrev,
Uppsala University, Sweden
Takuma Shibata,
The University of Tokyo, Japan

*CORRESPONDENCE
Valerie M. Lewis
valerie-lewis@omrf.org

RECEIVED 12 August 2024
ACCEPTED 30 January 2025
PUBLISHED 20 February 2025

CITATION
Scofield RH, Wren JD and Lewis VM (2025) The toll like receptor 7 pathway and the sex bias of systemic lupus erythematosus. *Front. Immunol.* 16:1479814. doi: 10.3389/fimmu.2025.1479814

The toll like receptor 7 pathway and the sex bias of systemic lupus erythematosus

R. Hal Scofield[1,2,3], Jonathan D. Wren[4] and Valerie M. Lewis[1,2,3]*

[1]Arthritis & Clinical Immunology Program, Oklahoma Medical Research Foundation, Oklahoma City, OK, United States, [2]Department of Medicine, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK, United States, [3]Research and Medical Services, Oklahoma City US Department of Veterans Health Care System, Oklahoma City, OK, United States, [4]Genes and Human Disease Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK, United States

Systemic lupus erythematosus (SLE) predominately affects women with a ratio of females-to-males of about 9:1. The complement of sex chromosomes may play and important role in the mechanism of the sex bias. Previous work has shown that men with Klinefleter's syndrome (47,XXY) as well as women with 47,XXX are found in excess among SLE patients well as among Sjogren's disease, systemic sclerosis and idiopathic inflammatory myositis. in cells with more than one X chromosome, all but one is inactivated. However, X chromosome inactivation, as mediated by the long noncoding RNA X-inactive specific transcript, or XIST, is not complete with approximately 10% of genes in the non-recombining region of the X chromosome escaping X inactivation. In the TLR7 signaling pathway, both the TLR7 and TLR adaptor interacting with endolysosomal SLC15A4 (TASL) escape X inactivation. Comparing male and female immune cells, there is increased TLR7 signaling related to increased expression of these genes in cells with more than one X chromosome. Cells with more than one X chromosome also express XIST, while cells with one X chromosome do not. XIST, as a source of ligand for TLR7, has also been shown to increase TLR7 signaling. Thus, we propose that both these mechanisms operating in immune cells with more than one X chromosome may act in a mutual way to mediate an X chromosome dose effect for the sex bias of autoimmune disease.

KEYWORDS

systemic lupus erythematosus, sex bias, TLR7, TASL, XIST

Sex bias in lupus

Systemic illness among patients with the rash of lupus erythematosus was first noted by Moriz Kaposi in Vienna during the late 19th century (1). During the remainder of the 19th century and through the middle of the 20th century, the entity of systemic lupus erythematosus was established (2). The bias of this disease to affect women was also noted during this period, with assembled cohorts comprised by ~90% of women (3). This ratio of ~9:1 women to men in

cohorts of SLE has continued to be true into the 21st century with modern epidemiological methods (4). This relationship holds true in all racial and ethnic groups studied.

Sex hormones in lupus

While there are sex hormone differences between SLE patients and matched controls, be they men or women (reviewed in (5)), a fundamental biological explanation for these findings and their relationship to the gender-bias of SLE has not been forthcoming (6). Clearly, some men with SLE have primary hypogonadism. For instance, Mok, et al, found that 5 of 35 men with SLE had low serum testosterone and high luteinizing hormone (LH) while none of 33 control men did (7). The etiology of the hypogonadism in these men was not determined. Higher serum prolactin is also found in both men and women with SLE compared to controls (8, 9). However, men with SLE have the same degree of hypogonadism and low testosterone as do men with other non-female biased chronic illnesses (10), suggesting chronic illness causes hypogonadism in SLE rather than vice versa. Furthermore, at the onset of disease, prior to treatment, there are no sex hormone differences between SLE patients and a matched control population (11).

X chromosome in lupus

Seeking another explanation to the sex bias of SLE, we examined the complement of sex chromosomes, initially among men with SLE. We found that these SLE-affected men were much more likely than matched control men to have Klinefelter's syndrome, that is, 47,XXY (12). Subsequent work found that 47,XXX was found in excess among women with SLE (13). We have also found the rare mosaic, 45XO/46XX/47XXX, is associated with SLE (14), while Turner's syndrome (female 45,XO) was not found in excess among SLE patients (15). We have now extended these findings to other female-biased autoimmune diseases (16, 17), and others have replicated the findings in SLE (18, 19). Thus, this work established that the number of X chromosomes was a risk factor for SLE, and that the number of X chromosomes might underly the female predominance of the disease.

Discussing the potential mechanisms by which an X chromosome dose effect might operate requires a brief review of the biology of the sex chromosomes, which are in mammals, of course, are the X and Y. The X and Y chromosomes pair in meiosis and mitosis by virtue of short regions at the distal ends of both chromosomes known as the pseudoautosomal regions (PAR); namely PAR1 and PAR2. Each PAR contains a handful of genes, which behave identically to autosomal genes. That is, there is expression of one copy on X and one copy on Y with genetic crossover occurring within PAR1 and PAR2 of the X and Y chromosomes. Meanwhile, on the X chromosome, centromeric to the two PARs are about 2000 genes that are X-linked. Similarly, on the Y chromosome centromeric to the two PARs are about 40 genes in the non-recombining region of Y. Almost all these Y genes are expressed in male gonadal tissue and function in spermatogenesis. In contrast, X-linked genes, like other chromosomes, are not functionally organized; and, generally do not have a Y homologue (although there are exceptions).

In cells with 2 or more X chromosomes, all but one is inactivated by methylation through the action of the X inactive-specific transcript (*Xist)* gene (Figure 1), which encodes a long non-coding RNA (20). That is, since women have two X-chromosomes and men have one, the imbalance in X chromosome gene expression is equalized by each cell with 2 or more X chromosomes randomly undergoing inactivation (which is mediated by methylation of CpG) of all but one X chromosome. However, despite the fact that the inactive X chromosome makes up the cytoplasmic Barr body, X inactivation is not an all-or-none phenomenon. On the inactivated X chromosome (X_i), about 15% of the genes escape methylation partially or completely giving women (and Klinefelter men) more phenotypic variability compared to normal (i.e., 46XY) men (21).

Continued presence of *Xist* transcripts were not thought to be needed for maintenance of X inactivation (22). However, recent data demonstrate that this may not be the case in immune cells. Yu and colleague showed that deletion of Xist in CD11c-positive atypical memory B lymphocytes along with TLR7 activation induced isotype switching. In addition, *Xist* down regulation was found among B cells from women with SLE (23). Also, Anguera has found different localization patterns of the Xist non-coding RNA in B cells with upregulation of 20 X chromosome genes in female cells (24, 25). In a published preprint, conditional knock of Xist in female mice (BALB/c and C57BL/6) produced a spontaneous lupus phenotype (26). Thus, there may be differences in the physiology of this long non-coding RNA in B cells that change X chromosome inactivation in such a way that predisposes to a SLE.

X chromosome and immune genes

The idea that immune genes are enriched on the X chromosome is frequently evoked. However, we find this is not the case. All human genes and Gene Ontology (GO) categories were downloaded from NCBI's FTP server (ftp.ncbi.nlm.nih.gov/Gene/DATA/) on August 6th, 2024. Only protein-coding and RNA-producing (eg, ncRNA) transcripts with at least one GO category annotation were selected for analysis. GO categories associated with all transcripts on each human chromosome were then identified, summed, and hypergeometric tests performed to determine relative chromosomal enrichments or depletions in each GO category. False Discovery Rate (FDR) corrections for the most significant p-value (enriched or depleted) were performed to correct for multiple testing. As a positive control, we find the Y chromosome highly enriched (p-value = 0) in the GO categories "spermatogenesis" and "gonadal mesoderm development". We find that, although there are many immune-related genes on the X chromosome, it is not particularly enriched for immune-related genes more than any other chromosome. This was true for all genes related to immune function with 50 of 1,482 (3.4%) on the X chromosome. Furthermore, no individual category of immune function had

10.3389/fimmu.2025.1479814

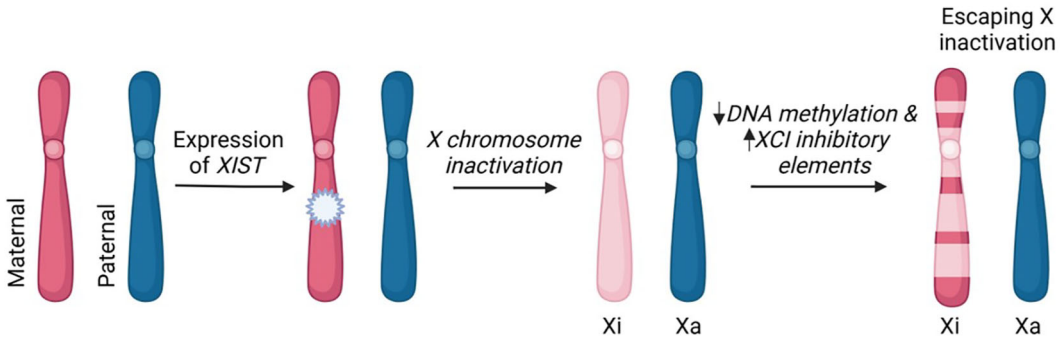


FIGURE 1
X chromosome inactivation and escaping X inactivation. The process of XCI occurs in mammalian cells that have two or more X chromosomes. In early stages of embryonic development, the maternal or paternal X chromosome is randomly silenced. This X-inactivation is initiated by long non-coding RNA, *XIST*, and subsequent DNA methylation and histone modifications. The incomplete inactivation of the X chromosome (pseudoautosomal region and variable genes throughout the X chromosome) results in approximately 15% of X-linked genes remaining transcriptionally active. These "escapee" genes contribute to differential expression of X-linked genes between men and women. Xi-inactive; Xa-active.This image was created in Biorender.com.

enrichment on the X chromosome (Table 1). In fact, we found significant immune-related transcript enrichment on other chromosomes, particularly chromosome 9 (Table 2), and we found other GO categories enriched on the X chromosome (Supplementary Table S1). Some of the categories in Supplementary Table S1 might impact immune processes (eg, miRNA-mediated gene silencing), but none are not immune-specific.

TABLE 1 Immune related gene categories for the X chromosome.

GO group name/ID	on X/total	OR	FDR p value
innate immune response/0045087	14/485	0.81	0.6494
immune response/0006955	8/310	0.72	0.6494
adaptive immune response/0002250	4/193	0.58	0.6494
AHIRMAP/0061844	2/99	0.56	0.6494
activation of innate immune response/0002218	2/32	1.82	0.6494
positive regulation of innate immune response/0045089	2/30	1.95	0.6494
positive regulation of Ig production/0002639	2/28	2.10	0.6494
immunoglobulin mediated immune response/0016064	2/24	2.48	0.6494
immunological synapse formation/0001771	1/13	2.28	0.6494
negative regulation of immune response/0050777	1/12	2.48	0.6494
negative regulation of Ig production/0002638	1/8	3.90	0.6494
positive regulation of adaptive immune response/0002821	1/8	3.90	0.6494
regulation of immunoglobulin production/0002637	1/7	4.56	0.6494
T cell mediated immunity/0002456	1/16	1.82	0.6497
regulation of innate immune response/0045088	1/21	1.37	0.6521
regulation of immune system process/0002682	1/39	0.72	0.6585
innate immune response in mucosa/0002227	1/27	1.05	0.6617
regulation of immune response/0050776	1/27	1.05	0.6617
immune response-regulating signaling pathway/0002764	1/37	0.76	0.6625
humoral immune response/0006959	2/60	0.94	0.6642
positive regulation of immune response/0050778	1/36	0.78	0.6656

AHIRMAP, antimicrobial humoral immune response mediated by antimicrobial peptide.

10.3389/fimmu.2025.1479814

TABLE 2 Gene ontology categories that are significantly found increased on a given chromosome.

chromosome	GO category/ID	#/total	OR	FDR p value	
9	0002286	TCA	17/24	57.99	0
9	0002323	NKCA	17/19	202.96	0
19	0002764	IRRSP	35/37	258.04	0
9	0006959	HIR	18/60	10.23	1.40E-08
6	0050778	PRIR	16/36	13.54	2.15E-08
19	0002682	RISP	15/39	9.21	4.00E-06
20	0045087	IIR	36/485	2.83	6.69E-05
9	0002250	AIR	25/193	3.56	0.0001
6	0002250	AIR	30/193	3.12	0.0001
4	0061844	AHIRMAP	14/99	4.15	0.006
8	0002227	IIRM	7/27	9.37	0.009
17	0045087	IIR	12/485	0.37	0.010
17	0061844	AHIRMAP	17/99	2.99	0.036
12	0061760	AIIR	6/18	8.62	0.045

TCA, T cell activation involved in immune response; NKCA, natural killer cell activation involved in immune response; IRRSP, immune response-regulating signaling pathway; HIR, humoral immune response; PRIR, positive regulation of immune response; RISP, regulation of immune system process; IIR, innate immune response; AIR, adaptive immune response; AHIRMAP, antimicrobial humoral immune response mediated by antimicrobial peptide; IIRM, innate immune response in mucosa; AIIR, antifungal innate immune response.

Candidate X genes in lupus

X chromosome genes that escape X inactivation; and, thus have expression of the gene from each of X chromosome, are candidates to mediate the X chromosome dose effect. Our attention was drawn to two genes in the toll like receptor 7 (TLR7) pathway that routinely escape X inactivation; namely, *TLR7* itself and *TASL* (TLR Adaptor Interacting With Endolysosomal SLC15A4). The TLR7 pathway is critical for the pathogenesis of SLE, both in murine models and humans. For instance, rare gain-of-function TLR7 mutations can cause monogenic pediatric SLE (27–29) and mice with TLR7 over-expression due to a translocation between the X and Y chromosome develop a lupus-like illness (30, 31). The TLR7 protein is localized to the endosome and is critical for recognition of viruses and subsequent activation of the innate immune system. TLR7 binds single-stranded RNA or metabolites thereof, which activates the pathway, leading to production of interferon as well as other cytokines (32). Furthermore, common population variants of genes encoding protein that function in the TLR7 pathway show genetic association to the SLE phenotype. These include TLR7, TASL, SLC15a4 (a binding partner of TASL (33)), and UNC93B1, a regulator of TLR7 movement into the endosome (34–37). Many functional studies also implicate the TLR7 pathway in SLE pathogenesis in both human and murine lupus models (30, 31, 38–43).

Given the critical nature of the TLR7 pathway in SLE and the association of X chromosome number with the sex bias of the disease, we elected to study the role of TASL in the TLR7 pathway. As described above, the TASL gene routinely escapes X inactivation and TASL is expressed in several immune cells, including B lymphocytes and monocytes, contains an SLE risk allele (19, 35) and binds SLC15A4 on the lysosomal surface (44). SLC15a4 regulates lysosomal pH, to which TLR7 signaling is highly sensitive (45, 46). In addition, knockout of the gene is known to abrogate TLR7 signaling (47).

Given these data, we undertook studies to examine the role of TASL in the TLR7 pathway (48). In particular, since TASL and SLC15a4 are binding partners and SLC15a4, at least in part, determines lysosomal pH, we studied lysosomal pH. First, we examined expression of the TASL protein in human primary monocytes, B cells and lymphoblastoid cells lines. In each case, TASL was expressed more highly in female cells compared to male cells (49). Additional studies from Odham et al, also found TASL was more highly expressed in female cells and this sexual dimorphism was magnified when stimulated with type I interferons (50). Using a ratiometric measurement of lysosomal pH via fluorescence in unstimulated female monocytes, we found lysosomal pH averaged 4.9 versus 5.6 in male cells (p=0.0001) (48). A similar difference in lysosomal pH was also found between male and female B cells and dendritic cells, while we did not find a female: male dichotomy for lysosomal pH in NK or T cells, neither of which express TASL (48). Thus, the sex difference in lysosomal pH is likely to be associated with increased TLR7 signaling, and may be dependent upon increased expression of TASL in female cells.

In order to determine if, in fact, TASL participates in lysosomal pH regulation and TLR7 signaling, we undertook a series of knockdown experiments using CRISPR-Cas9 and primary human monocytes (CD14+/CD16−). In female cells treated with a TLR7 agonist, TASL knockdown abrogated interferon-alpha, IL-6 and TNF production (49). Thus, TASL is critical for TLR7 pathway signaling. Furthermore, knockdown of TASL expression resulted in a rise in lysosomal pH in female monocytes to the pH we found in

male monocytes. And, intracellular transport of NOD1 antigens, a function of SLC15a4, was also abrogated by TASL knockdown (49). However, it should be noted that these results have not been independently replicated; and, thus, are not confirmed.

Several other lines of evidence support a sex-biased function of the TRL7 pathway (51–53). Our studies in primary monocytes and LCLs suggest TASL is involved in the TLR7 in a sexually dimorphic manner such that lysosomal pH is lower and TLR7 signaling greater in female versus male cells. As of late, studies on TASL have shown that the once uncharacterized protein functions as enzyme that regulates interferon regulatory factor 5 (IRF5), colocalizes with TLR7 and is interferon inducible. TASL ability to increase interferon production (our work and others) and its own protein level to be subsequently amplified by interferon stimulations suggest a positive feedforward response that would result in increased production that is often found in SLE affected subjects. Thus, increased expression of both TLR7 (54) and TASL (48, 49) may underlie not only improved outcome of women compared to men in some infections (55) but also female disposition to autoimmunity mediated via TLR7 (56).

XIST in lupus

Other investigators have taken a different tack in studying the role of the X chromosome in the sex bias of SLE (57, 58); however, the data generated also concern the TLR7 pathway. As mentioned above, XIST long non-coding RNA mediates X chromosome inactivation (Figure 1); and, thus, is expressed only in cells with more than one X chromosome. Dou and colleagues preformed a series of experiments that indicate XIST is a source of ligand for TLR7; and, of course, this is a sex specific source of ligand (57, 58).

First, these investigators noted that XIST is rich in potential TLR7 ligands. A putative TLR7 stimulatory motif, the UU dinucleotide, was found 2,140 times in XIST RNA. XIST was the sex-biased transcript with the highest degree of UU dinucleotide gene expression; and,

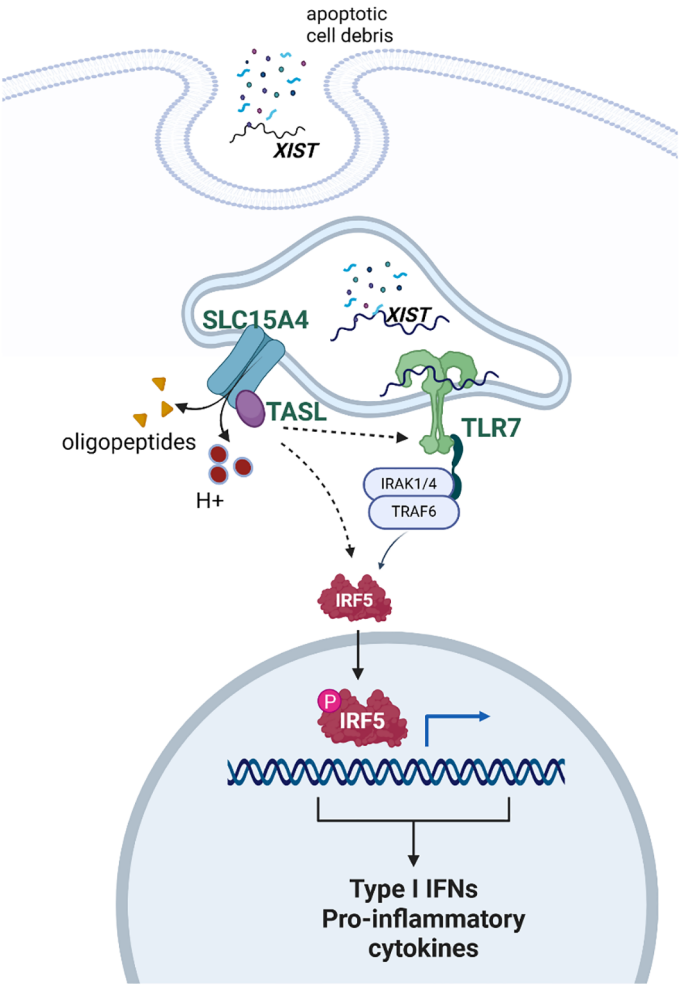


FIGURE 2
Schematic depicting the proposed interaction of XIST, TLR7, and TASL in response to self-antigen. XIST provides ligand for TLR7. Once TLR7 signaling is activated, there is a feed forward stimulation of the pathway. The genes for both TLR7 and TASL are on the X chromosome and escape X inactivation. Thus, some data suggest that TRL7 signaling is more robust in female cells, compared to male cells, on this basis. Created with Biorender.com.

further, was the only sex-biased expression source of the extended TLR7 motif 5′-GUCCUUCAA-3′ (57, 58). Overall, XIST was the strongest sex biased source of self TLR7 ligand.

Next, these investigators turned to stimulation of TLR7 by XIST nucleotides using HEK-hTLR7 cells as a reporter. The extended TLR7 motif found in XIST as well as a longer sequence of XIST (containing the A-repeat, UU dinucleotide rich region) were also found to stimulate TLR7 signaling as indicated by production of interferon-alpha. Further, not only was the response due to specific binding of XIST nucleotide and dose-dependent, the TLR7 response was inhibited by depletion of XIST as well by hydroxychloroquine (57). Additional studies found that XIST levels were higher in peripheral leukocytes among women with SLE compared to non-SLE affected matched controls, and that levels of XIST correlated with disease activity. The investigators concluded, and we certainly agree, that the XIST long non-coding RNA is the most potent source of sex biased TLR7 ligands in female cells.

XIST, TLR7, TASL in lupus and other autoimmune diseases – an hypothesis

We further conclude that these two sets of data suggest synergism for a female biased expansion of the TLR7 signaling pathway that could underlie the X chromosome dose effect found in various autoimmune diseases, including SLE (12–15), Sjögren's disease (13, 16), polymyositis/dermatomyositis (17), and systemic sclerosis (17). The idea, we think, is straight forward. XIST RNA supplies TLR7 ligand in female cells. In addition, female B lymphocytes, dendritic cells, and monocytes have enhanced TLR7 pathway signaling by virtue of the over-expression (compared to male cells) of not only TLR7 but also TASL. Enhanced TLR7 signaling activity deploys a feed forward loop in the TLR7 pathway that leads to increased expression and activity of the pathway (59). Thus, both increased ligand and enhanced activity support further enhancement of TLR7 signaling in female cells. Of course, these phenomena are universal in cells with more than one X chromosome; that is, from women or Klinefelter men. So, other factors must be in play such as other genetics or environmental exposure.

TLR7 signaling and environmental triggers in lupus

What environmental exposure might interact with this sex-biased enhancement of TLR7 signaling induced by Xist and genes in the TLR7 pathway that escape X inactivation? One candidate is Epstein Barr virus (EBV). Epidemiological evidence supports the idea that this near ubiquitous infection is necessary but not sufficient for the expression of SLE as well as multiple sclerosis, and there some evidence in Sjögren's disease (60–65). Recent studies have found that single nucleotide polymorphisms demonstrating genetic association with SLE or Sjögren's disease are more likely to be found in promoter regions bound by the EBV transcription factor EBV nuclear antigen 2 (EBNA2) (66, 67). Overall, the preponderance of evidence indicates that EBV infection is likely one of the environmental triggers for disease. Furthermore, EBV infects B lymphocytes, a cell type with expression of TASL, engaging and increasing expression of TLR7 (68). B cell hyperplasia is one of the hallmarks of systemic autoimmune disease (69). Thus, these data concerning enhanced expressed XIST, TLR7 and TASL in female cells impacting TLR7 signaling may interact with data concerning a role of EBV in promoting SLE and other autoimmune diseases (62, 63, 70). Of course, estrogen and differential expression of estrogen-regulated genes remain a potential biological trigger of the disease. The sex bias of SLE is present in prepubescent children at about 5 to 1, but of course is less pronounced than after puberty (71). These data suggest an effect of estrogen. Further, there are clear effects of estrogen on B lymphocytes and humeral immunity (72, 73) with effects on development, immune tolerance, immunoglobulin somatic hypermutation, and class switching. In addition, some estrogen effects in B cells may be mediated through cell surface (as opposed to nuclear) estrogen receptors (74).

Summary

The evidence is strong that the number of X chromosomes is important for the female bias of some, but not all, autoimmune diseases. The mechanism by which a dose effect for the X chromosome is not understood. Available evidence suggests that multiple factors may play roles that are complementary. These include expression of XIST, which provides TRL7 ligand, and escape of X inactivation by genes whose protein products are critical for TLR7 signaling (see Figure 2).

Author contributions

RHS: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. JW: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing. VL: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported in part by research award BX001451 to RS from the US Department of Veterans Affairs as well as NIH grants AR053483, AI082717, and GM104938. The funders had no role in the concept, the data gathering, the data analysis, the writing, or the decision to publish this manuscript.

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2025.1479814/full#supplementary-material

References

1. Kaposi M. Lupus eythematosus. In: Hebra F, editor. *On Diseases of the Skin including the Exanthemata*. The New syndenham society, London (1875). p. 14–35.

2. Scofield RH, Oates J. The place of William Osler in the description of systemic lupus erythematosus. *Am J Med Sci*. (2009) 338:409–12. doi: 10.1097/MAJ.0b013e3181acbd71

3. Harvey AM, Shulman LE, Tumulty PA, Conley CL, Schoenrich EH. Systemic lupus erythematosus: review of the literature and clinical analysis of 138 cases. *Med (Baltimore)*. (1954) 33:291–437. doi: 10.1097/00005792-195412000-00001

4. Dall'Era M, Cisternas MG, Snipes K, Herrinton LJ, Gordon C, Helmick CG. The incidence and prevalence of systemic lupus erythematosus in San Francisco County, California: the California lupus surveillance project. *Arthritis Rheumatol*. (2017) 69:1996–2005. doi: 10.1002/art.v69.10

5. Lu LJ, Wallace DJ, Ishimori ML, Scofield RH, Weisman MH. Review: Male systemic lupus erythematosus: a review of sex disparities in this disease. *Lupus*. (2010) 19:119–29. doi: 10.1177/0961203309350755

6. Lahita RG, Bradlow HL. Klinefelter's syndrome: hormone metabolism in hypogonadal males with systemic lupus erythematosus. *J Rheumatol*. (1987) 14 Suppl 13:154–7.

7. Mok CC, Lau CS. Profile of sex hormones in male patients with systemic lupus erythematosus. *Lupus*. (2000) 9:252–7. doi: 10.1191/096120300680198926

8. Jara LJ, Benitez G, Medina G. Prolactin, dendritic cells, and systemic lupus erythematosus. *Autoimmun Rev*. (2008) 7:251–5. doi: 10.1016/j.autrev.2007.11.018

9. McMurray RW. Prolactin in murine systemic lupus erythematosus. *Lupus*. (2001) 10:742–7. doi: 10/1191/096120301717164985

10. Mackworth-Young CG, Parke AL, Morley KD, Fotherby K, Hughes GR. Sex hormones in male patients with systemic lupus erythematosus: a comparison with other disease groups. *Eur J Rheumatol Inflammation*. (1983) 6:228–32.

11. Chang DM, Chang CC, Kuo SY, Chu SJ, Chang ML. Hormonal profiles and immunological studies of male lupus in Taiwan. *Clin Rheumatol*. (1999) 18:158–62. doi: 10.1007/s100670050075

12. Scofield RH, Bruner GR, Namjou B, Kimberly RP, Ramsey-Goldman R, Petri M, et al. Klinefelter's syndrome (47,XXY) in male systemic lupus erythematosus patients: support for the notion of a gene-dose effect from the X chromosome. *Arthritis Rheumatism*. (2008) 58:2511–7. doi: 10.1002/art.23701

13. Liu K, Kurien BT, Zimmerman SL, Kaufman KM, Taft DH, Kottyan LC, et al. X chromosome dose and sex bias in autoimmune diseases: increased prevalence of 47, XXX in systemic lupus erythematosus and sjogren's syndrome. *Arthritis Rheumatol*. (2016) 68:1290–300. doi: 10.1002/art.39560

14. Sharma R, Harris VM, Cavett J, Kurien BT, Liu K, Koelsch KA, et al. Rare X chromosome abnormalities in systemic lupus erythematosus and sjogren's syndrome. *Arthritis Rheumatol*. (2017) 69:2187–92. doi: 10.1002/art.v69.11

15. Cooney CM, Bruner GR, Aberle T, Namjou-Khales B, Myers LK, Feo L, et al. 46,X,del (X)(q13) Turner's syndrome women with systemic lupus erythematosus in a pedigree multiplex for SLE. *Genes Immunity*. (2009) 10:478–81. doi: 10.1038/gene.2009.37

16. Harris VM, Sharma R, Cavett J, Kurien BT, Liu K, Koelsch KA, et al. Klinefelter's syndrome (47,XXY) is in excess among men with Sjogren's syndrome. *Clin Immunol*. (2016) 187:137–8. doi: 10.1016/J.clim.2016.01.002

17. Scofield RH, Lewis VM, Cavitt J, Kurien BT, Assassi S, Martin J, et al. 47XXY and 47XXX in scleroderma and myositis. *ACR Open Rheumatol*. (2022) 4:528–33. doi: 10.1002/acr2.11413

18. Tangtanatakul P, Lei Y, Jaiwan K, Yang W, Boonbangyang M, Kunhapan P, et al. Association of genetic variation on X chromosome with systemic lupus erythematosus in both Thai and Chinese populations. *Lupus Sci Med*. (2024) 11(1):e001061. doi: 10.1136/lupus-2023-001061

19. Bentham J, Morris DL, Cunninghame Graham DS, Pinder CL, Tombleson P, Behrens TW, et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat. Genet*. (2015) 47(12):1457–65. doi: 10.1038/ng.3434

20. Kanata E, Duffié R, Schulz EG. Establishment and maintenance of random monoallelic expression. *Development*. (2024) 151(10):dev201741. doi: 10.1242/dev.201741

21. Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature*. (2005) 434:400–4. doi: 10.1038/nature03479

22. Csankovszki G, Panning B, Bates B, Pehrson JR, Jaenisch R. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. *Nat Genet*. (1999) 22:323–4. doi: 10.1038/11887

23. Yu B, Qi Y, Li R, Shi Q, Satpathy AT, Chang HY. B cell-specific XIST complex enforces X-inactivation and restrains atypical B cells. *Cell*. (2021) 184:1790–803.e17. doi: 10.1016/j.cell.2021.02.015

24. Wang J, Syrett CM, Kramer MC, Basu A, Atchison ML, Anguera MC. Unusual maintenance of X chromosome inactivation predisposes female lymphocytes for increased expression from the inactive X. *Proc Natl Acad Sci USA*. (2016) 113(14): E2029-38. doi: 10.1073/pnas.1520113113

25. Syrett CM, Sierra I, Beethem ZT, Dubin AH, Anguera MC. Loss of epigenetic modifications on the inactive X chromosome and sex-biased gene expression profiles in B cells from NZB/W F1 mice with lupus-like disease. *J Autoimmun*. (2020) 107:102357. doi: 10.1016/j.jaut.2019.102357

26. Lovell CD, Jiwrajka N, Amerman HK, Cancro MP, Anguera MC. Xist deletion in B cells results in systemic lupus erythematosus phenotypes. *bioRxiv*. (2024). doi: 10.1101/2024.05.15.594175

27. Stremenova Spegarova J, Sinnappurajar P, Al Julandani D, Navickas R, Griffin H, Ahuja M, et al. A *de novo* TLR7 gain-of-function mutation causing severe monogenic lupus in an infant. *J Clin Invest*. (2024) 134(13):e179193. doi: 10.1172/jc1179193

28. Brown GJ, Cañete PF, Wang H, Medhavy A, Bones J, Roco JA, et al. TLR7 gain-of-function genetic variation causes human lupus. *Nature*. (2022) 605:349–56. doi: 10.1038/s41586-022-04642-z

29. David C, Arango-Franco CA, Badonyi M, Fouchet J, Rice GI, Didry-Barca B, et al. Gain-of-function human UNC93B1 variants cause systemic lupus erythematosus and chilblain lupus. *J Exp Med*. (2024) 221(8):e20232066. doi: 10.1084/jem.20232066

30. Pisitkun P, Deane JA, Difilippantonio MJ, Tarasenko T, Satterthwaite AB, Bolland S. Autoreactive B cell responses to RNA-related antigens due to TLR7 gene duplication. *Science*. (2006) 312:1669–72. doi: 10.1126/science.1124978

31. Subramanian S, Tus K, Li QZ, Wang A, Tian XH, Zhou J, et al. A Tlr7 translocation accelerates systemic autoimmunity in murine lupus. *Proc Natl Acad Sci United States America*. (2006) 103:9970–5. doi: 10.1073/pnas.0603912103

32. Li F, Song B, Zhou WF, Chu LJ. Toll-like receptors 7/8: A paradigm for the manipulation of immunologic reactions for immunotherapy. *Viral Immunol*. (2023) 36:564–78. doi: 10.1089/vim.2023.0077

33. Zhang H, Bernaleau L, Delacrétaz M, Hasanovic E, Drobek A, Eibel H, et al. SLC15A4 controls endolysosomal TLR7-9 responses by recruiting the innate immune adaptor TASL. *Cell Rep*. (2023) 42:112916. doi: 10.1016/j.celrep.2023.112916

34. Al-Azab M, Idiiatullina E, Liu Z, Lin M, Hrovat-Schaale K, Xian H, et al. Genetic variants in UNC93B1 predispose to childhood-onset systemic lupus erythematosus. *Nat Immunol*. (2024) 25:969–80. doi: 10.1038/s41590-024-01846-5

35. Morris DL, Sheng Y, Zhang Y, Wang YF, Zhu Z, Tombleson P, et al. Genome-wide association meta-analysis in Chinese and European individuals identifies ten new loci associated with systemic lupus erythematosus. *Nat Genet*. (2016) 48:940–6. doi: 10.1038/ng.3603

36. Baccala R, Gonzalez-Quintial R, Blasius AL, Rimann I, Ozato K, Kono DH, et al. Essential requirement for IRF8 and SLC15A4 implicates plasmacytoid dendritic cells in the pathogenesis of lupus. *Proc Natl Acad Sci U S A*. (2013) 110:2940–5. doi: 10.1073/pnas.1222798110

37. Blasius AL, Krebs P, Sullivan BM, Oldstone MB, Popkin DL. Slc15a4, a gene required for pDC sensing of TLR ligands, is required to control persistent viral infection. *PloS Pathog.* (2012) 8:e1002915. doi: 10.1371/journal.ppat.1002915

38. von Hofsten S, Fenton KA, Pedersen HL. Human and murine toll-like receptor-driven disease in systemic lupus erythematosus. *Int J Mol Sci.* (2024) 24(10):5351. doi: 10.3390/ijms25105351

39. Sakata K, Nakayamada S, Miyazaki Y, Kubo S, Ishii A, Nakano K, et al. Up-regulation of TLR7-mediated IFN-alpha production by plasmacytoid dendritic cells in patients with systemic lupus erythematosus. *Front Immunol.* (2018) 9:1957. doi: 10.3389/fimmu.2018.01957

40. Clancy RM, Markham AJ, Buyon JP. Endosomal Toll-like receptors in clinically overt and silent autoimmunity. *Immunol Rev.* (2016) 269:76–84. doi: 10.1111/imr.2016.269.issue-1

41. Miquel CH, Faz-Lopez B, Guéry JC. Influence of X chromosome in sex-biased autoimmune diseases. *J Autoimmun.* (2023) 137:102992. doi: 10.1016/j.jaut.2023.102992

42. Wolf C, Lim EL, Mokhtari M, Kind B, Odainic A, Lara-Villacanas E, et al. UNC93B1 variants underlie TLR7-dependent autoimmunity. *Sci Immunol.* (2024) 9: eadi9769. doi: 10.1126/sciimmunol.adi9769

43. Mishra H, Schlack-Leigers C, Lim EL, Thieck O, Magg T, Raedler J, et al. Disrupted degradative sorting of TLR7 is associated with human lupus. *Sci Immunol.* (2024) 9:eadi9575. doi: 10.1126/sciimmunol.adi9575

44. Huttlin EL, Ting L, Bruckner RJ, Gebreab F, Gygi MP, Szpyt J, et al. The bioPlex network: A systematic exploration of the human interactome. *Cell.* (2015) 162:425–40. doi: 10.1016/j.cell.2015.06.043

45. Kobayashi T, Tanaka T, Toyama-Sorimachi N. How do cells optimize luminal environments of endosomes/lysosomes for efficient inflammatory responses? *J Biochem.* (2013) 154:491–9. doi: 10.1093/jb/mvt099

46. Kobayashi T, Tsutsui H, Shimabukuro-Demoto S, Sugitani-Yoshida R, Karyu H, Furuyama-Tanaka K, et al. Lysosome biogenesis regulated by the amino-acid transporter SLC15A4 is critical for functional integrity of mast cells. *Int Immunol.* (2017) 29(12):551–6. doi: 10.1093/intimm/dxx063

47. Kobayashi T, Shimabukuro-Demoto S, Yoshida-Sugitani R, Furuyama-Tanaka K, Karyu H, Sugiura Y, et al. The histidine transporter SLC15A4 coordinates mTOR-dependent inflammatory responses and pathogenic antibody production. *Immunity.* (2014) 41:375–88. doi: 10.1016/j.immuni.2014.08.011

48. Harris VM, Harley ITW, Kurien BT, Koelsch KA, Scofield RH. Lysosomal pH is regulated in a sex dependent manner in immune cells expressing CXorf21. *Front Immunol.* (2019) 10:578. doi: 10.3389/fimmu.2019.00578

49. Harris VM, Koelsch KA, Kurien BT, Harley ITW, Wren JD, Harley JB, et al. Characterization of cxorf21 provides molecular insight into female-bias immune response in SLE pathogenesis. *Front Immunol.* (2019) 10:2160. doi: 10.3389/fimmu.2019.02160

50. Odhams CA, Roberts AL, Vester SK. Interferon inducible X-linked gene CXorf21 may contribute to sexual dimorphism in Systemic Lupus Erythematosus. *Nat. Commun.* (2019) 10:2164. doi: 10.1038/s41467-019-10106-2

51. Guéry JC. Sex differences in primary HIV infection: revisiting the role of TLR7-driven type 1 IFN production by plasmacytoid dendritic cells in women. *Front Immunol.* (2021) 12:729233. doi: 10.3389/fimmu.2021.729233

52. Laffont S, Guéry JC. Deconstructing the sex bias in allergy and autoimmunity: From sex hormones and beyond. *Adv Immunol.* (2019) 142:35–64. doi: 10.1016/bs.ai.2019.04.001

53. Rubtsova K, Marrack P, Rubtsov AV. TLR7, IFNγ, and T-bet: their roles in the development of ABCs in female-biased autoimmunity. *Cell Immunol.* (2015) 294:80–3. doi: 10.1016/j.cellimm.2014.12.002

54. Souyris M, Cenac C, Azar P, Daviaud D, Canivet A, Grunenwald S, et al. TLR7 escapes X chromosome inactivation in immune cells. *Sci Immunol.* (2018) 3(19): eaap8855. doi: 10.1126/sciimmunol.aap8855

55. Spiering AE, de Vries TJ. Why females do better: the X chromosomal TLR7 gene-dose effect in COVID-19. *Front Immunol.* (2021) 12:756262. doi: 10.3389/fimmu.2021.756262

56. Souyris M, Mejia JE, Chaumeil J, Guery JC. Female predisposition to TLR7-driven autoimmunity: gene dosage and the escape from X chromosome inactivation. *Semin. Immunopathol.* (2019) 41:153–64. doi: 10.1007/s00281-018-0712-y

57. Crawford JD, Wang H, Trejo-Zambrano D, Cimbro R, Talbot CC Jr., Thomas MA, et al. The XIST lncRNA is a sex-specific reservoir of TLR7 ligands in SLE. *JCI Insight.* (2023) 8(20):e169344. doi: 10.1172/jci.insight.169344

58. Dou DR, Zhao Y, Belk JA, Zhao Y, Casey KM, Chen DC, et al. Xist ribonucleoproteins promote female sex-biased autoimmunity. *Cell.* (2024) 187:733–49.e16. doi: 10.1016/j.cell.2023.12.037

59. Assil S, Coléon S, Dong C, Décembre E, Sherry L, Allatif O, et al. Plasmacytoid dendritic cells and infected cells form an interferogenic synapse required for antiviral responses. *Cell Host Microbe.* (2019) 25:730–45.e6. doi: 10.1016/j.chom.2019.03.005

60. Almohmeed YH, Avenell A, Aucott L, Vickers MA. Systematic review and meta-analysis of the sero-epidemiological association between Epstein Barr virus and multiple sclerosis. *PloS One.* (2013) 8:e61110. doi: 10.1371/journal.pone.0061110

61. Draborg A, Izarzugaza JM, Houen G. How compelling are the data for Epstein-Barr virus being a trigger for systemic lupus and other autoimmune diseases? *Curr Opin Rheumatol.* (2016) 28:398–404. doi: 10.1097/BOR0000000000000289

62. Jog NR, James JA. Epstein barr virus and autoimmune responses in systemic lupus erythematosus. *Front Immunol.* (2020) 11:623944. doi: 10.3389/fimmu.2020.623944

63. Laurynenka V, Ding L, Kaufman KM, James JA, Harley JB. A high prevalence of anti-EBNA1 heteroantibodies in systemic lupus erythematosus (SLE) supports anti-EBNA1 as an origin for SLE autoantibodies. *Front Immunol.* (2022) 13:830993. doi: 10.3389/fimmu.2022.830993

64. Singh D, Oudit O, Hajtovic S, Sarbaugh D, Salis R, Adebowale T, et al. Antibodies to an Epstein Barr Virus protein that cross-react with dsDNA have pathogenic potential. *Mol Immunol.* (2021) 132:41–52. doi: 10.1016/j.molimm.2021.01.013

65. Wood RA, Guthridge L, Thurmond E, Guthridge CJ, Kheir JM, Bourn RL, et al. Serologic markers of Epstein-Barr virus reactivation are associated with increased disease activity, inflammation, and interferon pathway activation in patients with systemic lupus erythematosus. *J Transl Autoimmun.* (2021) 4:100117. doi: 10.1016/j.jtauto.2021.100117

66. Harley JB, Chen X, Pujato M, Miller D, Maddox A, Forney C, et al. Transcription factors operate across disease loci, with EBNA2 implicated in autoimmunity. *Nat. Genet.* (2018) 50:699–707. doi: 10.1038/s41588-018-0102-3

67. Hong T, Parameswaran S, Donmez OA, Miller D, Forney C, Lape M, et al. Epstein-Barr virus nuclear antigen 2 extensively rewires the human chromatin landscape at autoimmune risk loci. *Genome Res.* (2021) 31:2185–98. doi: 10.1101/gr.264705.120

68. Liu L, Wang Y, Wang W, Ying W, Sun B, Wang X, et al. Increased expression of the TLR7/9 signaling pathways in chronic active EBV infection. *Front Pediatr.* (2022) 10:1091571. doi: 10.3389/fped.2022.1091571

69. Brauner S, Folkersen L, Kvarnstrom M, Meisgen S, Petersen S, Franzen-Malmros M, et al. H1N1 vaccination in Sjogren's syndrome triggers polyclonal B cell activation and promotes autoantibody production. *Ann Rheumatic Dis.* (2017) 76:1755–63. doi: 10.1136/annrheumdis-2016-210509

70. Poole BD, Templeton AK, Guthridge JM, Brown EJ, Harley JB, James JA. Aberrant Epstein-Barr viral infection in systemic lupus erythematosus. *Autoimmun Rev.* (2009) 8:337–42. doi: 10.1016/j.autrev.2008.12.008

71. Smith EMD, Lythgoe H, Midgley A, Beresford MW, Hedrich CM. Juvenile-onset systemic lupus erythematosus: Update on clinical presentation, pathophysiology and treatment options. *Clin Immunol.* (2019) 209:108274. doi: 10.1016/j.clim.2019.108274

72. Hughes GC, Choubey D. Modulation of autoimmune rheumatic diseases by oestrogen and progesterone. *Nat Rev Rheumatol.* (2014) 10:740–51. doi: 10.1038/nrrheum.2014.144

73. Sakiani S, Olsen NJ, Kovacs WJ. Gonadal steroids and humoral immunity. *Nat Rev Endocrinol.* (2013) 9:56–62. doi: 10.1038/nrendo.2012.206

74. Seto K, Hoang M, Santos T, Bandyopadhyay M, Kindy MS, Dasgupta S. Non-genomic oestrogen receptor signal in B lymphocytes: An approach towards therapeutic interventions for infection, autoimmunity and cancer. *Int J Biochem Cell Biol.* (2016) 76:115–8. doi: 10.1016/j.biocel.2016.04.018

frontiers | Frontiers in Immunology

TYPE Review
PUBLISHED 09 May 2025
DOI 10.3389/fimmu.2025.1571700

OPEN ACCESS

EDITED BY
Frederick Miller,
National Institute of Environmental Health Sciences (NIH), United States

REVIEWED BY
Adam Schiffenbauer,
National Institutes of Health (NIH), United States
Minkyo Song,
National Institute on Aging (NIH), United States

*CORRESPONDENCE
Maria De Santis
maria.de_santis@hunimed.eu

RECEIVED 05 February 2025
ACCEPTED 15 April 2025
PUBLISHED 09 May 2025

CITATION
Tonutti A, Ceribelli A, Gremese E, Colafrancesco S, De Santis M and Selmi C (2025) Cancer in connective tissue disease. *Front. Immunol.* 16:1571700. doi: 10.3389/fimmu.2025.1571700

Cancer in connective tissue disease

Antonio Tonutti[1,2], Angela Ceribelli[1,2], Elisa Gremese[1,2], Serena Colafrancesco[1,2], Maria De Santis[1,2]* and Carlo Selmi[1,2]

[1]Department of Biomedical Sciences, Humanitas University, Milan, Italy, [2]Rheumatology and Clinical Immunology, IRCCS Humanitas Research Hospital, Milan, Italy

The association between cancer and autoimmunity is well-recognized, as represented by the increased incidence of cancer among patients with systemic autoimmune diseases; however, the underlying mechanisms remain only partially understood. On the one hand, malignancy may trigger a breakdown of immune tolerance in predisposed individuals, as autoimmune syndromes often emerge shortly after cancer diagnosis, suggesting that tumor antigens might initiate an autoimmune response. However, by involving persistent responses and the creation of a pro-inflammatory environment, the chronic immune activation characteristic of autoimmunity may promote oncogenesis. This scenario is further complicated by the use of immunosuppressive therapies for autoimmune conditions, which, as seen in transplant immunology, are associated with a higher risk of cancer, although data in rheumatology have not yielded definitive conclusions. Connective tissue diseases include systemic lupus erythematosus, primary Sjögren syndrome, idiopathic inflammatory myopathies, systemic sclerosis, mixed connective tissue disease, and undifferentiated forms. These conditions have been variably associated with an increased risk of cancer, both at the time of disease onset and in patients with long-standing autoimmune conditions, providing a paradigm for investigating this complex interplay. Despite recent progress, many unmet needs remain that warrant further research.

KEYWORDS

malignancy, immunology, autoimmunity, autoantibodies, connective tissue disease (CTD)

Why cancer and connective tissue disease

The relationship between malignancy and autoimmunity is well established, as supported by the increased incidence of cancer in patients with autoimmune diseases (1); however, several questions remain unanswered regarding the fundamental mechanisms of this association and their translation into clinical practice. In line with the established pathogenic model of autoimmune diseases, malignancy may trigger the breakdown of tolerance in predisposed individuals (2). This is illustrated by the occurrence of autoimmune syndromes, often with distinctive features, in close temporal proximity to cancer diagnosis (3). On the other hand, autoimmunity may serve as a fertile ground for the

development of malignancy, possibly due to persistent immune activation against autoantigens and the setting of a pro-inflammatory milieu, thus acting as a precancerous condition (4). Furthermore, autoimmune diseases are often treated using immunosuppressive therapies. While evidence from transplant immunology indicates that immunosuppression increases the risk of cancer (5), data are inconclusive when it comes to rheumatology and clinical immunology (6).

Connective tissue diseases (CTDs) are classic forms of systemic autoimmune disorders, including systemic lupus erythematosus (SLE), primary Sjögren syndrome (pSS), idiopathic inflammatory myopathies (IIM), systemic sclerosis (SSc), mixed connective tissue disease (MCTD), and undifferentiated forms (UCTD) (7–12). These diseases are characterized by unique clinical features and pathogenic mechanisms but also share a female predominance, overlapping clinical manifestations (e.g., arthralgia and arthritis, fatigue, interstitial lung disease, myositis, and Raynaud's phenomenon) (7–12), and similar immunological pathways (e.g., type I interferon activation, B-cell infiltration, activation, and proliferation) (13, 14). Within this shared framework, an increased risk of malignancy has frequently been reported across CTDs, reflecting the intricate interplay between cancer and autoimmunity (Figure 1). We speculate that some entities reflect the causal relationship of autoimmunity as a paraneoplastic phenomenon, as seen in cancer-associated myositis (CAM) or -scleroderma, where the temporal closeness between the two diagnoses is linked to peculiar environmental and pathophysiological changes (15). In other scenarios, subclinical chronic inflammation may constitute a precancerous condition contributing to the development of cancer-associated mutations and malignancy late in disease history (16, 17).

By evaluating the spectrum of CTDs, we present a critical analysis of the relationship between cancer and autoimmunity, with a focus on clinical associations, relevance of serum autoantibodies, impact of disease-specific risk factors, and role of immunosuppressive therapies. Different scenarios will be presented to support the proposed concept that certain CTDs can represent a paraneoplastic phenomenon, whereas the onset of malignancy is observed more frequently in specific longstanding CTD-related contexts. To ensure a consistent approach, similar sections will be summarized for different diseases. However, there are major differences in the available evidence, and considering that our work aims to provide a critical review of the state of the art while identifying clinical and research needs, the content of certain sections will need to be heterogeneous and vary from one condition to another. This is particularly evident in the section on immunological features, which lacks a uniform distribution in myositis and SSc compared to pSS and SLE. Table 1 summarizes the unmet needs in the management of malignancy in patients with CTDs and outlines a contextual research agenda based on the discussions presented throughout the text.

Methods and search strategy

We conducted a comprehensive critical review by searching PubMed for "idiopathic inflammatory myopathies," "systemic sclerosis," "Sjogren Disease," "systemic lupus erythematosus," and "cancer." The search focused on articles published in English from January 2010 to October 2024 and yielded 3,652 results. Papers of key relevance published outside of this period were included if they focused on relevant findings and approaches that could have influenced subsequent publications. Thus, 196 papers were included in the final review. A balanced discussion was provided by including studies that supported or challenged our perspective, ensuring a comprehensive and evidence-based analysis. Multiple reviewers (AT, AC, EG, and SC) independently evaluated the included studies; their interpretation was discussed by the full author panel to minimize bias and reach consensus, and different viewpoints were considered during the synthesis of the results. Owing to the heterogeneity of study designs, patient populations, and outcome measures, which made direct comparisons

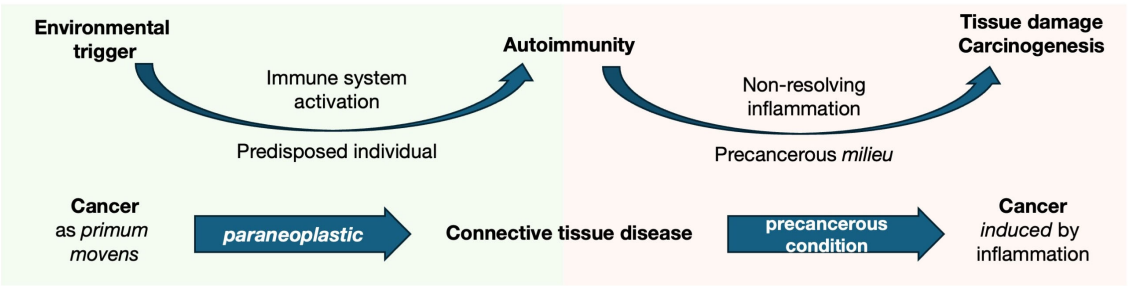


FIGURE 1
Cancer as both an environmental trigger and pathological consequence of autoimmunity in the paradigm of CTDs. The pathogenesis of autoimmune diseases involves a hypothetical environmental trigger that induces immune system response. In genetically predisposed individuals, this leads to an aberrant immune activation, which becomes dysregulated and persists over time, resulting in chronic inflammation. The chronic inflammatory milieu causes tissue damage due to ongoing inflammation but synchronously provides a precancerous condition (i.e., an environment that predisposes to the development of cancerous lesions). From this perspective, CTD are at a crossroads between cancer and autoimmunity. On the one hand, strong evidence supports the role of cancer as a trigger of autoimmune responses (as seen in cancer-associated myositis and scleroderma). However, the disease itself increases the risk of malignancies, particularly in tissues undergoing chronic inflammatory remodeling (such as the lung in SSc and lymphopoiesis in pSS).

TABLE 1 The unmet needs and research agenda in cancer management in patients with CTDs.

	IIM	SSc	pSS	SLE
Risk assessment	Can we measure the risk of cancer in new-onset IIM? Can we further stratify the risk in patients with specific phenotypes? (e.g., different autoantibodies associated with DM, ASyS) Can we better stratify patients at high risk according to disease phenotypes and autoantibodies? (e.g., anti-TIF1γ DM who do not develop cancer) Can we assess the risk of IIM in patients newly diagnosed with cancer?	Can we measure the risk of cancer in new-onset SSc? Can we measure the risk of cancer in longstanding SSc? What is the contribution of SSc to this risk? Which sites are the most involved? **In cancer-associated scleroderma**: what are the risk factors beyond anti-POLR3+ dcSSc? **Cancer-associated scleroderma**: which anti-POLR3+ patients will not develop cancer? **Late-onset cancer in SSc**: which patients should be thoroughly and repeatedly screened?	**Lymphoma**: can we measure or score the risk of developing lymphoma in pSS? Can we predict the time-to-lymphoma interval in pSS? Will any novel autoantibody provide more insights in estimating the risk of lymphoma in SSc? **Non-lymphoma**: Is pSS a risk factor for solid neoplasms? Which are the most common neoplasms? What are they associated with?	Which disease categories and phenotypes are at high vs. low risk of cancer? Can we identify any patient cluster? What is the timing of cancer onset in patients with SLE? Which autoantibodies are associated with cancer in patients with SLE, if any? Are overlap diseases (pSS, thyroiditis, autoimmune hepatitis or cholangitis) a concern in patients with SLE? Is elderly-onset SLE a risk factor for cancer?
Screening	How long and how often should patients with IIM be screened for malignancy? Should patients receive long-term screening for specific cancers in case of select internal organ involvement (e.g., ILD)? Does this apply to all IIM patients independently from the risk estimated according to the IMACS guidelines?	**Cancer-associated scleroderma**: How long and often should we screen patients? How should we screen patients? **Longstanding SSc:** When should we start screening patients? How often should we screen them? For which cancers should we screen them? Which diagnostic tests should be used and how should they be used? Which age, if any, should we start or stop searching for cancer in SSc?	**Lymphoma**: How should we screen patients? Which sites should be screened? When should we start screening patients? How often should patients be screened? **Non-lymphoma**: Should patients with pSS be offered a dedicated cancer screening because of specific risk factors?	Should patients with SLE be offered a dedicated cancer screening because of specific risk factors? In patients with overlapping CTDs, should patients follow the same screening procedures according to the overlapping entity?
Treatment	Does immunosuppressive treatment increase the risk of cancer in patients with CTDs? What is the correct management of immunosuppressive therapies in CTD patients with newly diagnosed malignancy? To what extent does disease activity enhance the risk of cancer in patients with CTDs? Which immunosuppressive treatments contribute to reduce vs. increase such risk by controlling disease activity?			

anti-POLR3, anti-RNA polymerase III autoantibodies; ASyS, antisynthetase syndrome; CTDs, connective tissue diseases; DM, dermatomyositis; IIM, idiopathic inflammatory myopathies; ILD, interstitial lung disease; IMACS, International Myositis Assessment and Clinical Studies Group; pSS, primary Sjogren Syndrome; SLE, systemic lupus erythematosus; SSc, systemic sclerosis.

10.3389/fimmu.2025.1571700

challenging, a narrative approach was adopted instead of a systematic review. To ensure a broad and speculative perspective on the topic, rigid predefined inclusion and exclusion criteria were not applied. However, studies included were original peer-reviewed research articles, systematic reviews, and meta-analyses. Case reports and small case series were considered only when they provided unique insights into novel clinical associations. Non-peer-reviewed sources and studies were excluded to maintain the robustness of the analysis.

Cancer and idiopathic inflammatory myopathies: the key role of synchronous malignancy

The heterogenous family of IIM encompasses dermatomyositis (DM), polymyositis (PM), antisynthetase syndrome (ASyS), immune-mediated necrotizing myopathy (IMNM), inclusion body myositis (IBM), juvenile inflammatory myositis, and paraneoplastic myositis or CAM (10, 18, 19). CAM is defined as a malignancy occurs within three years from the onset of myositis in adult patients (20, 21), and the risk of developing CAM varies according to the disease phenotype and the presence of selected myositis-specific autoantibodies (MSA) (22–24). Since the earliest reports dating back to 1916 (25), several studies have confirmed a strong link between cancer and IIM, particularly with DM and in the presence of autoantibodies targeting transcription intermediary factor 1γ (TIF1-γ) and the nuclear matrix protein 2 (NXP2) (26, 27).

Clinical features of paraneoplastic myositis

DM is the most common IIM clinical phenotype associated with the risk of CAM, presenting as heliotrope rash, Gottron's sign, or papules (28, 29). Patients with inclusion body myositis and ASyS do not seem to have an increased risk of malignancy (26, 30), even when presenting with signs of DM (30), whereas the risk remains unclear in subjects diagnosed with IMNM (31). In addition to the diagnosis of DM, risk factors for CAM include older age at IIM onset, male sex, smoking history, signs of cutaneous necrosis (32), dysphagia (33), rapidly progressive disease, and elevated inflammatory markers (34–37). Histological features on muscle biopsy, such as minimal lymphocytic infiltration, should also raise suspicion for CAM (38) while interstitial lung disease, arthritis, and Raynaud's phenomenon correlate with a lower risk of malignancy (34, 36, 37). Different types of malignancies have been reported with CAM, most commonly solid neoplasms, which seem to reflect the incidence observed in the general population. For instance, a large cohort from Northern Europe reported a high risk of ovarian, gastric, colorectal, and pancreatic cancers, and non-Hodgkin's lymphoma (NHL) (39). In contrast, nasopharyngeal carcinoma was confirmed as the most common neoplasm diagnosed in patients with IIM in the Taiwanese population, followed by lung, breast, and hepatic malignancies (40, 41). Moreover, slight differences in the type of incident neoplasms have been hypothesized by comparing patients with CAM according to the clinical phenotype, i.e., DM vs. PM (39). These differences warrant further investigation across different clinical subsets and ethnicities (Table 1).

Immunological features of paraneoplastic myositis

The immune pathogenesis of CAM involves several complex mechanisms, including the presence of shared antigens between tumor cells and normal tissues, molecular mimicry, and exposure to neo-self-antigens (42). These can be presented to tumor-infiltrating lymphocytes through class I (CD8+ cells) and class II (CD4+ cells) HLA complexes. This process leading to lymphocyte activation may result able to provide cancer elimination; on the other hand, activated lymphocyte may cross react with self-antigens and pathologically infiltrate normal tissues (e.g., skeletal muscle, skin), leading to inflammation and damage (42–44).

Serum autoantibodies, including both myositis-specific (MSA) and myositis-associated (MAA) autoantibodies, are of major use in the diagnosis of IIM and correlate with the development of particular manifestations among different clinical subsets (22). Most importantly, the presence of autoantibodies can further stratify patients with IIM according to cancer risk, as summarized in Table 2.

While malignancies often occur in association with DM, a 2012 meta-analysis including 312 adult patients with DM found that 80% of DM patients with cancer were anti-TIF1-γ-positive, whereas only 10% without cancer had this autoantibody (45). Overall, among patients with DM, the presence of anti-TIF1-γ autoantibodies had a positive predictive value for CAM of 58% and a negative predictive value of 93% (45). These findings were confirmed in another large cohort study, particularly raising concern for breast and ovarian neoplasms (26), and in an up-to-date meta-analysis (34). Moreover, it seems that the risk of cancer significantly increases in patients displaying high anti-TIF1-γ autoantibody titers, specifically in patients with the IgG2 isotype, compared with their respective counterparts (46, 47). TIF1-γ, also known as TRIM33, is an enzyme involved in post-translational peptide modifications, an E3-ubiquitin ligase and being involved in small ubiquitin-like modifications (SUMO). In particular, TIF1-γ has been demonstrated to participate in cell cycle regulation, DNA repair, and the regulation of TGF-β signaling (44). Alterations in the TIF1-γ gene have been described in cancer cells from patients with CAM, possibly representing the neo-self and thus triggering the anti-cancer immune response, which can culminate in autoimmunity to native TIF1-γ antigens (48). As a proof of concept, high expression of TIF1-γ has been observed in the skin and skeletal muscle, which represent the main targets of anti-TIF1-γ DM compared to other tissues (49, 50). Recently, the role of anti-TIF1-γ as a risk factor for synchronous cancer in DM patients has been redefined. Indeed, the coexisting immune response against autoantigens, such as Sp4 and CCAR1, would reduce the risk of cancer, perhaps accounting for a more robust antitumor immunological response (51–53). Further

10.3389/fimmu.2025.1571700

TABLE 2 Myositis-specific and -associated autoantibodies, associated phenotypes and current risk of cancer in IIM patients.

Autoantibody	Target molecule and function	Clinical phenotype	Clinical associations	Cancer risk
anti-TIF1γ/α	Transcription intermediary factor 1γ/α—transcriptional elongation, DNA repair	DM, JDM	DM, no ILD	High
anti-MJ/NXP2	Nuclear matrix protein-2—transcriptional regulation and activation of the tumor suppressor p53	DM, JDM	DM, calcinosis, subcutaneous edema, severe myopathy, dysphagia	High
anti-SAE	Small ubiquitin-like modifier 1 activating enzyme—post-translational modifications	DM	Severe cutaneous disease, dysphagia, systemic symptoms, mild myopathy, mild ILD (50%)	Intermediate
Anti-PUF60 (FIRs)	poly-U-binding factor protein	DM, pSS	Less ILD; in pSS frequently with Ro60, Ro52, La	Intermediate-High (200)
Anti-HMGCR	HMG-CoA reductase—rate-limiting enzyme for cholesterol synthesis	IMNM (statin-induced myopathy)	Necrotizing myopathy	Intermediate
anti-Jo-1	Histidyl-tRNA synthetase	ASyS	Classic ASyS with frequent muscle involvement	Standard
anti-PL-7	Threonyl-tRNA synthetase	ASyS	Severe ILD	Standard
anti-PL-12	Alanyl-tRNA synthetase	ASyS	May present with ILD only	Standard
anti-EJ	Glycyl-tRNA synthetase	ASyS	ASyS, ILD (with anti-Ro52)	Standard
anti-OJ	Isoleucyl-tRNA synthetase	ASyS	ASyS (severe myositis), ILD	Standard
anti-KS	Asparaginyl-tRNA synthetase	ASyS	CADM, ILD, overlap subset with *sicca*	Standard
anti-ZO	Phenylalanyl-tRNA synthetase	ASyS	Classic ASyS, rare (<1% ASyS)	Unknown
anti-YRS (Ha)	Tyrosyl-tRNA synthetase	ASyS	ASyS, rash, arthritis, rare	Unknown
anti-KJ	Translocation factor	ASyS-like	Rare	Unknown
anti-MDA5/IFIH1	Melanoma differentiation-associated gene 5—innate immune responses against viruses	DM, JDM	CADM, severe ILD, peculiar skin involvement (reverse Gottron, vasculitis, ear lesions), mechanic's hands, MIP-C	Intermediate
anti-TIF1-β	Transcription intermediary factor 1β—regulation of gene expression and chromatin structure	DM	CADM, no ILD	Unknown
anti-Ku	Heterodimer complex of 2 subunits that binds to free DNA termini—DNA repair, transcription regulation	SLE, SSc, MCTD, PM	Raynaud, arthralgia, myopathy, overlap with other connective tissue diseases	Standard
Anti-SRP	Signal recognition particle—co-translational translocation of proteins across the endoplasmic reticulum	IMNM	Necrotizing myositis, myocarditis, low ILD	Standard
anti-PM/Scl	complex of 100 KDa and 75 KDa—processing and degradation of RNAs	PM, DM, SSc, PM/SSc overlap, SLE	ASyS-like (myositis, Raynaud, arthritis, ILD, mechanic's hands)	Standard
anti-Mi-2	helicase of the nucleosome remodeling deacetylase—transcriptional regulation	DM	Classic DM (no ILD)	Standard
Anti-cN-1A	Cytosolic 5'-Nucleotidase 1A protein – nucleotide hydrolysis	IBM	Bulbar muscle weakness, wrist flexor involvement	Unknown
Anti-FHL1	Four-and-a-Half LIM domain 1—intracellular protein–protein interactions mainly with cytoskeletal proteins	DM, PM	Severe myositis, dysphagia, vasculitis	Unknown
Anti-RuvBL1/2	Ruv BL1/2 double hexame—DNA repair, chromatin remodeling, gene transcription	SSc, PM	Higher age at onset, men, diffuse SSc and myositis overlap, GI dysmotility, myocarditis	Unknown

(Continued)

TABLE 2 Continued

Autoantibody	Target molecule and function	Clinical phenotype	Clinical associations	Cancer risk
anti-SMN	Survival of motoneuron complex—transcriptional regulation and small nuclear RNP formation	MCTD, PM	MCTD with clinical features of all components of SLE, SSc and IIM; high prevalence of PAH and ILD	Unknown
anti-Nup	Nucleoporins	Not known	Myositis, ILD, Raynaud	Unknown

Cancer risk is reported as 'high' (i.e., increased compared to same-age general population), 'intermediate', or 'standard' (i.e., not different to same-age general population), according to the recent International Myositis Assessment and Clinical Studies Group (IMACS) guidelines (36). Otherwise, for rarer or novel autoantibodies, an estimate of the risk of cancer is given according to the references in the Table, linked to observational cohort studies, whereas 'Unknown' risk is reported if little (e.g., case reports, small case series) or no evidence showing cancer association is available.

ASyS, antisynthetase syndrome (myositis, ILD, polyarthritis, Raynaud's phenomenon, mechanic's hands and the presence of an antisynthetase antibody); CADM, clinically amyopathic/hypomyopathic DM; DM, dermatomyositis; GI, gastrointestinal; IBM, inclusion body myositis; ILD, interstitial lung disease; IMNM, immune-necrotizing myopathy; JDM, juvenile dermatomyositis; MCTD, mixed connective tissue disease; MIP-C, MDA5-associated autoimmunity and interstitial pneumonitis contemporaneous to the COVID-19 pandemics; PAH, pulmonary arterial hypertension; PM, polymyositis; pSS, Sjogren syndrome; SLE, systemic lupus erythematosus; SSc, systemic sclerosis.

implementation of these observations in clinical practice is required (Table 1).

NXP2, also known as MORC3, is a nuclear protein involved in the activation of the tumor suppressor protein p53 (54), a key regulator of cell cycle and senescence. Downregulation of NXP2 has been described in different malignancies, correlating with an enhanced type I IFN signature and, most importantly, with increased expression of the immune checkpoint antigen PD-L1, which is known to suppress T-cell response by binding to the cognate receptor PD-1 (55). Autoantibodies against MJ/NXP2 have been extensively associated with the risk of cancer in IIM patients (27, 56–58), even though some large studies (59) and meta-analyses (60) failed to demonstrate an association with malignancy compared to other patient subsets. The heterogeneity of the results obtained when detecting myositis autoantibodies using different methods (59, 61) suggests that one possible explanation for this discrepancy may be the varying techniques used to identify anti-NXP2 autoantibodies across different studies (58). For instance, in one of the largest studies conducted on anti-NXP2-positive DM, the presence of these autoantibodies was confirmed by immunoprecipitation in only 62% of the patients who tested positive using commercial line blots (59).

Recent studies have reported the risk of malignancy in patients with other rare serum autoantibodies. A higher incidence of cancers was observed with anti-SAE, a hallmark of erythrodermic DM (62–64), with malignancies diagnosed also many years after the onset of myositis in an American cohort (65). SAE1 is a subunit of the E1 complex constituting a SUMO activator protein that plays crucial roles in the activation of type I IFN synthesis but is also involved in tumorigenesis (66). For instance, overexpression of SAE1 has been observed in different types of cancers, correlating with a higher disease burden, metastatic disease, and worse prognosis (67–69). Concerning IMNM, it has been suggested that the risk of developing malignancies increases only in seronegative forms (70, 71), despite some reports suggesting a slightly higher rate in subjects with anti-HMGCR (71–73). Nevertheless, other autoantibodies, namely anti-Ku and anti-Mi-2, have been confirmed not to harbor any increased risk of malignancy in patients with IIM (22, 34, 74). Rare and novel MSA have been identified in short reports of small IIM cohorts, but their association with cancer is still unknown and needs to be studied more extensively in larger cohorts worldwide (Table 1). For instance, this is the case with anti-FHL1 (75), anti-RuvBL1/2 (76–78), anti-Nup (79), and anti-SMN (80, 81) autoantibodies, which have been identified in small subsets of IIM patients, as well as in SSc and MCTD.

Cancer screening in IIM: the IMACS initiative

In 2023, the International Guideline for Idiopathic Inflammatory Myopathy-Associated Cancer Screening was released by the International Myositis Assessment and Clinical Studies Group (IMACS) (36) to provide guidance on the management of patients with suspected CAM. These guidelines enable the stratification of each patient with new-onset IIM into a 'standard,' 'moderate,' or 'high' risk of malignancy, by combining the clinical features, autoantibody status, and demographic factors such as age and sex. For instance, patients should be considered at high risk if they meet at least two of the following criteria: DM phenotype, positivity for anti-TIF1-γ or anti-NXP2, age >40 years at the onset of IIM, persistent high disease activity despite therapy, dysphagia, and cutaneous necrosis. Second, the guidelines outline a 'basic' and an 'enhanced' screening panel to be performed in a tailored manner in patients with IIM, according to their previously established cancer risk.

Therefore, all patients with IIM should participate in country- or region-specific age- and sex-appropriate cancer screening programs regardless of their individual cancer risk. Additionally, basic or enhanced screening panels should be conducted at the time of diagnosis. The 'basic screening panel' should include comprehensive history taking and physical examination, routine laboratory investigations (i.e., complete blood count, liver function tests, acute phase reactants, serum protein electrophoresis, and urinalysis), and chest X-ray. Instead, the 'enhanced screening panel' includes total body CT scan, cervical screening, mammography, dosage of the prostate-specific antigen or CA-125 (while other neoplastic markers are not recommended for general screening), pelvic or transvaginal ultrasonography, and search for fecal occult blood. Additional screening with ^{18}FDG-PET/CT and upper and lower gastrointestinal endoscopy should be considered in selected patients, based on clinical evaluation.

When evaluated in retrospective cohorts, these recommendations displayed excellent sensitivity in identifying patients with malignancy but with lower specificity. Indeed, most patients with IIM were classified as high or intermediate risk of cancer, with only a minority of subjects being represented in the standard-risk group. The ability of these guidelines to detect patients developing long-term cancers seems comparable to their effectiveness in identifying malignancies occurring close to the onset of IIM (82, 83). Further multicentric, long-term cohort studies are needed to evaluate the application of the IMACS guidelines for cancer screening and their impact on follow-up strategies (Table 1). Additionally, there is a recognized need to incorporate emerging evidence on novel risk factors to improve patient stratification (Table 1), particularly concerning serum autoantibodies, as outlined in Table 2.

Cancer and Sjogren syndrome: a model of autoimmunity-induced malignancy

PSS is a chronic autoimmune disease characterized by lymphocytic infiltration of exocrine glands, leading to glandular dysfunction and development of systemic manifestations (9). In patients with pSS the overall risk of cancer is higher compared to the general population, with an estimated standardized incidence ratio (SIR) of 2.17 (95% confidence interval—CI 1.57–3.00) (84).

Clinical features of cancer in pSS

Hematological malignancies are the most frequent life-threatening complication of pSS, with one-third of cancers being B-cell lymphomas (85). Among these, NHL is the most frequently reported, with an SIR of 13.71 (95%CI 8.83–21.29) (84), reflecting a seven to 15 times higher incidence compared with the general population (86). Although autoimmunity-promoting lymphoma is frequently observed in autoimmune diseases, this association is highly expressed in patients with pSS. Mucosal-associated lymphoid tissue (MALT) lymphoma constitutes the majority of pSS-associated NHL cases (up to 65%) and mainly originates from the salivary glands. However, additional mucosal sites can be affected, including the stomach, thyroid gland, and lungs (85). In MALT-NHL, lymphomagenesis represents the last stage of the persistent polyclonal activation of marginal zone B cells. In pSS, this activation can evolve into monoclonality, typically resulting in low- or intermediate-grade lymphomas.

In recent years, efforts have been made to identify clinical features and serological biomarkers that predict the development of MALT lymphoma in patients with pSS. Data from the HarmonicSS cohort identified positive serum rheumatoid factors as the earliest and most persistent independent predictor of lymphoma. Simultaneously, B-cell manifestations (including cryoglobulinemia and glandular, cutaneous, and hematological manifestations) appear to signal a more advanced stage in the lymphomagenesis process (87). Additional biomarkers predictive of a higher risk of NHL development have also been identified, including leukopenia, low complement C4 levels, and presence of anti-La/SSB autoantibodies (88). Major salivary gland enlargement and salivary gland focus score evaluated at the time of diagnosis have also been established as independent risk factors for lymphoma in patients with pSS. In particular, a shorter time interval from pSS to lymphoma has been described with an increasing focus score (89), highlighting the importance of histological evaluation in these patients.

A higher risk of hematologic malignancies, other than lymphoma, has been reported in patients with pSS. In these patients, the detection of monoclonal gammopathy of undetermined significance (MGUS) is common, and as a result, the documented higher prevalence of multiple myeloma is not surprising. The risk of MGUS seems restricted to patients with anti-Ro/SSA and anti-La/SSB autoantibodies (90); however, studies on its evolution to multiple myeloma are limited. Thus, further epidemiological investigations are required to precisely determine the incidence and prevalence of this complication in patients with pSS.

Solid cancers were also more frequently observed in patients with pSS (SIR 1.39). In particular, an association between thyroid and other ENT cancers, nonmelanoma skin cancer, hepatocellular carcinoma, lung cancer, prostate carcinoma, kidney, and urothelial cancers has been reported (84). Among these, thyroid cancer is the most frequently recognized, with a 2.6 SIR reported in a pSS cohort of over 7,000 patients (91). These data were confirmed by Britton Zeron et al., who described thyroid cancer as the most common solid tumor in pSS after hematological neoplasms (SIR 5.05) (92). The explanation for this association remains unclear. However, considering that the risk of developing thyroid cancer is higher in patients with autoimmune thyroiditis (93), and that autoimmune thyroiditis is one of the most frequent comorbidities in pSS (94), it is reasonable to hypothesize that the co-occurring autoimmune disease affecting the thyroid might contribute to the development of this neoplastic manifestation.

Current evidence on the established and putative risk factors for malignancy in patients with pSS is summarized in Supplementary Table 1.

Immunological features of cancer in pSS

MALT lymphoma is thought to result from local antigen-driven B-cell selection within tertiary lymphoid structures (TLS), which are typically referred to as ectopic germinal centers (GCs). It is now recognized that during pSS, ectopic GCs form in the minor salivary and/or parotid glands of approximately 30%–40% of patients (95). Since these structures host crucial phenomena, such as oligoclonal B cell expansion and somatic hypermutation of Ig variable genes (96), ectopic GCs are currently considered the 'beating heart' of the autoimmune reaction (97). However, despite these functions, the association between ectopic GC formation and lymphoma development remains unclear. While some studies have indicated

that the presence of ectopic GCs in minor salivary gland biopsies is a risk factor for NHL lymphoma development (98, 99), more recent studies have not confirmed their predictive value (100). Nevertheless, the view that ectopic GCs are markers of more active and severe diseases is widely accepted (101). Peripheral biomarkers associated with ectopic GCs formation, such as CXCL13, have been identified (102) and are currently being used in clinical trials to monitor disease progression. Notably, elevated peripheral levels of CXCL13 appear to be associated with an increased risk of NHL, further strengthening the relationship between ectopic GC formation and hematologic malignancy development (103, 104).

Cancer screening in pSS

Lymphoproliferative disease surveillance remains a challenge in patients with pSS even after stratification according to patient risk. Recent studies have shown that patients without clinical suspicion of lymphoma or increased systemic disease activity are unlikely to benefit from major salivary gland imaging screening for detecting this complication (105). This issue is compounded by evidence of the poor reliability of salivary gland ultrasound protocols and scores in identifying lymphoma in patients with pSS and high clinical suspicion (106). It has been proposed that combining salivary gland ultrasound with histology could improve the detection of patients at the highest risk of lymphoma (106). However, evidence is still lacking regarding optimal screening strategies, imaging modalities, and timing. Efforts should also focus on detecting lymphoproliferative diseases at sites other than the major salivary glands, including both the nodal and extranodal sites. Furthermore, identifying the risk factors and screening protocols for non-lymphoproliferative neoplasms should also constitute a priority in the research agenda (Table 1).

Cancer and systemic sclerosis: a unique scenario for both malignancy-induced autoimmunity and autoimmunity-induced malignancy

Systemic sclerosis (SSc) is associated with an increased risk of malignancy, with cancers being diagnosed at a significantly younger age compared to the general population (17, 107–112), and is a leading cause of death among patients (113–115). Cancer strongly affects the disease course of SSc (110, 116), particularly when diagnosed close to the onset of rheumatological manifestations (117). Breast, lung, and hematologic cancers, including lymphoid and myeloid neoplasms, are most frequently diagnosed in patients with SSc (17, 118–120), but increased rates of melanoma and non-melanoma skin cancers, hepatocellular carcinoma, urothelial (119), and thyroid cancers, particularly in cases of coexistent autoimmune thyroiditis (121), have also been reported.

Risk factors for cancer in patients with SSc include demographic and clinical features, disease duration, selected complications, and the presence (or absence) of particular autoantibodies (120, 122–124). However, a clear profile of the patient with SSc 'at risk of malignancy' remains elusive due to the complex interplay between such characteristics and additional risk factors (e.g., family history, exposure to smoking, air pollutants, ionizing radiation, etc.). Compelling evidence suggests that in patients with SSc, some cancers are diagnosed close to the onset of autoimmune manifestations, akin to paraneoplastic phenomena, whereas others exhibit a characteristic delay, often correlating with an increased burden of organ damage (125). These aspects will be discussed in the following sections and summarized in Figure 2.

Clinical features of cancer in SSc

Given the short interval that is seldom observed between the onset of SSc and the diagnosis of cancer, a subset of SSc cases is thought to represent a paraneoplastic syndrome (120, 125, 126), referred to as 'cancer-associated scleroderma.' This subset may include patients in whom the antitumor immune response culminates in the onset of autoimmunity (127). From a clinical perspective, early diffuse and rapidly progressive SSc is associated with a high risk of synchronous malignancy (128, 129), particularly in the presence of certain serum autoantibodies.

A second peak of incident malignancies occurs in patients with a long history of SSc and related complications (125), such as pulmonary arterial hypertension and interstitial lung disease (ILD) (119, 124), particularly in cases of progressive fibrosis (120). Chronic inflammation has long been associated with an increased risk of malignancy (4), and what is observed in the SSc scenario could fit within this frame. For instance, this is the case for lung cancer, which arises more frequently in patients with ILD and established disease (123). However, while esophageal involvement is common in SSc, no increased risk of esophageal malignancy has been reported to date. Further research is warranted to test whether the presence of factors considered as 'protective' from cancer (i.e., limited cutaneous disease, anticentromere autoantibodies—ACA) (130, 131) is linked to smoldered cancer incidence in this patient subset.

Immunological features of cancer in SSc

Positivity for anti-RNA polymerase III (POLR3) autoantibodies has traditionally been linked to an increased risk of overall (120, 130, 132, 133) and synchronous cancers (111, 130, 133–135), mostly in patients with diffuse disease (131). Support for the association between the two conditions was elegantly provided by the evidence of alterations in the *POLR3A* locus in samples of synchronous cancers derived from patients with anti-POLR3^{+} SSc, but not in negative cases (15). However, conflicting data on the risk of malignancy with anti-POLR3 autoantibodies have been reported

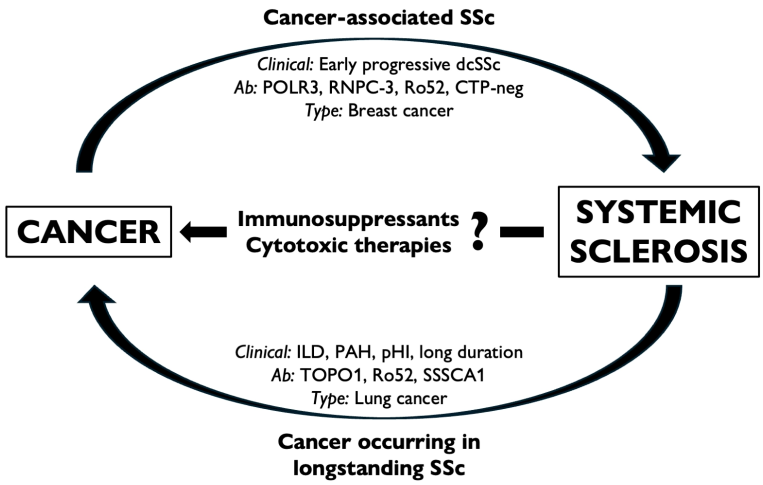


FIGURE 2
The interplay between cancer and SSc. Some forms of SSc can be regarded as cancer-associated (or paraneoplastic) scleroderma, in which the putative etiological role of malignancy is supposed to trigger the onset of autoimmunity in predisposed individuals (a). Cancer can also occur in longstanding SSc, particularly at specific sites and is associated with the selection of risk factors, phenotypes, and disease complications (b). Immunosuppressive and cytotoxic treatments are commonly adopted to treat SSc-related complications; however, the putative role of such therapies remains elusive (c). CTP-neg, 'CTP-negative' patients; dcSSc, diffuse cutaneous SSc; ILD, interstitial lung disease; PAH, pulmonary arterial hypertension; pHI, primary heart involvement.

in some cohorts (111, 136, 137). Apart from possibly reflecting genetic or epigenetic differences, such heterogeneity could also indicate the role of multiple autoantibody specificities in modulating the rate of cancers (127, 138). Indeed, similar to what was recently described in DM (51, 52), multiple serum autoantibody specificities likely confer a protective role against malignancy in patients with another autoantibody traditionally linked to an increased risk of cancer. A significant difference in the rate of neoplasms has been observed in anti-POLR3 positive patients with or without concurrent autoantibodies (130, 137). An increased risk of cancer-associated scleroderma has been also reported in patients without anticentromere (ACA), anti-Topoisomerase-I (TOPO1), and anti-POLR3 autoantibodies, the so-called CTP-negative cases (131), as well as in ANA-negative SSc cases (139). Mecoli et al. demonstrated a protective role of anti-Th/To in cancer-associated scleroderma (140). Since the Th/To complex is composed of four molecular subunits (140), it would be useful to investigate correlations between the rate of malignancies based on the presence of single vs. multiple autoantibodies directed towards the different subunits. Similar considerations could be made in patients with anti-POLR3, notably directed to RP155 and/or RP11 subunits of RNA polymerase III (141), and autoantibodies to the PM/Scl complex, which includes a 75 KDa and a 100 KDa subunit and have been associated with malignancy in Spanish patients (120, 142).

Among the rarer autoantibodies, anti-U3-RNP/fibrillarin (138) and anti-RNPC-3, usually associated with limited cutaneous disease but severe organ involvement, have been correlated with cancer-associated scleroderma, along with a worse prognosis, comparable to that observed with anti-POLR3 (143). In particular, a short SSc-cancer interval has been described for anti-RNPC-3 in an American cohort (143), although no association with malignancy was found in another European cohort (144). However, while the first study primarily focused on the characteristics of anti-RNPC-3+ patients and their association with cancer, the European study aimed to characterize the features of patients who tested positive vs. negative for that autoantibody. Moreover, different autoantibody detection methods have been used (143, 144), which could have influenced the results.

Breast cancer is the most frequent malignancy diagnosed as cancer-associated scleroderma, particularly in the presence of anti-POLR3 (136) and diffuse disease (131). Interestingly, breast cancer and SSc share select molecular pathways, including hyperactivation of the mammalian target of rapamycin (mTOR), phosphatidylinositol 3-kinase (PI3K), and transforming growth factor beta (TGF-β) (145). In addition, tumor-infiltrating lymphocytes are more abundant in breast cancers of patients compared than in those without autoimmune disease (145). These observations support the hypothesis of a possible interplay between the anticancer response and the onset of autoimmunity in cancer-associated scleroderma. Further research is required to understand the prognostic role and therapeutic impact of these observations from both the oncological and rheumatological perspectives.

Serum autoantibodies also played a significant role in stratifying patients according to the risk of late-onset malignancy (Table 1). Anti-topoisomerase I (TOPO1) positivity is a potential risk factor, particularly for lung cancer. However, it is unclear whether autoantibodies themselves, their association with ILD, or both are putative risk factors for malignancy (110, 146). Late-onset cancer occurs more frequently with the recently described anti-SSSCA1 antibody, an emerging predictor of SSc-related primary heart involvement, which may support the hypothesis of a correlation between long-standing SSc, organ damage, and incident malignancies (147). Anti-SSA/Ro autoantibodies, often detected in patients with

SSc and high burden of visceral involvement (148, 149), have been associated to late-onset cancers in a French SSc cohort. A large case–control study attributed this correlation specifically to positivity for the anti-Ro52 subset (130). This result was retrospectively validated by our group in an independent cohort of patients with SSc (137), suggesting a more intricate role of anti-Ro52 positivity. Indeed, cancer-associated scleroderma was more frequently reported when anti-Ro52 was found to be the sole autoantibody, whereas its positivity in combination with other specificities correlated with higher rates of overall cancer throughout the disease history of patients with SSc (137).

Table 3 summarizes the current evidence on the association between serum autoantibodies and cancer risk in patients with SSc.

Cancer screening in SSc

Patients with SSc represent an ideal population for implementing tailored cancer screening strategies because of the potential existence of different risk categories, as recently proposed for IIM (36). Recommendations for cancer screening were proposed by a panel of experts and are specifically meant for patients with new-onset SSc and anti-POLR3 autoantibodies (133). The panel pointed to the need to exclude synchronous malignancy, particularly of the breast, with regular screening suggested thereafter according to age- and sex-related risk factors (133). Despite preliminary evidence demonstrating the predictive role of seriate monitoring of tumor-associated antigen serum levels (150), a panel of experts discouraged their dosage *a priori* in patients with SSc, similar to that in the general population (133). However, the proposed recommendations are only applicable to anti-POLR3 positive patients. Thus, a tailored cancer-screening strategy for SSc remains largely speculative.

Cancer screening should be a priority, and tools to allow patient stratification into different risk clusters are needed. Such clusters may ideally benefit from different screening strategies at different time points during the disease course. As mentioned in the previous sections, the interplay of a wide range of features should be considered to assess the risk of malignancy in patients with SSc, including the disease phenotype, presence and severity of complications, serum autoantibodies, and traditional risk factors, such as tobacco exposure and family history. Finally, it would be interesting to verify whether repeated testing for serum autoantibodies could intercept changes in the autoimmune repertoire, which might help stratify the risk of incident cancer in patients with SSc during the follow-up period (Table 1).

Immunosuppressive treatments and cancer in SSc

Patients with SSc-related organ involvement are treated with immunosuppressive and/or cytotoxic therapies, raising concern for secondary cancers (151, 152) as supported by the observation of urothelial cancers occurring after exposure to cyclophosphamide (119, 120, 153). Mycophenolate mofetil (MMF) is commonly used for the treatment of SSc and is particularly effective in ILD (154, 155). Evidence mostly derived from transplant immunology has not raised

TABLE 3 Systemic sclerosis-specific and -associated autoantibodies, clinical associations and current evidence regarding cancer risk.

Autoantibody	Target antigen	Clinical associations	Cancer risk
anti-TOPO1/Scl-70	Topoisomerase I	dcSSc, ILD	Likely increased** (110, 146)
anti-CENP-A/B	Centromere proteins	lcSSc, PAH, DU, calcinosis, gastrointestinal disease	Not increased (130)
anti-POLR3	RNA polymerase III	Rapidly progressive dcSSc, SRC, GAVE	Increased*** (133, 137, 138)
anti-Th/To	RNase P Nucleolar Protein Complex	lcSSc, ILD, PAH	Not increased (140)
anti-NOR90	Nucleolar Organizer Region 90 KDa	lcSSc, mild disease	Not increased (130)
anti-PM/Scl	Nucleolar macro-molecular complex of 75 KDa and 100 KDa	arthritis, myositis, ILD	Likely increased (120)
anti-Ro52	Tripartite motif-containing protein 21	lcSSc, ILD, PAH, overlap pSS	Likely increased# (124, 130, 137)
anti-U3-RNP	Fibrillarin	higher mRSS, myositis	Likely increased## (138)
anti-RNPC-3	RNA Binding Region Containing 3 (U11/U12-RNP)	ILD, gastrointestinal dysmotility	Increased## (143)
anti-SSSCA1	autoantigen p27 (centromere-associated protein)	cardiac involvement*, pSS overlap	Increased## (147)

Due to relatively poor evidence concerning cancer risk, compared to IIM, cancer risk is reported as 'increased,' 'possibly increased,' or 'not increased,' according to relevant literature discussed in the main text. Results are mainly derived from observational cohort or case-control studies. In particular, multicentric cohort studies were available for anti-TOPO1, anti-POLR3, anti-CENP-A/B, anti-Th/To, and anti-PM/Scl autoantibodies.

dcSSc, diffuse cutaneous systemic sclerosis; DU, digital ulcers; GAVE, gastric antral vascular ectasia; ILD, interstitial lung disease; lcSSc, limited cutaneous systemic sclerosis; mRSS, modified Rodnan skin score; PAH, pulmonary arterial hypertension; pSS, Sjogren syndrome; SRC, scleroderma renal crisis.

* Defined as evidence of impaired left ventricle function and/or signs of right failure and/or clinically significant arrhythmia.

** Evidence suggests particularly for long-term incidence of lung cancer.

*** Conflicting evidence pointing towards increased risk only in the absence of multiple autoantibody positivity.

Evidence suggesting increased risk particularly in patients without multiple autoantibody positivity.

Evidence from single studies or small case series.

major concerns regarding the oncological risk of MMF (155–157), except for the possibly increased rate of non-melanoma skin cancers (158). While no study has specifically evaluated the risk of cancer in patients with SSc treated with MMF, drug safety was suggested in a large cohort of patients treated for fibrotic lung diseases (159), as well as in patients with SSc (138). We hypothesized that the antiproliferative effects of MMF (155) modulate the humoral immune response without affecting cell-mediated immunity (160), thus minimally impairing immune surveillance towards malignancy. Finally, current data are insufficient to establish any association between cancer incidence and more innovative treatments (e.g., rituximab and tocilizumab) in patients with SSc (161) (Table 1).

Cancer and systemic lupus erythematosus: still an unclear scenario

The dual role of immune activation in SLE—driving autoimmunity while potentially influencing tumor suppression or promotion—creates a paradox that is central to understanding the relationship between SLE and cancer. A recent meta-analysis revealed a pronounced increase (2.87-fold; 95%CI 2.49–3.24) in the standardized mortality ratio (SMR) for all-cause mortality among SLE patients compared to the general population (162). Despite the heterogeneity among the included studies, an elevated cancer-related mortality risk (SMR 1.7-fold) was reported in SLE patients (163). The overall cancer risk profile in SLE is shaped by a heterogeneous set of factors, including disease activity and damage, immunosuppressive treatments, genetic predisposition, and environmental exposure (164).

From an epidemiological perspective, SLE displays a unique cancer risk profile. Hematologic malignancies (NHL, Hodgkin lymphoma, leukemia, and myeloma), and lung, cervical, thyroid, gastrointestinal, hepatobiliary, and liver cancers occur more frequently in SLE, which is partly attributed to chronic immune activation and persistent inflammation. Conversely, breast, endometrial, and prostate cancers and melanoma are less common, possibly due to alterations in hormonal pathways and immune surveillance mechanisms (163, 165).

Clinical features of cancer in SLE

Specific features of SLE, such as hematological and pulmonary manifestations, may contribute to cancer risk, namely NHL and lung cancer. However, despite the well-established association between idiopathic pulmonary fibrosis and lung neoplasms, pulmonary fibrosis is rarely reported in SLE and has not shown statistically significant associations, despite evidence of increasing trends (166). A higher SLICC/ACR Damage Index has emerged as a risk factor for cancer (167, 168); however, the relationship with disease activity risk remains unclear (168) (Table 1).

Secondary and overlapping autoimmune diseases, such as Sjogren's syndrome, autoimmune liver disease, scleroderma, and autoimmune thyroiditis, may contribute to cancer risk in SLE (169) (Table 1). For instance, secondary Sjogren's syndrome increases the risk of NHL (168), although the predominance of the DLBCL subtype raises questions about Sjogren's status as the primary driver (170). Autoimmune thyroiditis is strongly linked to thyroid cancer in SLE patients, as supported by evidence of thyroid autoimmunity in most cases of thyroid cancer in this population (171).

Childhood-onset SLE (cSLE) is a disease subset that warrants particular attention regarding cancer risk. Lymphomas and solid tumors have been reported at a significant rate, with a median time of 10 years after cSLE diagnosis. Distinct clinical presentations, risk factors, and treatment challenges have been outlined in this population, underscoring the need for heightened vigilance and tailored management strategies for young patients (172).

Finally, patients with SLE may be more susceptible to oncogenic viruses such as Epstein–Barr virus (EBV) (169), human papillomavirus (HPV), and hepatitis B virus (HBV). Impaired immune surveillance could lead to higher rates of viral persistence and reactivation, contributing to the development of lymphomas (173), cervical dysplasia and cancer (174), and hepatocellular carcinoma. By weakening the antiviral defenses, immunosuppressive therapies may further increase this risk.

Current evidence on the established and putative risk factors for malignancy in patients with SLE is summarized in Supplementary Table 2.

Immunological features of cancer in SLE

Chronic inflammation plays a key role in fostering a pro-oncogenic microenvironment via DNA damage, oxidative stress, and cytokine-mediated pathways (175, 176). For instance, the increased risk of lymphoma may be driven by cytokines upregulated in SLE, such as BAFF, APRIL, IL-6, and IL-10, which promote B-cell survival, proliferation, and inflammation (177). These factors are linked to non-germinal center B-cell-like DLBCL, the predominant lymphoma subtype in SLE (169, 178).

SLE-associated autoantibodies, a hallmark of the disease, are hypothesized to promote tumor development by entering cells and causing DNA damage (179). Notably, an anti-DNA autoantibody named 3E10 has been shown to enter cell nuclei, bind to DNA, and impair key DNA repair pathways, thereby contributing to genomic instability. By increasing susceptibility to DNA damage, 3E10 provides a compelling link between SLE autoimmunity and malignancy (180).

Moreover, specific genetic variants (e.g., SNPs in CD40 and HLA alleles) have been associated with both SLE and malignancy, particularly DLBCL and lung cancer (181), although some findings suggest pleiotropy or linkage disequilibrium rather than direct biological causation (182). Emerging research has also identified epigenetic mechanisms, particularly microRNA dysregulation, implicated in both SLE pathogenesis and hematologic cancers, highlighting the potential role of shared post-transcriptional regulatory pathways in the concurrent development of autoimmunity and malignancy (183).

SLE might also confer protection against hormone-sensitive cancers, possibly because of lower exposure to estrogens and androgens. Indeed, women with SLE often experience earlier menopause (184) and are less frequently prescribed estrogen-containing medications (185), whereas men with SLE have lower androgen levels (186). Moreover, certain autoimmune mechanisms may yield protective effects, as in the case of 5C6 anti-DNA autoantibodies that selectively target tumor cells with defects in DNA repair processes (e.g., BRCA2-deficient cancer cells) (187). While the rates of hormone-susceptible breast cancers are similar among SLE patients and the general population, patients with SLE experience a significantly lower incidence of triple-negative cancers, which are mostly characterized by genetic mutations in DNA repair pathways (188).

Cancer screening in SLE

Established recommendations for cancer screening in patients with SLE are unavailable. Thus, these procedures largely rely on expert opinions, substantially overlapping with what is recommended in the general population (189). In particular, cervical screening and/or HPV vaccinations, periodic mammograms, and fecal occult blood testing are advised for all patients according to age- and sex-specific local guidelines (189). Moreover, clinical screening through regular lymph node examination and routine chemistry is recommended for hematological malignancies, while thyroid enzymes, autoantibodies, and ultrasound should be performed because of the risk of thyroid neoplasms (189). Apart from pursuing smoking cessation, lung cancer screening with annual chest CT scans is recommended only in patients with a high-risk profile (i.e., aged 50 years–75 years and with a history of smoking) (189), while hepatobiliary screening is not recommended unless in cases of positive HBV or HCV serologies (189), and urinary cytology is recommended periodically in patients who have undergone cyclophosphamide.

However, a large cohort study demonstrated that adherence to cancer screening is an issue in patients with SLE, with at least 25% of patients not being regularly screened, particularly in cases of established and longstanding disease (190). This seems particularly crucial regarding cervical cancer screening, since patients with SLE are at higher risk of abnormal test results compared with controls (191).

Immunosuppressive treatments and cancer in SLE

Immunosuppressive treatments can influence the risk of cancer in SLE (192) because their long-term use may impair immune surveillance (193). Prolonged and cumulative high-dose cyclophosphamide has been strongly linked to an elevated risk of bladder cancer (with oral cyclophosphamide) and hematological malignancies (189). Similarly, azathioprine has been associated with a risk of hematologic malignancies (164), highlighting the need for careful monitoring and optimal dosing. Moreover, the use of immunosuppressive therapies is associated with a higher risk of cervical neoplasia than antimalarials (194), underscoring the importance of regular screening in these patients.

Calcineurin inhibitors have been associated with an increased incidence of cancers in solid organ transplant recipients (195), with previous studies suggesting their role in impairing DNA repair, promoting angiogenesis, and facilitating tumor invasion (196). However, a recent large cohort study of SLE patients with consistent follow-up found no significant difference in cancer risk between those using calcineurin inhibitors and those who did not, even after adjusting for potential confounders (197). Biologics that target B-cell pathways, such as rituximab and belimumab, are generally considered safe; however, their effects on cancer remain the subject of ongoing investigation. Finally, owing to the close association between drug exposure and disease activity, many studies face challenges in distinguishing the individual contributions of these factors to cancer risk (Table 1).

Compared to immunosuppressants, hydroxychloroquine, which is universally prescribed for SLE, has been associated with a decreased cancer risk (198), particularly for breast and non-melanoma skin cancer (193), possibly because of its anti-proliferative and anti-angiogenic activity.

Limitations and concluding remarks

While this study aimed to provide insight into the dual-faceted clinical relationship between cancer and CTDs (i.e., cancer-associated CTDs vs. cancer occurring subsequently or within the context of CTDs), we acknowledge certain limitations. Although our literature review was comprehensive and sought to analyze evidence that supports and challenges our hypotheses, we did not follow a systematic review approach, which would be necessary to address more specific research questions based on the current evidence. A consistent approach was attempted across diseases, but the major differences in evidence availability led to some degree of heterogeneity, particularly in the immunological feature sections related to myositis and SSc *versus* pSS and SLE. Publication bias should also be considered, particularly regarding data on rare and emerging autoantibody specificities, along with the relatively greater abundance of studies on certain diseases, primarily IIM and SSc, compared to pSS and SLE. There are also biases in the races and ethnicities that have been studied in different diseases, which should be addressed in future investigations. The heterogeneity of analytical methods for autoantibody detection (e.g., immunoprecipitation, line blot, and ELISA) should also be considered when comparing different studies, as the sensitivity and specificity vary depending on the techniques used and the target autoantigen (199). Moreover, our objective was to highlight unmet needs and identify avenues for future research in autoimmunity and rheumatology, with potentially significant implications from the clinical, pathophysiological, and therapeutic perspectives.

Patients with CTDs exhibit distinct cancer risk profiles, which are influenced by the etiological role of malignancy in certain contexts and the precancerous environment created by chronic inflammation and autoimmune activation. Similarities in immune pathogenesis are thought to occur among patients with

paraneoplastic forms of CTDs, as seen when comparing findings from anti-TIF1-γ+ DM and anti-POLR3+ SSc, in which the complex interplay between cancer-related mutations and aberrant tumor immune editing is thought to culminate in the activation of self-reactive lymphocytes, ultimately leading to tissue damage and CTD onset. On the other hand, chronic immune activation reflecting specific pathogenic clues can be considered a potentially premalignant condition, as suggested by the evidence of an increased risk of lung cancer in patients with longstanding SSc-ILD. From this point of view, the example provided by pSS is paradigmatic, since the disease itself is responsible for the generation of autoreactive lymphocyte clones with lymphoma-prone behavior, ultimately culminating in MALT-NHL onset. Most importantly, a correlation between disease activity and lymphoma risk has been clearly demonstrated in pSS. The role of immunosuppressive therapies in cancer risk in these patients remains unclear. Therefore, further research is needed to unravel the complex interplay between CTDs and malignancy, which requires a multidisciplinary approach that integrates clinical and pathophysiological aspects (Table 1). Addressing this challenge is essential to improve cancer screening, prevention, and treatment strategies in this patient population.

Author contributions

AT: Conceptualization, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing. AC: Conceptualization, Supervision, Visualization, Writing – original draft, Writing – review & editing. EG: Conceptualization, Supervision, Writing – original draft, Writing – review & editing. SC: Writing – original draft, Writing – review & editing. MD: Conceptualization, Supervision, Writing – original draft, Writing – review & editing. CS: Conceptualization, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was partially supported by "Ricerca Corrente" funding from the Italian Ministry of Health to the IRCCS Humanitas Research Hospital.

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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2025.1571700/full#supplementary-material

References

1. Masetti R, Tiri A, Tignanelli A, Turrini E, Argentiero A, Pession A, et al. Autoimmunity and cancer. *Autoimmun Rev*. (2021) 20:102882. doi: 10.1016/j.autrev.2021.102882

2. De Vita S. Treatment of mixed cryoglobulinemia: a rheumatology perspective. *Clin Exp Rheumatol*. (2011) 29:S99–103.

3. Carsons S. The association of Malignancy with rheumatic and connective tissue diseases. *Semin Oncol*. (1997) 24:360–72.

4. Coussens LM, Werb Z. Inflammation and cancer. *Nature*. (2002) 420:860–7. doi: 10.1038/nature01322

5. Dantal J, Soulillou JP. Immunosuppressive drugs and the risk of cancer after organ transplantation. *N Engl J Med*. (2005) 352:1371–3. doi: 10.1056/NEJMe058018

6. Buchanich JM, Newcomb CW, Washington TL, Foster CS, Sobrin L, Thorne JE, et al. Use of immunosuppression and subsequent cancer incidence: cohort study. *BMJ Oncol*. (2023) 2:e000037. doi: 10.1136/bmjonc-2023-000037

7. Furst DE, Matucci Cerinic M. Connective tissue diseases. *Best Pract Res Clin Rheumatol*. (2007) 21:969–70. doi: 10.1016/j.berh.2007.11.001

8. Hoi A, Igel T, Mok CC, Arnaud L. Systemic lupus erythematosus. *Lancet*. (2024) 403:2326–38. doi: 10.1016/S0140-6736(24)00398-2

9. Mariette X, Criswell LA. Primary Sjögren's syndrome. *N Engl J Med*. (2018) 378:931–9. doi: 10.1056/NEJMcp1702514

10. Lundberg IE, Fujimoto M, Vencovsky J, Aggarwal R, Holmqvist M, Christopher-Stine L, et al. Idiopathic inflammatory myopathies. *Nat Rev Dis Primers*. (2021) 7:86. doi: 10.1038/s41572-021-00321-x

11. Volkmann ER, Andréasson K, Smith V. Systemic sclerosis. *Lancet*. (2023) 401 (10373):304–18. doi: 10.1016/S0140-6736(22)01692-0

12. Tanaka Y. Is mixed connective tissue disease (MCTD) a subtype of systemic sclerosis? *Semin Arthritis Rheum*. (2025) 72S:152678. doi: 10.1016/j.semarthrit.2025.152678

13. Higgs BW, Liu Z, White B, Zhu W, White WI, Morehouse C, et al. Patients with systemic lupus erythematosus, myositis, rheumatoid arthritis and scleroderma share activation of a common type I interferon pathway. *Ann Rheum Dis*. (2011) 70:2029–36. doi: 10.1136/ard.2011.150326

14. Schett G, Nagy G, Krönke G, Mielenz D. B-cell depletion in autoimmune diseases. *Ann Rheum Dis*. (2024) 83:1409–20. doi: 10.1136/ard-2024-225727

15. Joseph CG, Darrah E, Shah AA, Skora AD, Casciola-Rosen LA, Wigley FM, et al. Association of the autoimmune disease scleroderma with an immunologic response to cancer. *Science*. (2014) 343:152–7. doi: 10.1126/science.1246886

16. De Vita S, Gandolfo S. Predicting lymphoma development in patients with Sjögren's syndrome. *Expert Rev Clin Immunol*. (2019) 15:929–38. doi: 10.1080/1744666X.2019.1649596

17. Bonifazi M, Tramacere I, Pomponio G, Gabrielli B, Avvedimento EV, La Vecchia C, et al. Systemic sclerosis (scleroderma) and cancer risk: systematic review and meta-analysis of observational studies. *Rheumatol (Oxford)*. (2013) 52:143–54. doi: 10.1093/rheumatology/kes303

18. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). *N Engl J Med*. (1975) 292:344–7. doi: 10.1056/NEJM197502132920706

19. Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). *N Engl J Med*. (1975) 292:403–7. doi: 10.1056/NEJM197502202920807

20. Ponyi A, Constantin T, Garami M, András C, Tállai B, Váncsa A, et al. Cancer-associated myositis: clinical features and prognostic signs. *Ann N Y Acad Sci*. (2005) 1051:64–71. doi: 10.1196/annals.1361.047

21. Selva-O'Callaghan A, Trallero-Araguás E, Ros J, Gil-Vila A, Lostes J, Agustí A, et al. Management of cancer-associated myositis. *Curr Treatm Opt Rheumatol*. (2022) 8:91–104. doi: 10.1007/s40674-022-00197-2

22. McHugh NJ, Tansley SL. Autoantibodies in myositis. *Nat Rev Rheumatol*. (2018) 14:290–302. doi: 10.1038/nrrheum.2018.56

23. Li Y, Jia X, Sun X, Shi L, Lin F, Gan Y, et al. Risk factors for cancer-associated myositis: A large-scale multicenter cohort study. *Int J Rheum Dis*. (2021) 24:268–73. doi: 10.1111/1756-185X.14046

24. Kardes S, Gupta L, Aggarwal R. Cancer and myositis: Who, when, and how to screen. *Best Pract Res Clin Rheumatol*. (2022) 36:101771. doi: 10.1016/j.berh.2022.101771

25. Sigurgeirsson B, Lindelöf B, Edhag O, Allander E. Risk of cancer in patients with dermatomyositis or polymyositis. A population-based study. *N Engl J Med*. (1992) 326:363–7. doi: 10.1056/NEJM199202063260602

26. Mecoli CA, Igusa T, Chen M, Wang XY, Albayda J, Paik JJ, et al. Subsets of idiopathic inflammatory myositis enriched for contemporaneous cancer relative to the general population. *Arthritis Rheumatol*. (2023) 75(4):620–9. doi: 10.1002/art.42311

27. Fiorentino DF, Chung LS, Christopher-Stine L, Zaba L, Li S, Mammen AL, et al. Most patients with cancer-associated dermatomyositis have antibodies to nuclear matrix protein NXP-2 or transcription intermediary factor 1γ. *Arthritis Rheumatol*. (2013) 65:2954–62. doi: 10.1002/art.v65.11

28. Mainetti C, Terziroli Beretta-Piccoli B, Selmi C. Cutaneous manifestations of dermatomyositis: a comprehensive review. *Clin Rev Allergy Immunol*. (2017) 53:337–56. doi: 10.1007/s12016-017-8652-1

29. O'Connell KA, LaChance AH. Dermatomyositis. *N Engl J Med*. (2021) 384:2437. doi: 10.1056/NEJMicm2033425

30. Hum RM, Lilleker JB, Lamb JA, Oldroyd AGS, Wang G, Wedderburn LR, et al. Comparison of clinical features between patients with anti-synthetase syndrome and dermatomyositis: Results from the MYONET registry. *Rheumatol (Oxford)*. (2024) 63(8):2093–100. doi: 10.1093/rheumatology/kead481

31. Shelly S, Beecher G, Milone M, Liewluck T, Ernste F, Triplett J, et al. Cancer and immune-mediated necrotizing myopathy: a longitudinal referral case-controlled outcomes evaluation. *Rheumatol (Oxford)*. (2022) 62:281–9. doi: 10.1093/rheumatology/keac144

32. Mahé E, Descamps V, Burnouf M, Crickx B. A helpful clinical sign predictive of cancer in adult dermatomyositis: cutaneous necrosis. *Arch Dermatol*. (2003) 139(4):539. doi: 10.1001/archderm.139.4.539-a

33. Labeit B, Pawlitzki M, Ruck T, Muhle P, Claus I, Suntrup-Krueger S, et al. The impact of dysphagia in myositis: A systematic review and meta-analysis. *J Clin Med*. (2020) 9:2150. doi: 10.3390/jcm9072150

34. Oldroyd AGS, Allard AB, Callen JP, Chinoy H, Chung L, Fiorentino D, et al. A systematic review and meta-analysis to inform cancer screening guidelines in idiopathic inflammatory myopathies. *Rheumatol (Oxford)*. (2021) 60:2615–28. doi: 10.1093/rheumatology/keab166

35. Moghadam-Kia S, Oddis CV, Ascherman DP, Aggarwal R. Risk factors and cancer screening in myositis. *Rheum Dis Clin North Am*. (2020) 46:565–76. doi: 10.1016/j.rdc.2020.05.006

36. Oldroyd AGS, Callen JP, Chinoy H, Chung L, Fiorentino D, Gordon P, et al. International guideline for idiopathic inflammatory myopathy-associated cancer screening: an international myositis assessment and clinical studies group (IMACS) initiative. *Nat Rev Rheumatol*. (2023) 19(12):805–17. doi: 10.1038/s41584-023-01045-w

37. Lu X, Yang H, Shu X, Chen F, Zhang Y, Zhang S, et al. Factors predicting Malignancy in patients with polymyositis and dermatomyostis: a systematic review and meta-analysis. *PloS One*. (2014) 9:e94128. doi: 10.1371/journal.pone.0094128

38. Uchino M, Yamashita S, Uchino K, Mori A, Hara A, Suga T, et al. Muscle biopsy findings predictive of Malignancy in rare infiltrative dermatomyositis. *Clin Neurol Neurosurg*. (2013) 115:603–6. doi: 10.1016/j.clineuro.2012.07.019

39. Hill CL, Zhang Y, Sigurgeirsson B, Pukkala E, Mellemkjaer L, Airio A, et al. Frequency of specific cancer types in dermatomyositis and polymyositis: a population-based study. *Lancet*. (2001) 357:96–100. doi: 10.1016/S0140-6736(00)03540-6

40. Chen YJ, Wu CY, Huang YL, Wang CB, Shen JL, Chang YT. Cancer risks of dermatomyositis and polymyositis: a nationwide cohort study in Taiwan. *Arthritis Res Ther*. (2010) 12:R70. doi: 10.1186/ar2987

41. Hsu JL, Liao MF, Chu CC, Kuo HC, Lyu RK, Chang HS, et al. Reappraisal of the incidence, various types and risk factors of Malignancies in patients with dermatomyositis and polymyositis in Taiwan. *Sci Rep*. (2021) 11:4545. doi: 10.1038/s41598-021-83729-5

42. Selva-O'Callaghan A, Ros J, Gil-Vila A, Vila-Pijoan G, Trallero-Araguás E, Pinal-Fernandez I. Malignancy and myositis, from molecular mimicry to tumor infiltrating lymphocytes. *Neuromuscul Disord*. (2019) 29:819–25. doi: 10.1016/j.nmd.2019.09.014

43. Levine SM. Cancer and myositis: new insights into an old association. *Curr Opin Rheumatol*. (2006) 18:620–4. doi: 10.1097/01.bor.0000245721.02512.77

44. De Vooght J, Vulsteke JB, De Haes P, Bossuyt X, Lories R, De Langhe E. Anti-TIF1-γ autoantibodies: warning lights of a tumour autoantigen. *Rheumatol (Oxford)*. (2020) 59:469–77. doi: 10.1093/rheumatology/kez572

45. Trallero-Araguás E, Rodrigo-Pendás JÁ, Selva-O'Callaghan A, Martínez-Gómez X, Bosch X, Labrador-Horrillo M, et al. Usefulness of anti-p155 autoantibody for diagnosing cancer-associated dermatomyositis: a systematic review and meta-analysis. *Arthritis Rheumatol*. (2012) 64:523–32. doi: 10.1002/art.33379

46. Aussy A, Fréret M, Gallay L, Bessis D, Vincent T, Jullien D, et al. The IgG2 isotype of anti-transcription intermediary factor 1γ Autoantibodies is a biomarker of cancer and mortality in adult dermatomyositis. *Arthritis Rheumatol*. (2019) 71:1360–70. doi: 10.1002/art.40895

47. Cordel N, Dechelotte B, Jouen F, Lamb JA, Chinoy H, New P, et al. Anti-transcription intermediary factor 1-gamma IgG2 isotype is associated with cancer in adult dermatomyositis: an ENMC multinational study. *Rheumatol (Oxford)*. (2023) 62:1711–5. doi: 10.1093/rheumatology/keac577

48. Pinal-Fernandez I, Ferrer-Fabregas B, Trallero-Araguas E, Balada E, Martínez MA, Milisenda JC, et al. Tumour TIF1 mutations and loss of heterozygosity related to cancer-associated myositis. *Rheumatol (Oxford)*. (2018) 57:388–96. doi: 10.1093/rheumatology/kex413

49. Mohassel P, Rosen P, Casciola-Rosen L, Pak K, Mammen AL. Expression of the dermatomyositis autoantigen transcription intermediary factor 1γ in regenerating muscle. *Arthritis Rheumatol*. (2015) 67:266–72. doi: 10.1002/art.38863

50. Scholtissek B, Ferring-Schmitt S, Maier J, Wenzel J. Expression of the autoantigen TRIM33/TIF1γ in skin and muscle of patients with dermatomyositis is upregulated, together with markers of cellular stress. *Clin Exp Dermatol*. (2017) 42:659–62. doi: 10.1111/ced.2017.42.issue-6

51. Fiorentino D, Mecoli CA, Igusa T, Albayda J, Paik JJ, Tiniakou E, et al. Association of anti-CCAR1 autoantibodies with decreased cancer risk relative to the general population in patients with anti-transcriptional intermediary factor 1γ-positive dermatomyositis. *Arthritis Rheumatol*. (2023) 75:1238–45. doi: 10.1002/art.42474

52. Hosono Y, Sie B, Pinal-Fernandez I, Pak K, Mecoli CA, Casal-Dominguez M, et al. Coexisting autoantibodies against transcription factor Sp4 are associated with decreased cancer risk in patients with dermatomyositis with anti-TIF1γ autoantibodies. *Ann Rheumatic Dis*. (2023) 82(2):246–52. doi: 10.1136/ard-2022-222441

53. Fiorentino DF, Mecoli CA, Rosen MC, Chung LS, Christopher-Stine L, Rosen A, et al. Immune responses to CCAR1 and other dermatomyositis autoantigens are associated with attenuated cancer emergence. *J Clin Invest*. (2022) 132:e150201. doi: 10.1172/JCI150201

54. Takahashi K, Yoshida N, Murakami N, Kawata K, Ishizaki H, Tanaka-Okamoto M, et al. Dynamic regulation of p53 subnuclear localization and senescence by MORC3. *Mol Biol Cell*. (2007) 18:1701–9. doi: 10.1091/mbc.e06-08-0747

55. Fu W, Chang X, Ye K, Zheng Z, Lai Q, Ge M, et al. Genome-wide analysis reveals the MORC3-mediated repression of PD-L1 expression in head and neck cancer. *Front Cell Dev Biol*. (2024) 12:1410130. doi: 10.3389/fcell.2024.1410130

56. Ichimura Y, Matsushita T, Hamaguchi Y, Kaji K, Hasegawa M, Tanino Y, et al. Anti-NXP2 autoantibodies in adult patients with idiopathic inflammatory myopathies: possible association with Malignancy. *Ann Rheum Dis*. (2012) 71:710–3. doi: 10.1136/annrheumdis-2011-200697

57. Ceribelli A, Fredi M, Taraborelli M, Cavazzana I, Franceschini F, Quinzanini M, et al. Anti-MJ/NXP-2 autoantibody specificity in a cohort of adult Italian patients with polymyositis/dermatomyositis. *Arthritis Res Ther*. (2012) 14:R97. doi: 10.1186/ar3822

58. Ichimura Y, Konishi R, Shobo M, Inoue S, Okune M, Maeda A, et al. Anti-nuclear matrix protein 2 antibody-positive inflammatory myopathies represent extensive myositis without dermatomyositis-specific rash. *Rheumatol (Oxford)*. (2022) 61:1222–7. doi: 10.1093/rheumatology/keab518

59. Fredi M, Cavazzana I, Ceribelli A, Cavagna L, Barsotti S, Bartoloni E, et al. An Italian multicenter study on anti-NXP2 antibodies: clinical and serological associations. *Clin Rev Allergy Immunol*. (2022) 63:240–50. doi: 10.1007/s12016-021-08920-y

60. Zhong L, Yu Z, Song H. Association of anti-nuclear matrix protein 2 antibody with complications in patients with idiopathic inflammatory myopathies: A meta-analysis of 20 cohorts. *Clin Immunol*. (2019) 198:11–8. doi: 10.1016/j.clim.2018.11.008

61. Ceribelli A, Tonutti A, Isailovic N, De Santis M, Selmi C. Interstitial lung disease associated with inflammatory myositis: Autoantibodies, clinical phenotypes, and progressive fibrosis. *Front Med (Lausanne)*. (2023) 10:1068402. doi: 10.3389/fmed.2023.1068402

62. Demortier J, Vautier M, Chosidow O, Gallay L, Bessis D, Berezne A, et al. Anti-SAE autoantibody in dermatomyositis: original comparative study and review of the literature. *Rheumatol (Oxford)*. (2023) 62:3932–9. doi: 10.1093/rheumatology/kead154

63. Zhang Y, Liu L, Duan X, Pi H, Jiang L, Li J, et al. Longitudinal study of patients with anti-SAE antibody positive dermatomyositis: a multicenter cohort study in China. *Rheumatol (Oxford)*. (2025) 64(3):1377–85. doi: 10.1093/rheumatology/keae232

64. Zhao Y, Su H, Yin X, Hou H, Wang Y, Xu Y, et al. Cancer associated autoantibodies in idiopathic inflammatory myopathies: A retrospective cohort from a single center in China. *Med Clin (Barc)*. (2023) 160:10–6. doi: 10.1016/j.medcli.2022.03.023

65. Albayda J, Mecoli C, Casciola-Rosen L, Danoff SK, Lin CT, Hines D, et al. A North American cohort of anti-SAE dermatomyositis: clinical phenotype, testing, and review of cases. *ACR Open Rheumatol*. (2021) 3:287–94. doi: 10.1002/acr2.11247

66. Tharuka MDN, Courelli AS, Chen Y. Immune regulation by the SUMO family. *Nat Rev Immunol*. (2025). doi: 10.1038/s41577-025-01155-4

67. Shi L, Shangguan J, Lu Y, Rong J, Yang Q, Yang Y, et al. ROS-mediated up-regulation of SAE1 by Helicobacter pylori promotes human gastric tumor genesis and progression. *J Transl Med*. (2024) 22:148. doi: 10.1186/s12967-024-04913-5

68. Liu H, Wang J, Li Y, Luo F, Xing L. Upregulated SAE1 drives tumorigenesis and is associated with poor clinical outcomes in breast cancer. *Breast J*. (2024) 2024:2981722. doi: 10.1155/2024/2981722

69. Ong JR, Bamodu OA, Khang NV, Lin YK, Yeh CT, Lee WH, et al. SUMO-activating enzyme subunit 1 (SAE1) is a promising diagnostic cancer metabolism biomarker of hepatocellular carcinoma. *Cells*. (2021) 10:178. doi: 10.3390/cells10010178

70. Allenbach Y, Keraen J, Bouvier AM, Jooste V, Champtiaux N, Hervier B, et al. High risk of cancer in autoimmune necrotizing myopathies: usefulness of myositis specific antibody. *Brain*. (2016) 139:2131–5. doi: 10.1093/brain/aww054

71. Allenbach Y, Benveniste O, Stenzel W, Boyer O. Immune-mediated necrotizing myopathy: clinical features and pathogenesis. *Nat Rev Rheumatol*. (2020) 16:689–701. doi: 10.1038/s41584-020-00515-9

72. Levin MI, Mozaffar T, Al-Lozi MT, Pestronk A. Paraneoplastic necrotizing myopathy: clinical and pathological features. *Neurology*. (1998) 50:764–7. doi: 10.1212/WNL.50.3.764

73. Vu HJ, Pham D, Makary R, Nguyen T, Shuja S. Paraneoplastic necrotizing myopathy presenting as severe muscle weakness in a patient with small-cell lung cancer: successful response to chemoradiation therapy. *Clin Adv Hematol Oncol*. (2011) 9:557–6.

74. Belizna C, Henrion D, Beucher A, Lavigne C, Ghaali A, Lévesque H. Anti-Ku antibodies: Clinical, genetic and diagnostic insights. *Autoimmun Rev*. (2010) 9:691–4. doi: 10.1016/j.autrev.2010.05.020

75. Galindo-Feria AS, Horuluoglu B, Day J, Fernandes-Cerqueira C, Wigren E, Gräslund S, et al. Autoantibodies against four-and-a-half-LIM domain 1 (FHL1) in inflammatory myopathies: results from an Australian single-centre cohort. *Rheumatol (Oxford)*. (2022) 61:4145–54. doi: 10.1093/rheumatology/keac003

76. Vulsteke JB, Piette Y, Bonroy C, Verschueren P, Blockmans D, Vanderschueren S, et al. Anti-RuvBL1/2 autoantibodies in patients with systemic sclerosis or idiopathic inflammatory myopathy and a nuclear speckled pattern. *Ann Rheum Dis*. (2022) 81:742–4. doi: 10.1136/annrheumdis-2021-220004

77. Nomura Y, Ueda-Hayakawa I, Yamazaki F, Ozaki Y, Hamaguchi Y, Takehara K, et al. A case of anti-RuvBL1/2 antibody-positive systemic sclerosis overlapping with myositis. *Eur J Dermatol*. (2020) 30:52–3. doi: 10.1684/ejd.2019.3684

78. Di Pietro L, Chiccoli F, Salvati L, Vivarelli E, Vultaggio A, Matucci A, et al. Anti-RuvBL1/2 autoantibodies detection in a patient with overlap systemic sclerosis and polymyositis. *Antibodies (Basel)*. (2023) 12:13. doi: 10.3390/antib12010013

79. Senécal JL, Isabelle C, Fritzler MJ, Targoff IN, Goldstein R, Gagné M, et al. An autoimmune myositis-overlap syndrome associated with autoantibodies to nuclear pore complexes: description and long-term follow-up of the anti-Nup syndrome. *Med (Baltimore)*. (2014) 93:383–94. doi: 10.1097/MD.0000000000000223

80. Satoh M, Chan JYF, Ross SJ, Ceribelli A, Cavazzana I, Franceschini F, et al. Autoantibodies to survival of motor neuron complex in patients with polymyositis: immunoprecipitation of D, E, F, and G proteins without other components of small nuclear ribonucleoproteins. *Arthritis Rheumatol*. (2011) 63:1972–8. doi: 10.1002/art.30349

81. El Kamouni H, Jalaledin D S, Albert A, Hoa S, Vo C, Bourré-Tessier J, et al. Anti-SMN autoantibodies in mixed connective tissue disease are associated with a severe systemic sclerosis phenotype. *RMD Open*. (2023) 9:e003431. doi: 10.1136/rmdopen-2023-003431

82. Ceribelli A, Tonutti A, Isailovic N, De Santis M, Selmi C. Established and novel insights to guide cancer assessment in patients with idiopathic inflammatory myopathies. *Semin Arthritis Rheumatism*. (2025) 71:152619. doi: 10.1016/j.semarthrit.2024.152619

83. Tang IYK, Chan SCW, Li PH, Li TWL, Luk LTH, Chan D, et al. Validation of the International Myositis Assessment and Clinical Studies Group guideline on cancer risk stratification. *Rheumatol (Oxford)*. (2025) 64(4):2106–14. doi: 10.1093/rheumatology/keae504

84. Zhong H, Liu S, Wang Y, Xu D, Li M, Zhao Y, et al. Primary Sjögren's syndrome is associated with increased risk of Malignancies besides lymphoma: A systematic review and meta-analysis. *Autoimmun Rev*. (2022) 21:103084. doi: 10.1016/j.autrev.2022.103084

85. Nocturne G, Pontarini E, Bombardieri M, Mariette X. Lymphomas complicating primary Sjögren's syndrome: from autoimmunity to lymphoma. *Rheumatol (Oxford)*. (2021) 60:3513–21. doi: 10.1093/rheumatology/kez052

86. Thandra KC, Barsouk A, Saginala K, Padala SA, Barsouk A, Rawla P. Epidemiology of non-Hodgkin's lymphoma. *Med Sci (Basel)*. (2021) 9:5. doi: 10.3390/medsci9010005

87. Goules AV, Chatzis L, Pezoulas VC, Patsouras M, Mavragani C, Quartuccio L, et al. Identification and evolution of predictors of Sjögren's disease-associated mucosa-associated lymphoid tissue lymphoma development over time: a case-control study. *Lancet Rheumatol*. (2024) 6:e693–702. doi: 10.1016/S2665-9913(24)00183-8

88. Quartuccio L, Isola M, Baldini C, Priori R, Bartoloni Bocci E, Carubbi F, et al. Biomarkers of lymphoma in Sjögren's syndrome and evaluation of the lymphoma risk in prelymphomatous conditions: results of a multicenter study. *J Autoimmun*. (2014) 51:75–80. doi: 10.1016/j.jaut.2013.10.002

89. Chatzis L, Goules AV, Pezoulas V, Baldini C, Gandolfo S, Skopouli FN, et al. A biomarker for lymphoma development in Sjogren's syndrome: Salivary gland focus score. *J Autoimmun*. (2021) 121:102648. doi: 10.1016/j.jaut.2021.102648

90. Mofors J, Björk A, Smedby KE, Kvarnström M, Forsblad-d'Elia H, Magnusson-Bucher S, et al. Increased risk of multiple myeloma in primary Sjögren's syndrome is limited to individuals with Ro/SSA and La/SSB autoantibodies. *Ann Rheum Dis*. (2020) 79:307–8. doi: 10.1136/annrheumdis-2019-216287

91. Weng MY, Huang YT, Liu MF, Lu TH. Incidence of cancer in a nationwide population cohort of 7852 patients with primary Sjogren's syndrome in Taiwan. *Ann Rheum Dis*. (2012) 71:524–7. doi: 10.1136/annrheumdis-2011-200402

92. Brito-Zerón P, Kostov B, Fraile G, Caravia-Durán D, Maure B, Rascón FJ, et al. Characterization and risk estimate of cancer in patients with primary Sjögren syndrome. *J Hematol Oncol*. (2017) 10:90. doi: 10.1186/s13045-017-0464-5

93. Hu X, Wang X, Liang Y, Chen X, Zhou S, Fei W, et al. Cancer risk in Hashimoto's thyroiditis: a systematic review and meta-analysis. *Front Endocrinol (Lausanne)*. (2022) 13:937871. doi: 10.3389/fendo.2022.937871

94. ColaFrancesco S, Celia AI, Baldini C, Quartuccio L, Bartoloni E, Carubbi F, et al. Clinical and histological features of patients with primary Sjögren's syndrome and autoimmune thyroiditis: a national multicentre cross-sectional study. *Clin Exp Rheumatol*. (2023) 41:2389–96. doi: 10.55563/clinexprheumatol/eh36vs

95. Bombardieri M, Lewis M, Pitzalis C. Ectopic lymphoid neogenesis in rheumatic autoimmune diseases. *Nat Rev Rheumatol*. (2017) 13:141–54. doi: 10.1038/nrrheum.2016.217

96. Stott DI, Hiepe F, Hummel M, Steinhauser G, Berek C. Antigen-driven clonal proliferation of B cells within the target tissue of an autoimmune disease. The salivary glands of patients with Sjögren's syndrome. *J Clin Invest*. (1998) 102:938–46. doi: 10.1172/JCI3234

97. Corsiero E, Nerviani A, Bombardieri M, Pitzalis C. Ectopic lymphoid structures: powerhouse of autoimmunity. *Front Immunol*. (2016) 7:430. doi: 10.3389/fimmu.2016.00430

98. Theander E, Vasaitis L, Baecklund E, Nordmark G, Warfvinge G, Liedholm R, et al. Lymphoid organisation in labial salivary gland biopsies is a possible predictor for the development of Malignant lymphoma in primary Sjögren's syndrome. *Ann Rheum Dis*. (2011) 70:1363–8. doi: 10.1136/ard.2010.144782

99. Sène D, Ismael S, Forien M, Charlotte F, Kaci R, Cacoub P, et al. Ectopic germinal center-like structures in minor salivary gland biopsy tissue predict lymphoma occurrence in patients with primary Sjögren's syndrome. *Arthritis Rheumatol*. (2018) 70:1481–8. doi: 10.1002/art.40528

100. Haacke EA, van der Vegt B, Vissink A, Spijkervet FKL, Bootsma H, Kroese FGM. Germinal centres in diagnostic labial gland biopsies of patients with primary Sjögren's syndrome are not predictive for parotid MALT lymphoma development. *Ann Rheum Dis*. (2017) 76:1781–4. doi: 10.1136/annrheumdis-2017-211290

101. Haacke EA, van der Vegt B, Vissink A, Spijkervet FKL, Bootsma H, Kroese FGM. Germinal Centers in Diagnostic Biopsies of Patients With Primary Sjögren's Syndrome Are Not a Risk Factor for Non-Hodgkin's Lymphoma but a Reflection of High Disease Activity: Comment on the Article by Sène et al. *Arthritis Rheumatol*. (2019) 71:170–1. doi: 10.1002/art.2019.71.issue-1

102. ColaFrancesco S, Priori R, Smith CG, Minniti A, Iannizzotto V, Pipi E, et al. CXCL13 as biomarker for histological involvement in Sjögren's syndrome. *Rheumatol (Oxford)*. (2020) 59:165–70. doi: 10.1093/rheumatology/kez255

103. Barone F, Bombardieri M, Rosado MM, Morgan PR, Challacombe SJ, De Vita S, et al. CXCL13, CCL21, and CXCL12 expression in salivary glands of patients with Sjogren's syndrome and MALT lymphoma: association with reactive and Malignant areas of lymphoid organization. *J Immunol*. (2008) 180:5130–40. doi: 10.4049/jimmunol.180.7.5130

104. Badarinza M, Serban O, Maghear L, Pelea MA, Rosca RI, Fodor D, et al. Diagnostic role of CXCL13 biomarker in primary Sjogren's syndrome patients with parotid non-Hodgkin's lymphoma complication. *Med Clin (Barc)*. (2023) 160:484–8. doi: 10.1016/j.medcli.2023.01.023

105. Hüper S, Nagler L, Strunz PP, Froehlich M, Labinsky H, Schmalzing M, et al. Lymphoma in Sjögren's syndrome: no need for repetitive screening ultrasounds of the major salivary glands and neck in asymptomatic patients. *Scand J Rheumatol*. (2025) 54:49–57. doi: 10.1080/03009742.2024.2370109

106. Giovannini I, De Martino M, Manfrè V, Lorenzon M, Cereser L, Di Loreto C, et al. Ultrasonographic scores and parotid histopathology in Sjögren's disease: challenges in lymphoma identification. *Clin Exp Rheumatol*. (2024) 42:2483–9. doi: 10.55563/clinexprheumatol/98vjav

107. Abu-Shakra M, Guillemin F, Lee P. Cancer in systemic sclerosis. *Arthritis Rheumatol*. (1993) 36:460–4. doi: 10.1002/art.1780360405

108. Rosenthal AK, McLaughlin JK, Gridley G, Nyrén O. Incidence of cancer among patients with systemic sclerosis. *Cancer*. (1995) 76:910–4. doi: 10.1002/1097-0142(19950901)76:5<910::AID-CNCR2820760528>3.0.CO;2-T

109. Derk CT, Rasheed M, Artlett CM, Jimenez SA. A cohort study of cancer incidence in systemic sclerosis. *J Rheumatol*. (2006) 33:1113–6.

110. Watad A, McGonagle D, Bragazzi NL, Tiosano S, Comaneshter D, Shoenfeld Y, et al. Autoantibody status in systemic sclerosis patients defines both cancer risk and survival with ANA negativity in cases with concomitant cancer having a worse survival. *Oncoimmunology*. (2019) 8:e1588084. doi: 10.1080/2162402X.2019.1588084

111. Morrisroe K, Hansen D, Huq M, Stevens W, Sahhar J, Ngian GS, et al. Incidence, risk factors, and outcomes of cancer in systemic sclerosis. *Arthritis Care Res (Hoboken)*. (2020) 72:1625–35. doi: 10.1002/acr.v72.11

112. Onishi A, Sugiyama D, Kumagai S, Morinobu A. Cancer incidence in systemic sclerosis: meta-analysis of population-based cohort studies. *Arthritis Rheumatol*. (2013) 65:1913–21. doi: 10.1002/art.37969

113. Tyndall AJ, Bannert B, Vonk M, Airò P, Cozzi F, Carreira PE, et al. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. *Ann Rheum Dis*. (2010) 69:1809–15. doi: 10.1136/ard.2009.114264

114. Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972-2002. *Ann Rheum Dis*. (2007) 66:940–4. doi: 10.1136/ard.2006.066068

115. Kortelainen S, Käyrä M, Rissanen T, Paltta J, Taimen K, Pirilä L, et al. Causes and predictors of death among Finnish patients with systemic sclerosis. *Scand J Rheumatol*. (2024) 53(4):269–75. doi: 10.1080/03009742.2024.2335781

116. Rubio-Rivas M, Corbella X, Guillén-del-Castillo A, Tolosa Vilella C, Colunga Argüelles D, Argibay A, et al. Spanish scleroderma risk score (RESCLESCORE) to predict 15-year all-cause mortality in scleroderma patients at the time of diagnosis based on the RESCLE cohort: Derivation and internal validation. *Autoimmun Rev*. (2020) 19:102507. doi: 10.1016/j.autrev.2020.102507

117. Czirják L, Kumánovics G, Varjú C, Nagy Z, Pákozdi A, Szekanecz Z, et al. Survival and causes of death in 366 Hungarian patients with systemic sclerosis. *Ann Rheum Dis*. (2008) 67:59–63. doi: 10.1136/ard.2006.066340

118. Zhang JQ, Wan YN, Peng WJ, Yan JW, Li BZ, Mei B, et al. The risk of cancer development in systemic sclerosis: a meta-analysis. *Cancer Epidemiol*. (2013) 37:523–7. doi: 10.1016/j.canep.2013.04.014

119. Guillen-Del-Castillo A, Simeón-Aznar CP. Identifying the risk of cancer-associated systemic sclerosis. *Joint Bone Spine*. (2023) 90:105618. doi: 10.1016/j.jbspin.2023.105618

120. Carbonell C, Marcos M, Guillén-Del-Castillo A, Rubio-Rivas M, Argibay A, Marín-Ballvé A, et al. Standardized incidence ratios and risk factors for cancer in patients with systemic sclerosis: Data from the Spanish Scleroderma Registry (RESCLE). *Autoimmun Rev*. (2022) 21:103167. doi: 10.1016/j.autrev.2022.103167

121. Antonelli A, Ferri C, Ferrari SM, Di Domenicantonio A, Giuggioli D, Galleri D, et al. Increased risk of papillary thyroid cancer in systemic sclerosis associated with autoimmune thyroiditis. *Rheumatol (Oxford)*. (2016) 55:480–4. doi: 10.1093/rheumatology/kev358

122. Sakr L, Hudson M, Wang M, Younanian E, Baron M, Bernatsky S. Interstitial lung disease is associated with an increased risk of lung cancer in systemic sclerosis: Longitudinal data from the Canadian Scleroderma Research Group. *J Scleroderma Relat Disord*. (2018) 3:221–7. doi: 10.1177/2397198318766825

123. Colaci M, Giuggioli D, Sebastiani M, Manfredi A, Vacchi C, Spagnolo P, et al. Lung cancer in scleroderma: results from an Italian rheumatologic center and review of the literature. *Autoimmun Rev*. (2013) 12:374–9. doi: 10.1016/j.autrev.2012.06.003

124. Lopez L, Barnetche T, Galli G, Seneschal J, Blanchard E, Shipley E, et al. Clinical and immunological features of patients with cancer-associated systemic sclerosis: An observational study. *Joint Bone Spine*. (2023) 90:105555. doi: 10.1016/j.jbspin.2023.105555

125. Partouche L, Goulabchand R, Maria ATJ, Rivière S, Jorgensen C, Rigau V, et al. Biphasic temporal relationship between cancers and systemic sclerosis: A clinical series from Montpellier university hospital and review of the literature. *J Clin Med*. (2020) 9:853. doi: 10.3390/jcm9030853

126. Shah AA, Rosen A. Cancer and systemic sclerosis: novel insights into pathogenesis and clinical implications. *Curr Opin Rheumatol*. (2011) 23:530–5. doi: 10.1097/BOR.0b013e32834a5081

127. Fiorentino DF, Casciola-Rosen L. Autoantibodies and cancer association: the case of systemic sclerosis and dermatomyositis. *Clinic Rev Allerg Immunol*. (2022) 63:330–41. doi: 10.1007/s12016-022-08944-y

128. Calderon LM, Domsic RT, Shah AA, Pope JE. Preventative care in scleroderma: what is the best approach to bone health and cancer screening? *Rheum Dis Clin North Am*. (2023) 49:411–23. doi: 10.1016/j.rdc.2023.01.011

129. Shah AA, Casciola-Rosen L. Cancer and scleroderma: a paraneoplastic disease with implications for Malignancy screening. *Curr Opin Rheumatol*. (2015) 27:563–70. doi: 10.1097/BOR.0000000000000222

130. Kim JS, Woods A, Gutierrez-Alamillo L, Laffoon M, Wigley FM, Hummers LK, et al. Distinct scleroderma autoantibody profiles stratify patients for cancer risk at scleroderma onset and during the disease course. *Arthritis Rheumatol*. (2023) 76(1):68–77. doi: 10.1002/art.42663

131. Igusa T, Hummers LK, Visvanathan K, Richardson C, Wigley FM, Casciola-Rosen L, et al. Autoantibodies and scleroderma phenotype define subgroups at high-risk and low-risk for cancer. *Ann Rheum Dis*. (2018) 77:1179–86. doi: 10.1136/annrheumdis-2018-212999

132. Airo' P, Ceribelli A, Cavazzana I, Taraborelli M, Zingarelli S, Franceschini F. Malignancies in Italian patients with systemic sclerosis positive for anti-RNA polymerase III antibodies. *J Rheumatol*. (2011) 38:1329–34. doi: 10.3899/jrheum.101144

133. Lazzaroni MG, Cavazzana I, Colombo E, Dobrota R, Hernandez J, Hesselstrand R, et al. Malignancies in patients with anti-RNA polymerase III antibodies and systemic sclerosis: analysis of the EULAR scleroderma trials and research cohort and possible recommendations for screening. *J Rheumatol*. (2017) 44:639–47. doi: 10.3899/jrheum.160817

134. Callejas-Moraga EL, Guillén-Del-Castillo A, Marín-Sánchez AM, Roca-Herrera M, Balada E, Tolosa-Vilella C, et al. Clinical features of systemic sclerosis patients with anti-RNA polymerase III antibody in a single centre in Spain. *Clin Exp Rheumatol*. (2019) 37 Suppl 119(4):41–8.

135. Nikpour M, Hissaria P, Byron J, Sahhar J, Micallef M, Paspaliaris W, et al. Prevalence, correlates and clinical usefulness of antibodies to RNA polymerase III in systemic sclerosis: a cross-sectional analysis of data from an Australian cohort. *Arthritis Res Ther*. (2011) 13:R211. doi: 10.1186/ar3544

136. Moinzadeh P, Fonseca C, Hellmich M, Shah AA, Chighizola C, Denton CP, et al. Association of anti-RNA polymerase III autoantibodies and cancer in scleroderma. *Arthritis Res Ther*. (2014) 16:R53. doi: 10.1186/ar4486

137. De Santis M, Tonutti A, Motta F, Rodolfi S, Isailovic N, Selmi C. Serum autoantibodies and the risk of cancer in systemic sclerosis over time. *Arthritis Rheumatol*. (2024) 76(2):314–5. doi: 10.1002/art.42701

138. Tonutti A, Motta F, Isailovic N, Ceribelli A, Ragusa R, Nappi E, et al. Autoantibodies, cutaneous subset and immunosuppressants contribute to the cancer risk in systemic sclerosis. *RMD Open*. (2024) 10:e004492. doi: 10.1136/rmdopen-2024-004492

139. Schneeberger D, Tyndall A, Kay J, Søndergaard KH, Carreira PE, Morgiel E, et al. Systemic sclerosis without antinuclear antibodies or Raynaud's phenomenon: a multicentre study in the prospective EULAR Scleroderma Trials and Research (EUSTAR) database. *Rheumatol (Oxford)*. (2013) 52:560–7. doi: 10.1093/rheumatology/kes315

140. Mecoli CA, Adler BL, Yang Q, Hummers LK, Rosen A, Casciola-Rosen L, et al. Cancer in systemic sclerosis: analysis of antibodies against components of the Th/To complex. *Arthritis Rheumatol*. (2021) 73:315–23. doi: 10.1002/art.41493

141. Liaskos C, Marou E, Simopoulou T, Barmakoudi M, Efthymiou G, Scheper T, et al. Disease-related autoantibody profile in patients with systemic sclerosis. *Autoimmunity*. (2017) 50:414–21. doi: 10.1080/08916934.2017.1357699

142. Bernal-Bello D, de Tena JG, Guillén-Del Castillo A, Selva-O'Callaghan A, Callejas-Moraga EL, Marín-Sánchez AM, et al. Novel risk factors related to cancer in scleroderma. *Autoimmun Rev*. (2017) 16:461–8. doi: 10.1016/j.autrev.2017.03.012

143. Shah AA, Xu G, Rosen A, Hummers LK, Wigley FM, Elledge SJ, et al. Brief report: anti-RNPC-3 antibodies as a marker of cancer-associated scleroderma. *Arthritis Rheumatol*. (2017) 69:1306–12. doi: 10.1002/art.40065

144. Callejas-Moraga EL, Guillén-Del-Castillo A, Perurena-Prieto J, Sanz-Martinez MT, Fonollosa-Pla V, Lorite-Gomez K, et al. Anti-RNPC-3 antibody predicts poor prognosis in patients with interstitial lung disease associated to systemic sclerosis. *Rheumatol (Oxford)*. (2021) 61:154–62. doi: 10.1093/rheumatology/keab279

145. Isca C, Spinella A, Toss A, de Pinto M, Ficarra G, Fabbiani L, et al. Molecular pathways of breast cancer in systemic sclerosis: exploratory immunohistochemical analysis from the Sclero-breast study. *J Pers Med*. (2022) 12:2007. doi: 10.3390/jpm12122007

146. Hoa S, Lazizi S, Baron M, Wang M, Fritzler MJ, Hudson M, et al. Association between autoantibodies in systemic sclerosis and cancer in a national registry. *Rheumatol (Oxford)*. (2022) 61:2905–14. doi: 10.1093/rheumatology/keab735

147. Wallwork RS, Shah AA, Casciola-Rosen L. Association between anti-SSSCA1 antibodies and cancer in systemic sclerosis. *Rheumatol (Oxford)*. (2023) 62:2539–43. doi: 10.1093/rheumatology/keac614

148. Burja B, Boubaya M, Bruni C, Carreira P, Bergmann C, Ananyeva LP, et al. OP0013 anti-Ro/Ssa antibodies are predictive of A more severe lung involvement in patients with systemic sclerosis: A study from the Eustar Hdatabase. *Ann Rheumatic Diseases*. (2024) 83:1–2. doi: 10.1136/annrheumdis-2024-eular.1670

149. Watanabe T, Ototake Y, Akita A, Suzuki M, Kanaoka M, Tamura J, et al. Clinical features of patients with systemic sclerosis positive for anti-SS-A antibody: a cohort study of 156 patients. *Arthritis Res Ther*. (2024) 26:93. doi: 10.1186/s13075-024-03325-6

150. De Luca G, Bosello SL, Berardi G, Rucco M, Canestrari G, Correra M, et al. Tumour-associated antigens in systemic sclerosis patients with interstitial lung disease: association with lung involvement and cancer risk. *Rheumatol (Oxford)*. (2015) 54:1991–9. doi: 10.1093/rheumatology/kev204

151. Munagala M, Phancao A. Malignancy: an adverse effect of immunosuppression. *Handb Exp Pharmacol*. (2022) 272:315–35. doi: 10.1007/164_2021_554

152. Travis LB, Curtis RE, Glimelius B, Holowaty EJ, Van Leeuwen FE, Lynch CF, et al. Bladder and kidney cancer following cyclophosphamide therapy for non-Hodgkin's lymphoma. *J Natl Cancer Inst*. (1995) 87:524–30. doi: 10.1093/jnci/87.7.524

153. Chou WH, McGregor B, Schmidt A, Carvalho FLF, Hirsch MS, Chang SL, et al. Cyclophosphamide-associated bladder cancers and considerations for survivorship care: A systematic review. *Urol Oncol*. (2021) 39:678–85. doi: 10.1016/j.urolonc.2021.05.017

154. Tashkin DP, Roth MD, Clements PJ, Furst DE, Khanna D, Kleerup EC, et al. Mycophenolate mofetil versus oral cyclophosphamide in scleroderma-related interstitial lung disease (SLS II): a randomised controlled, double-blind, parallel group trial. *Lancet Respir Med*. (2016) 4:708–19. doi: 10.1016/S2213-2600(16)30152-7

155. Bhat R, Tonutti A, Timilsina S, Selmi C, Gershwin ME. Perspectives on mycophenolate mofetil in the management of autoimmunity. *Clin Rev Allergy Immunol*. (2023) 65:86–100. doi: 10.1007/s12016-023-08963-3

156. Buell JF, Gross TG, Woodle ES. Malignancy after transplantation. *Transplantation*. (2005) 80:S254–264. doi: 10.1097/01.tp.0000186382.81130.ba

157. Hirunsatitpron P, Hanprasertpong N, Noppakun K, Pruksakorn D, Teekachunhatean S, Koonrungsesomboon N. Mycophenolic acid and cancer risk in solid organ transplant recipients: Systematic review and meta-analysis. *Br J Clin Pharmacol*. (2022) 88:476–89. doi: 10.1111/bcp.14979

158. Gibson JAG, Cordaro A, Dobbs TD, Griffiths R, Akbari A, Whitaker S, et al. The association between immunosuppression and skin cancer in solid organ transplant recipients: a control-matched cohort study of 2,852 patients. *Eur J Dermatol*. (2021) 31:712–21. doi: 10.1684/ejd.2021.4108

159. Lok SD, Wong AW, Khor YH, Ryerson CJ, Johannson KA, CARE-PF Investigators. Malignancy risk associated with mycophenolate mofetil or azathioprine in patients with fibrotic interstitial lung disease. *Chest*. (2022) 161:1594–7. doi: 10.1016/j.chest.2021.12.636

160. De Santis M, Motta F, Isailovic N, Clementi M, Criscuolo E, Clementi N, et al. Dose-dependent impairment of the immune response to the moderna-1273 mRNA vaccine by mycophenolate mofetil in patients with rheumatic and autoimmune liver diseases. *Vaccines (Basel)*. (2022) 10:801. doi: 10.3390/vaccines10050801

161. Moradzadeh M, Aghaei M, Mehrbakhsh Z, Arab-Bafrani Z, Abdollahi N. Efficacy and safety of rituximab therapy in patients with systemic sclerosis disease (SSc): systematic review and meta-analysis. *Clin Rheumatol*. (2021) 40:3897–918. doi: 10.1007/s10067-021-05698-4

162. Lee YH, Song GG. Mortality in patients with systemic lupus erythematosus: A meta-analysis of overall and cause-specific effects. *Lupus*. (2024) 33:929–37. doi: 10.1177/09612033241257134

163. Clarke AE, Pooley N, Marjenberg Z, Langham J, Nicholson L, Langham S, et al. Risk of Malignancy in patients with systemic lupus erythematosus: Systematic review and meta-analysis. *Semin Arthritis Rheumatol*. (2021) 51:1230–41. doi: 10.1016/j.semarthrit.2021.09.009

164. Hardenbergh D, Molina E, Naik R, Geetha D, Chaturvedi S, Timlin H. Factors mediating cancer risk in systemic lupus erythematosus. *Lupus*. (2022) 31:1285–95. doi: 10.1177/09612033221122163

165. Sirong C, Zhang K, Yang Y, Zhong Y, Sun P, Zanhong W. Association between systemic lupus erythematosus and common female reproductive system Malignancies. *Discovery Oncol*. (2024) 15:341. doi: 10.1007/s12672-024-01218-3

166. Bernatsky S, Ramsey-Goldman R, Petri M, Urowitz MB, Gladman DD, Fortin PR, et al. Smoking is the most significant modifiable lung cancer risk factor in systemic lupus erythematosus. *J Rheumatol*. (2018) 45:393–6. doi: 10.3899/jrheum.170652

167. Zhao Q, Liu H, Yang W, Zhou Z, Yang Y, Jiang X, et al. Cancer occurrence after SLE: effects of medication-related factors, disease-related factors and survival from an observational study. *Rheumatol (Oxford)*. (2023) 62:659–67. doi: 10.1093/rheumatology/keac316

168. Bernatsky S, Ramsey-Goldman R, Joseph L, Boivin JF, Costenbader KH, Urowitz MB, et al. Lymphoma risk in systemic lupus: effects of disease activity versus treatment. *Ann Rheum Dis*. (2014) 73:138–42. doi: 10.1136/annrheumdis-2012-202099

169. Draborg AH, Duus K, Houen G. Epstein-Barr virus and systemic lupus erythematosus. *Clin Dev Immunol*. (2012) 2012:370516. doi: 10.1155/2012/370516

170. Gayed M, Bernatsky S, Ramsey-Goldman R, Clarke A, Gordon C. Lupus and cancer. *Lupus*. (2009) 18:479–85. doi: 10.1177/0961203309102556

171. Pan XF, Gu JQ, Shan ZY. Patients with systemic lupus erythematosus have higher prevalence of thyroid autoantibodies: a systematic review and meta-analysis. *PloS One*. (2015) 10:e0123291. doi: 10.1371/journal.pone.0123291

172. Brufatto MZ, Lanças SHS, de Albuquerque Pedrosa Fernandes T, Sallum AME, Campos LMA, Sakamoto AP, et al. Childhood-onset systemic lupus erythematosus (cSLE) and Malignancy: a nationwide multicentre series review. *Adv Rheumatol*. (2024) 64:13. doi: 10.1186/s42358-024-00353-3

173. Zhu QY. Bioinformatics analysis of the pathogenic link between Epstein-Barr virus infection, systemic lupus erythematosus and diffuse large B cell lymphoma. *Sci Rep*. (2023) 13:6310. doi: 10.1038/s41598-023-33585-2

174. Tam LS, Chan PKS, Ho SC, Yu MMY, Yim SF, Cheung TH, et al. Natural history of cervical papilloma virus infection in systemic lupus erythematosus - a prospective cohort study. *J Rheumatol*. (2010) 37:330–40. doi: 10.3899/jrheum.090644

175. Kristinsson SY, Björkholm M, Hultcrantz M, Derolf ÅR, Landgren O, Goldin LR. Chronic immune stimulation might act as a trigger for the development of acute myeloid leukemia or myelodysplastic syndromes. *J Clin Oncol*. (2011) 29:2897–903. doi: 10.1200/JCO.2011.34.8540

176. Ladouceur A, Tessier-Cloutier B, Clarke AE, Ramsey-Goldman R, Gordon C, Hansen JE, et al. Cancer and systemic lupus erythematosus. *Rheum Dis Clin North Am*. (2020) 46:533–50. doi: 10.1016/j.rdc.2020.05.005

177. Wang LH, Wang WM, Lin SH, Shieh CC. Bidirectional relationship between systemic lupus erythematosus and non-Hodgkin's lymphoma: a nationwide population-based study. *Rheumatol (Oxford)*. (2019) 58:1245–9. doi: 10.1093/rheumatology/kez011

178. Tessier-Cloutier B, Twa DD, Baecklund E, Gascoyne R, Johnson NA, Backlin C, et al. Cell of origin in diffuse large B-cell lymphoma in systemic lupus erythematosus: molecular and clinical factors associated with survival. *Lupus Sci Med*. (2019) 6:e000324. doi: 10.1136/lupus-2019-000324

179. Noble PW, Bernatsky S, Clarke AE, Isenberg DA, Ramsey-Goldman R, Hansen JE. DNA-damaging autoantibodies and cancer: the lupus butterfly theory. *Nat Rev Rheumatol*. (2016) 12:429–34. doi: 10.1038/nrrheum.2016.23

180. Hansen JE, Chan G, Liu Y, Hegan DC, Dalal S, Dray E, et al. Targeting cancer with a lupus autoantibody. *Sci Transl Med*. (2012) 4:157ra142. doi: 10.1126/scitranslmed.3004385

181. Bernatsky S, Velásquez García HA, Spinelli JJ, Gaffney P, Smedby KE, Ramsey-Goldman R, et al. Lupus-related single nucleotide polymorphisms and risk of diffuse large B-cell lymphoma. *Lupus Sci Med*. (2017) 4:e000187. doi: 10.1136/lupus-2016-000187

182. Rosenberger A, Sohns M, Friedrichs S, Hung RJ, Fehringer G, McLaughlin J, et al. Gene-set meta-analysis of lung cancer identifies pathway related to systemic lupus erythematosus. *PloS One*. (2017) 12:e0173339. doi: 10.1371/journal.pone.0173339

183. Di Marco M, Ramassone A, Pagotto S, Anastasiadou E, Veronese A, Visone R. MicroRNAs in autoimmunity and hematological Malignancies. *Int J Mol Sci*. (2018) 19:3139. doi: 10.3390/ijms19103139

184. Mayorga J, Alpízar-Rodríguez D, Prieto-Padilla J, Romero-Díaz J, Cravioto MC. Prevalence of premature ovarian failure in patients with systemic lupus erythematosus. *Lupus*. (2016) 25:675–83. doi: 10.1177/0961203315622824

185. Petri M, Kim MY, Kalunian KC, Grossman J, Hahn BH, Sammaritano LR, et al. Combined oral contraceptives in women with systemic lupus erythematosus. *N Engl J Med*. (2005) 353:2550–8. doi: 10.1056/NEJMoa051135

186. Mok CC, Lau CS. Profile of sex hormones in male patients with systemic lupus erythematosus. *Lupus*. (2000) 9:252–7. doi: 10.1191/096120300680198926

187. Noble PW, Young MR, Bernatsky S, Weisbart RH, Hansen JE. A nucleolytic lupus autoantibody is toxic to BRCA2-deficient cancer cells. *Sci Rep*. (2014) 4:5958. doi: 10.1038/srep05958

188. Gadalla SM, Amr S, Langenberg P, Baumgarten M, Davidson WF, Schairer C, et al. Breast cancer risk in elderly women with systemic autoimmune rheumatic diseases: a population-based case-control study. *Br J Cancer*. (2009) 100:817–21. doi: 10.1038/sj.bjc.6604906

189. Ladouceur A, Bernatsky S, Ramsey-Goldman R, Clarke AE. Managing cancer risk in patients with systemic lupus erythematous. *Expert Rev Clin Immunol*. (2018) 14:793–802. doi: 10.1080/1744666X.2018.1519394

190. Bernatsky SR, Cooper GS, Mill C, Ramsey-Goldman R, Clarke AE, Pineau CA. Cancer screening in patients with systemic lupus erythematosus. *J Rheumatol*. (2006) 33:45–9.

191. Bruera S, Lei X, Zogala R, Pundole X, Zhao H, Giordano SH, et al. Cervical cancer screening in women with systemic lupus erythematosus. *Arthritis Care Res (Hoboken)*. (2021) 73:1796–803. doi: 10.1002/acr.v73.12

192. Hsu CY, Lin MS, Su YJ, Cheng TT, Lin YS, Chen YC, et al. Cumulative immunosuppressant exposure is associated with diversified cancer risk among 14–832 patients with systemic lupus erythematosus: a nested case-control study. *Rheumatol (Oxford)*. (2017) 56:620–8. doi: 10.1093/rheumatology/kew457

193. Bernatsky S, Ramsey-Goldman R, Urowitz MB, Hanly JG, Gordon C, Petri MA, et al. Cancer risk in a large inception systemic lupus erythematosus cohort: effects of demographic characteristics, smoking, and medications. *Arthritis Care Res (Hoboken)*. (2021) 73:1789–95. doi: 10.1002/acr.v73.12

194. Wadström H, Arkema EV, Sjöwall C, Askling J, Simard JF. Cervical neoplasia in systemic lupus erythematosus: a nationwide study. *Rheumatol (Oxford)*. (2017) 56:613–9. doi: 10.1093/rheumatology/kew459

195. Sherston SN, Carroll RP, Harden PN, Wood KJ. Predictors of cancer risk in the long-term solid-organ transplant recipient. *Transplantation*. (2014) 97:605–11. doi: 10.1097/01.TP.0000436907.56425.5c

196. Hojo M, Morimoto T, Maluccio M, Asano T, Morimoto K, Lagman M, et al. Cyclosporine induces cancer progression by a cell-autonomous mechanism. *Nature*. (1999) 397:530–4. doi: 10.1038/17401

197. Ichinose K, Sato S, Igawa T, Okamoto M, Takatani A, Endo Y, et al. Evaluating the safety profile of calcineurin inhibitors: cancer risk in patients with systemic lupus erythematosus from the LUNA registry-a historical cohort study. *Arthritis Res Ther*. (2024) 26:48. doi: 10.1186/s13075-024-03285-x

198. Li XB, Cao NW, Chu XJ, Zhou HY, Wang H, Yu SJ, et al. Antimalarials may reduce cancer risk in patients with systemic lupus erythematosus: a systematic review and meta-analysis of prospective studies. *Ann Med*. (2021) 53:1687–95. doi: 10.1080/07853890.2021.1981547

199. Angeli F, Pedretti E, Garrafa E, Fredi M, Ceribelli A, Franceschini F, et al. Comparison of lineblot and immunoprecipitation methods in the detection of myositis-specific and myositis-associated antibodies in patients with idiopathic inflammatory myopathies: consistency with clinical diagnoses. *Diagnostics (Basel)*. (2024) 14:2192. doi: 10.3390/diagnostics14192192

200. Kobayashi S, Hoshino T, Hiwasa T, Satoh M, Rahmutulla B, Tsuchida S, et al. Anti-FIRs (PUF60) auto-antibodies are detected in the sera of early-stage colon cancer patients. *Oncotarget*. (2016) 7:82493–503. doi: 10.18632/oncotarget.12696

frontiers | Frontiers in Immunology

TYPE General Commentary
PUBLISHED 22 July 2025
DOI 10.3389/fimmu.2025.1641619

OPEN ACCESS

EDITED BY
Gunnar Houen,
University of Copenhagen, Denmark

REVIEWED BY
Nicole Trier,
University of Copenhagen, Denmark

*CORRESPONDENCE
Jiayi Chen
cjy13912736738@163.com

RECEIVED 06 June 2025
ACCEPTED 07 July 2025
PUBLISHED 22 July 2025

CITATION
Chen J (2025) Commentary: Cancer in connective tissue disease.
Front. Immunol. 16:1641619.
doi: 10.3389/fimmu.2025.1641619

Commentary: Cancer in connective tissue disease

Jiayi Chen*

Department of Stomatology, Suzhou Wujiang District Hospital of Traditional Chinese Medicine, Suzhou, China

KEYWORDS

autoantibodies, autoimmunity, connective tissue disease (CTD), immunology, malignancy

A Commentary on

Cancer in connective tissue disease

By Tonutti A, Ceribelli A, Gremese E, Colafrancesco S, De Santis M and Selmi C (2025). *Front. Immunol.* 16:1571700. doi: 10.3389/fimmu.2025.1571700

I read with great interest the comprehensive review by Tonutti et al. titled "Cancer in Connective Tissue Disease" (1), which provides a timely analysis of the bidirectional relationship between malignancy and autoimmunity in connective tissue diseases (CTDs). The authors adeptly synthesize current evidence on cancer risk stratification, autoantibody profiles, and screening challenges across systemic lupus erythematosus, systemic sclerosis, idiopathic inflammatory myopathies (IIM), and Sjögren's syndrome (SS). Their work underscores the critical need for multidisciplinary collaboration to address unmet needs in early detection and management.

I commend the authors for highlighting the paradoxical role of autoimmunity—where chronic inflammation may promote oncogenesis, yet autoimmune responses can also exert antitumor effects. This duality is exemplified by the contrasting implications of autoantibodies like anti-TIF1-γ (high cancer risk in IIM) and anti-Sp4/CCAR1 (potentially protective). However, I emphasize the urgent need for standardized autoantibody detection methods. As noted, discrepancies in anti-NXP2 results across assays (e.g., line blot vs. immunoprecipitation) complicate clinical interpretation (2). Harmonizing laboratory techniques is essential to refine risk stratification and validate guidelines like the IMACS cancer-screening algorithm (3).

I also support the call for disease-specific screening frameworks. While IMACS offers a model for IIM, similar protocols are lacking for systemic sclerosis and Sjögren's syndrome, where lymphoma risk escalates with biomarkers like ectopic germinal centers or CXCL13. Tailored strategies must integrate serological, clinical, and imaging data (e.g., salivary gland ultrasound in SS) while balancing cost-effectiveness and accessibility.

Finally, the impact of immunosuppressants on cancer risk warrants deeper exploration. Although the review notes inconclusive data on therapies like mycophenolate in systemic sclerosis, real-world studies are needed to clarify risks associated with newer biologics (e.g.,

rituximab) and the potential protective role of hydroxychloroquine. Pharmacovigilance registries could illuminate these associations.

In conclusion, Tonutti et al. have delivered an invaluable review that crystallizes the complex cancer-CTD interplay. Future efforts should prioritize validating autoantibody panels, expanding screening guidelines, and elucidating treatment-related oncogenic risks through international cohorts.

Author contributions

JC: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

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

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. The author would like to thank Deepseek for its assistance in the preparation of this letter.

Publisher's note

References

1. Tonutti A, Ceribelli A, Gremese E, ColaFrancesco S, De Santis M, Selmi C. Cancer in connective tissue disease. *Front Immunol.* (2025) 16: 1571700.

2. Cavazzana I, Fredi M, Ceribelli A, Mordenti C, Ferrari F, Carabellese N, et al. Testing for myositis specific autoantibodies: Comparison between line blot and immunoprecipitation assays in 57 myositis sera. *J Immunol Methods.* (2016) 433:1–5.

3. Tang IYK, Chan SCW, Li PH, Li WL, Luk LTH, Chan D, et al. Validation of the International Myositis Assessment and Clinical Studies Group guideline on cancer risk stratification. *Rheumatol (Oxford).* (2025) 64:2106–14.

frontiers | Frontiers in Immunology

TYPE Brief Research Report
PUBLISHED 22 May 2025
DOI 10.3389/fimmu.2025.1531390

OPEN ACCESS

EDITED BY
Esther Erdei,
University of New Mexico Health Sciences Center, United States

REVIEWED BY
Lovro Lamot,
University of Zagreb, Croatia
Lili Zhi,
Shandong Provincial Qianfoshan Hospital, China

*CORRESPONDENCE
Kristine Løkås Haftorn
kristine.haftorn@gmail.com
Hamid Khoshfekr Rudsari
khoshfekr1994@gmail.com

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

RECEIVED 20 November 2024
ACCEPTED 02 May 2025
PUBLISHED 22 May 2025

CITATION
Haftorn KL, Rudsari HK, Jaholkowski PP, Dåstøl VØ, Hestetun SV, Andreassen OA, Weinberg CR and Sanner H (2025) Nonlinearity and sex differences in the performance of a polygenic risk score for juvenile idiopathic arthritis.
Front. Immunol. 16:1531390.
doi: 10.3389/fimmu.2025.1531390

Nonlinearity and sex differences in the performance of a polygenic risk score for juvenile idiopathic arthritis

Kristine Løkås Haftorn[1*†], Hamid Khoshfekr Rudsari[1*†], Piotr Pawel Jaholkowski[2], Vilde Øverlien Dåstøl[1], Sigrid Valen Hestetun[1], Ole A. Andreassen[2,3], Clarice R. Weinberg[4] and Helga Sanner[1,5]

[1]Department of Rheumatology, Oslo University Hospital, Oslo, Norway, [2]Center for Precision Psychiatry, Division of Mental Health and Addiction, Oslo University Hospital, and Institute of Clinical Medicine, University of Oslo, Oslo, Norway, [3]K. G. Jebsen Centre for Neurodevelopmental Disorders, University of Oslo and Oslo University Hospital, Oslo, Norway, [4]Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Bethesda, MD, United States, [5]Norway and Institute of Clinical Medicine, University of Oslo, Oslo, Norway

Background: Juvenile idiopathic arthritis (JIA) is an immune-mediated pediatric disease believed to result from a complex interplay of genetic and environmental factors. Genome-wide association studies have enabled calculation of polygenic risk scores (PRS) for JIA. Understanding how the PRS associates with JIA and whether it performs similarly across sexes is essential for its utility in future studies.

Methods: We studied the relationship between a PRS developed from a previously published genome-wide association study of JIA and JIA in children from the Norwegian Mother, Father and Child Cohort Study (MoBa; total n = 57,630; JIA cases = 238). Generalized linear models (GLM) and generalized additive models (GAM) were used in logistic regression to assess the association. Furthermore, we investigated whether the relationship between PRS and JIA differed by sex by applying GAM models with interaction terms.

Results: PRS was significantly associated with JIA using both GLM ($p< 2e-16$) and GAM ($p< 2e-16$) models, and our results indicated a nonlinear relationship between PRS and JIA (effective degrees of freedom, EDF = 1.96). We found a significant interaction between sex and JIA PRS in relation to JIA ($p = 0.017$), and indications of a stronger and more logit-nonlinear relationship in females (EDF = 1.82) versus males (EDF = 1.06).

Conclusion: The relationship between PRS and JIA was slightly logit-nonlinear for females and logit-linear for males. The PRS for JIA can likely be used either as a continuous or discrete variable in analyses, but sex-stratification is recommended for future studies.

KEYWORDS

juvenile idiopathic arthritis, polygenic risk score, nonlinearity, sex differences, gene-sex interaction

1 Introduction

Juvenile idiopathic arthritis (JIA) is an immune-mediated disease characterized by joint inflammation lasting for at least six weeks and presenting before the age of 16 (1). It is a heterogenous disease with seven subtypes, and it is more prevalent in girls (2, 3). JIA imposes a significant burden on patients, their families, and society. It is believed to result from a complex interplay of genetic and environmental factors, although causal factors and underlying mechanisms remain largely unknown (4).

Familial, twin, and genome-wide association studies (GWAS) have helped to approach and dissect the genetic contribution to complex diseases, including JIA (5, 6). The monozygotic twin concordance rate of JIA has been estimated as 25-40%, and the sibling recurrence risk ratio as 11.6 (1). In the so far largest GWAS of JIA, including 3305 cases and 9196 controls, López-Isac et al. identified numerous susceptibility loci for JIA with a total SNP-based heritability of 0.61 (7).

The results from GWAS studies can be exploited by constructing polygenic risk scores (PRS), comprising aggregated effects of variants across the genome, which can be used to estimate the individual's genetic risk for the outcome of interest (8). PRS have been widely applied in studies of a range of different diseases and phenotypes and can be particularly useful in studies assessing the relationship between genetic and environmental risk factors for disease (9). Although PRSs have been suggested as potential clinical tools in the future, there are several obstacles that need to be addressed before they can be implemented into a clinical setting (9). PRSs are therefore so far mainly useful as research tools for studying genetic risk.

Recently, we developed a PRS for the children in the Norwegian Mother, Father and Child cohort study (MoBa) based on results from the aforementioned GWAS by López-Isac et al. (7, 10). When including a PRS in statistical models, either as a main effect or interaction variable, it is important to know how it relates to the outcome, in our case JIA. Understanding how the risk of JIA changes depending on the PRS can inform whether the PRS can be used as a continuous variable in the model or if it should be grouped into a discrete variable, and if so, how the discrete variable should be defined (11). Traditional logistic regression assumes a linear relationship between predictors and the log-odds of the outcome. However, some biological associations, including those between genetic risk scores and disease, may not follow a strictly linear pattern. Using nonlinear methods for modelling can therefore be useful because they are flexible enough to capture more complex relationships between the PRS and JIA. Furthermore, the PRS may be performing differently in specific subgroups, such as males and females, which can also be important to uncover when including the PRS in studies of risk and disease development (12).

Sex-specific genetic associations appear to play a role in a number of autoimmune and immune-mediated diseases, but the degree to which these differences contribute to JIA susceptibility has not been fully studied (13). A recent study on JIA patients found that the presence of antinuclear antibodies (ANA) was associated with specific genes, and this was observed more frequently in females, suggesting an interaction between certain genes and sex (14). Furthermore, a female-specific association between the *PTPN22* SNP rs2476601 and JIA has been confirmed across several different populations (15, 16), and evidence of a sex-specific association of *PSMA6/PSMC6/PSMA3* genetic variants with subtypes of JIA has also been reported (17). However, genome-wide studies of JIA, including the GWAS on which our PRS is based, were not stratified by sex (7). To address potential sex differences, it is thus important to assess whether the PRS performs similarly in males and females.

To fill these knowledge gaps, our aims of this study were 1) to investigate the relationship between the PRS for JIA and the probability of a JIA diagnosis, and 2) to explore whether the relationship between the PRS and JIA risk is different between males and females.

2 Methods

2.1 Study population and design

MoBa is a large-scale pregnancy cohort study led by the Norwegian Institute of Public Health (NIPH), which recruited participants across Norway between 1999 and 2008. 41% of the eligible women participated. The cohort comprises around 114,500 children, 95,200 mothers, and 75,200 fathers (18, 19). The present study uses version 12 of the MoBa data files, which underwent quality assurance and were made available for research in January 2019. We included MoBa children who had previously been genotyped (20).

2.2 Outcome

Information about JIA status was collected by linkage to the Norwegian Patient Registry (NPR), which includes data with personal ID numbers from all Norwegian public hospitals and specialists with public funding from 2008 (21). In Norway, the university hospitals with specialists within pediatric rheumatology have the main responsibility of diagnosing and following JIA patients. Cases were born between 1999 and 2009 and diagnosed with JIA before December 2021. We defined a JIA case as having at least two International Classification of Diseases (ICD)-10 codes (≥2 M08, ≥2 M09, or ≥1 M08 *and* ≥1 M09). In a recent validation of this case definition, we found a positive predictive value of 93.4% (10), ensuring a low number of false positive diagnoses. It is therefore reasonable to assume that our case definition largely reflects accurate diagnoses. For cases who received their first ICD-10 code in 2021, we accepted a single relevant ICD-10 code (M08 or M09), as we received our latest updates from NPR in December 2021. Controls were defined as non-JIA cases, and we removed all controls who had one ICD-10 code (M08 or M09) because they might have JIA.

2.3 Polygenic risk score for JIA

Umbilical cord blood samples were collected at birth, and the extracted DNA was frozen and stored at NIPH. The genotyping, quality control and imputation of the genetics data of the samples in MoBa have been extensively described previously (20). We calculated PRSs from the results of a previously published GWAS of JIA (7) by applying PRSice, version 2.3.3 (22). We chose *p*-value thresholds of 5E-8, 1E-6, 1E-5, 1E-4, 1E-3, 1E-2, 5E-2, 1E-1, and 1 to calculate PRSs and then extracted the first principal component (PC) for PRSs across all the thresholds, using this first PRS-PC as our final PRS for JIA (23). We then, using the whole dataset, standardized the PRS to a mean of zero and a standard deviation (SD) of 1 (24) and we used the standardized PRS for all analyses. In sensitivity analyses, the PRS was categorized into (1) quartiles, forming four equal-sized categories, (2) three categories containing the top 10%, middle 80% and bottom 10% of observations, and (3) a binary variable based on the median (Supplementary Table 1).

2.4 Statistical analysis

R version 4.2.3 was used to conduct all statistical analyses (25), and all scripts are available in our GitHub repository (https://github.com/KristineLH/PRS-JIA-sex). We used multiple logistic regression and generalized additive models (GAM) to examine the relationship between PRS and JIA. The top 10 PCs from the whole genotype dataset, together with sex, and year of birth were included as covariates in the models.

Nonlinear modeling approach

To account for potential logit-nonlinearity, we applied GAM using the *gam* function from the mgcv package (26). GAM extends traditional regression by allowing flexibility in how predictors influence the outcome, fitting smooth, data-driven curves rather than assuming a fixed logit-linear form. In our model, PRS was modeled as a smooth function using a regression spline, which adapts to the shape of the data. The effective degrees of freedom (EDF) from the GAM output served as an indicator of nonlinearity, with an EDF of 1 representing a linear relationship and values greater than 1 suggesting a nonlinear relationship (27).

Modeling sex differences

To investigate whether the relationship between the PRS and JIA differed by sex, we first included an interaction term between the PRS and sex in the multiple logistic regression model. The Wald test was used to assess statistical significance of the interaction, and a p-value < 0.05 was regarded as significant. However, interaction terms in standard regression models assume a constant, linear modification of the association by sex, which may not fully capture potential differences in the way the PRS is associated with JIA in males and females. To address this, we further investigated sex-specific patterns by fitting separate smooth splines for the PRS in males and females. Specifically, we created new variables by multiplying PRS with dummy variables for each sex and then modeled these products as smooth terms in the GAM framework. This allowed us to estimate the association between the PRS and JIA in each sex separately.

Visualization

To aid interpretation, we visualized the relationship between PRS and JIA for each model. Using the *predict* function, we calculated the probability of JIA across a range of PRS values (-4.5 to 4.5 with an increment of 0.1), while keeping other covariates (10 PCs, year of birth) at their mean values. This enabled direct comparison of PRS effects across methods (Figure 1) and sexes (Figure 2).

3 Results

3.1 Study sample characteristics

Our final analytical sample included 57,630 children of whom 238 were identified as JIA-cases (Table 1). Male participants comprised 51.0% (n = 29,139) of the controls, compared to only 39.9% (n = 91) of the JIA cases. The JIA cases had a mean PRS of 0.58 (+/- 1.10 SD), whereas the mean PRS in controls was -0.002 (+/- 1.00 SD).

3.2 Association between PRS and JIA

We assessed the association between PRS and JIA using a standard logistic regression model (GLM) and a generalized additive model (GAM), results shown in Figure 1. In both

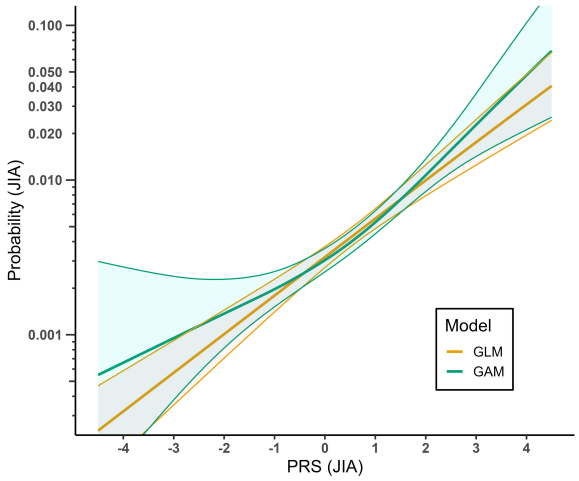


FIGURE 1
Relationship between PRS for JIA and probability of JIA modelled by a generalized linear model (GLM) compared to a generalized additive model (GAM). The lines show the fitted prediction model of JIA probability ~ PRS + sex + year of birth + top 10 principal components for each of the models. The colored areas represent the 95% confidence intervals for the corresponding models.

10.3389/fimmu.2025.1531390

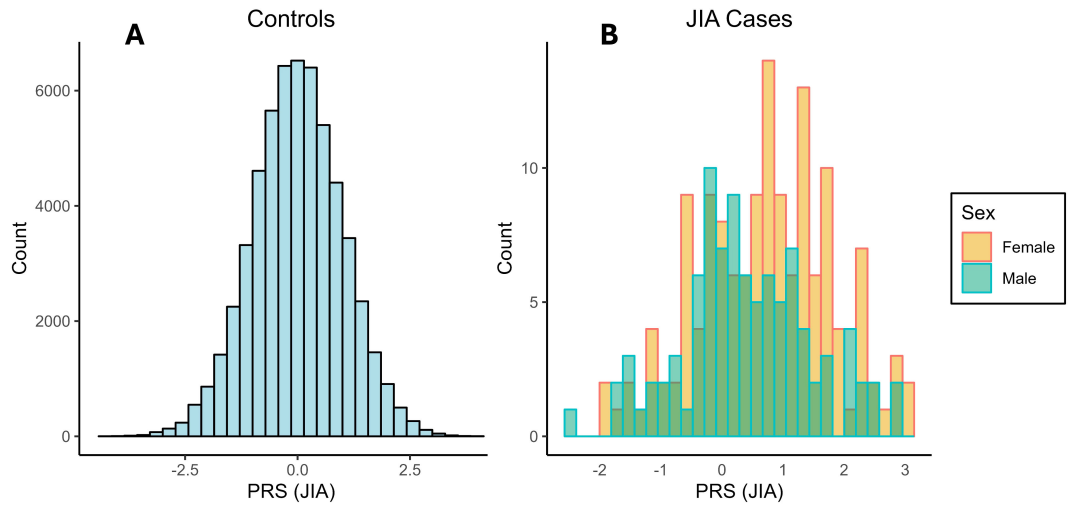


FIGURE 2
Distribution of JIA PRS in **(A)** controls, and **(B)** JIA cases grouped by sex.

models, PRS was significantly associated with JIA ($p < 2e\text{-}16$ for both models), and the results were similar for the categorized PRS variables (Supplementary Figure 1). The EDF in our GAM model was 1.939, indicating a logit-nonlinear relationship between PRS and risk of JIA.

3.3 The association between PRS and JIA differs by sex

In Figure 2, we show the distributions of PRS in controls, as well as cases stratified by sex. The PRS distributions for controls show a mean of 0.01 in males and -0.01 in females. In contrast, JIA cases demonstrate higher PRS means. Specifically, the PRS mean for male cases is 0.40, while for female cases, it is 0.70, indicating a stronger association between PRS and JIA diagnosis in females compared to males.

We further investigated the interaction between sex and PRS in association with JIA. In a simple logit-linear model, the interaction term between sex and PRS was significantly associated with JIA ($p = 0.017$). We then investigated this interaction further by conducting a semi-stratified analysis allowing for nonlinear relationships (Figure 3). This model showed that PRS was significantly associated with JIA in both females ($p < 2e\text{-}16$) and males ($p < 0.001$). Interestingly, the relationship between PRS and JIA was approximately logit-linear in males (EDF = 1.06) but showed a larger tendency of logit-nonlinearity in females (EDF = 1.82). We detected a similar pattern when defining the PRS as high- and low-risk variable divided into top 10%, bottom 10% and middle 80% of observations (Supplementary Figure 2).

4 Discussion

Our results show that the relationship between PRS and JIA is weakly logit-nonlinear. The notable difference in PRS distribution between male and female JIA cases underscores a sex-specific variation in PRS among JIA cases in the MoBa cohort. Furthermore, we show a significant interaction between sex and PRS in relation to JIA, with sex acting as a PRS effect measure modifier. Interestingly, the logit-nonlinearity of the relationship seems to be driven by the females, whereas in males the relationship seems to be logit-linear.

Understanding the relationship between a PRS and the outcome of interest is important when the PRS is to be used in further analyses, such as when investigating interactions between environmental exposures and genetic predisposition to develop JIA. Particularly, for the PRS to be used as a continuous variable

TABLE 1 Study sample characteristics.

Characteristics	JIA cases			Controls		
	All	Male	Female	All	Male	Female
Sample size (n,%)	238 (100)	95 (39.9)	143 (60.1)	57,392	29,319 (51.0)	28,073 (48.9)
Year of birth (mean, SD)	2005 (2.18)	2004 (2.19)	2005 (2.15)	2005 (2.17)	2005 (2.17)	2005 (2.16)
PRS (mean, SD)	0.579 (1.10)	0.399 (1.08)	0.699 (1.11)	-0.002 (1.00)	0.005 (1.00)	-0.010 (0.99)

10.3389/fimmu.2025.1531390

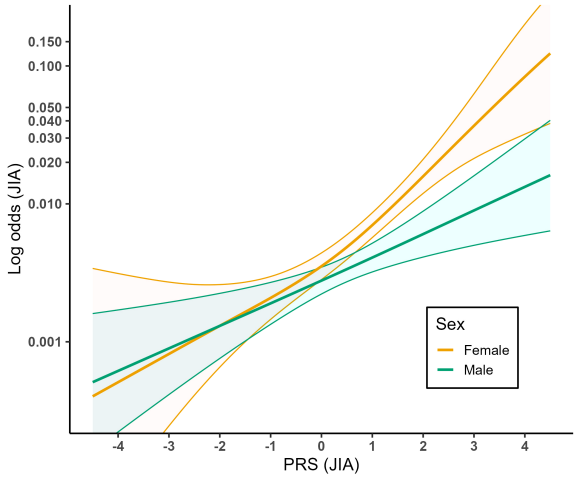


FIGURE 3
Relationship between PRS of JIA and probability of JIA in females and males. The lines show the fitted prediction model of JIA probability ~ PRS + sex + year of birth + top 10 principal components for each of the sexes. The colored areas represent the 95% confidence intervals for the corresponding models.

in analyses of JIA, the relationship between PRS and JIA should be well modelled (11). A non-linear relationship between the PRS and JIA could suggest that, for risk prediction, the PRS should be grouped into a discrete variable. Our results indicate a somewhat logit-nonlinear relationship between the PRS for JIA and risk of JIA, with a stronger effect with higher PRS compared to lower PRS. However, as shown in Figure 1, the GAM model taking logit-nonlinear associations into account is not vastly different from the simple logit-linear model. It should be noted, however, that the relatively small number of JIA cases in our dataset may have limited our power to detect subtle nonlinear interactions, particularly for males. Although grouping the PRS into a categorical variable as shown in Supplementary Figure 1 gave a similar fit, the predicted probabilities of JIA were lower than with the continuous PRS, especially for the high-risk groups. This indicates some loss of information and shrinkage towards the mean due to grouping the PRS. Thus, we suggest using PRS as a continuous variable in future studies when possible. Grouping the PRS into high- and low-risk groups of top 10%, bottom 10% and middle 80% gave the most similar fit compared to using the PRS as a continuous variable and may therefore be an alternative way of modelling the PRS. However, males and females appear to require distinct models for use of this PRS for JIA.

Sex-specific and sex-dependent effects of PRSs for other diseases, like schizophrenia and coronary artery disease have also been reported (28–30). The difference we observe in PRS performance between the sexes could reflect differences in the sex ratio among cases and controls in the GWAS our PRS is based on (12), with the girl cases outnumbering the boys and consequently having more influence on the formation of the score. However, the sex ratios were not stated in the GWAS paper, which may limit our results (7). Furthermore, different subtypes of JIA are associated with different genetic loci, and sex distribution also differs depending on the subtype (3). Some subtypes, such as oligoarticular and polyarticular JIA, which constitute around 70% of all cases, occur 2–3 times more frequently in girls, but not all JIA subtypes are more common in females (3). Thus, the PRS may be mainly reflecting genetic predisposition for the more common subtypes which are also more common in females and therefore show a stronger association with JIA in females compared to males. We did not have access to information on subtypes in our dataset and were thus not able to account for this in our analyses. Given that certain JIA subtypes differ in their genetic patterns, this represents a limitation of our study. Furthermore, gene-environment interactions involving exposures that differ by sex, such as hormones, have not been accounted for and may have influenced our results. Finally, our results may indicate that the effect of genetic predisposition on JIA development is dependent on biological processes that differ between the sexes.

When using the PRS for JIA in association and interaction analyses, researchers should be aware of the sex-specific associations and consider sex-stratification when possible. Our findings suggest that future studies on the genetic predisposition to JIA, including GWAS and the development of PRS, should incorporate sex-specific analyses to identify genetic loci that may contribute to disease development in males and females separately, as well as those shared between sexes (31, 32). Developing a set of distinct PRS scores specifically for sex-by-subtype categories could prove to be even more usefully predictive, but this would require a very large genetic dataset with detailed information on sex and JIA subtypes. We also suggest exploring potential susceptibility loci for JIA on the X-chromosome (33) as this was not included in our study nor, to our knowledge, in any GWAS of JIA thus far. As sex differences are common in autoimmune diseases in general, investigating sex-specific associations of PRS may be relevant also for other autoimmune and immune-mediated diseases (34).

In conclusion, our results show that the relationship between our PRS and JIA is slightly logit-nonlinear, but only for females.

The PRS for JIA can likely be used either as a continuous or discrete variable in analyses, but sex-stratification should be considered. Future studies should further investigate sex-differences in genetic predisposition of JIA and other autoimmune diseases.

Data availability statement

Access to MoBa data can be obtained by applying to the Norwegian Institute of Public Health (NIPH). Restrictions apply regarding the availability of these data, which were used under specific approvals for the current study and therefore not publicly available. Access can only be given after approval by the Regional Committees for Medical and Health Research Ethics (REC) in compliance with the EU General Data Protection Regulation (GDPR) and approval from the data owners. The consent given by the participants does not open for storage of data on an individual level in repositories or journals. Requests to access these datasets should be directed to helsedata.no/en.

Ethics statement

The studies involving humans were approved by The Regional Committees of Medical and Health Research Ethics South East (project number 28469) and the Norwegian Data Protection Agency. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

KH: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. HR: Conceptualization, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing, Methodology. PJ: Writing – original draft, Writing – review & editing. VD: Writing – original draft, Writing – review & editing. SH: Writing – original draft, Writing – review & editing. OA: Data curation, Funding acquisition, Resources, Writing – original draft, Writing – review & editing. CW: Methodology, Writing – original draft, Writing – review & editing. HS: Data curation, Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was funded by the South-Eastern Norway Regional Health Authority (HSO: PREVENT-JIA Prospective evaluation of early-life modifiable environmental factors and genetic risk in juvenile idiopathic arthritis), grant number 340421. CRW are supported by the intramural research program of the National Institute of Environmental Health Sciences, National Institutes of Health.

Acknowledgments

We thank Prof. Håkon K. Gjessing and Dr. Yunsung Lee for useful discussions regarding this work. We are grateful to all the participating families in MoBa, and we thank the Norwegian Institute of Public Health (NIPH) for generating high-quality genomic data. This research is part of the HARVEST collaboration, supported by the Research Council of Norway (#229624). We also thank the NORMENT Centre for providing genotype data, funded by the Research Council of Norway (#223273), South-East Norway Health Authorities and Stiftelsen Kristian Gerhard Jebsen. We further thank the Center for Diabetes Research, the University of Bergen for providing genotype data and performing quality control and imputation of the data funded by the ERC AdG project SELECTionPREDISPOSED, Stiftelsen Kristian Gerhard Jebsen, Trond Mohn Foundation, the Research Council of Norway, the Novo Nordisk Foundation, the University of Bergen, and the Western Norway Health Authorities.

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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2025.1531390/full#supplementary-material

References

1. Martini A, Lovell DJ, Albani S, Brunner HI, Hyrich KL, Thompson SD, et al. Juvenile idiopathic arthritis. *Nat Rev Dis Primers*. (2022) 8:5. doi: 10.1038/s41572-021-00332-8

2. Cattalini M, Soliani M, Caparello MC, Cimaz R. Sex differences in pediatric rheumatology. *Clin Rev Allergy Immunol*. (2019) 56:293–307. doi: 10.1007/s12016-017-8642-3

3. Ravelli A, Martini A. Juvenile idiopathic arthritis. *Lancet*. (2007) 369:767–78. doi: 10.1016/S0140-6736(07)60363-8

4. Horton DB, Shenoi S. Review of environmental factors and juvenile idiopathic arthritis. *Open Access Rheumatol*. (2019) 11:253–67. doi: 10.2147/OARRR.S165916

5. La Bella S, Rinaldi M, Di Ludovico A, Di Donato G, Di Donato G, Salpietro V, et al. Genetic background and molecular mechanisms of juvenile idiopathic arthritis. *Int J Mol Sci*. (2023) 24. doi: 10.3390/ijms24031846

6. Cobb JE, Hinks A, Thomson W. The genetics of juvenile idiopathic arthritis: current understanding and future prospects. *Rheumatol (Oxford)*. (2014) 53:592–9. doi: 10.1093/rheumatology/ket314

7. López-Isac E, Smith SL, Marion MC, Wood A, Sudman M, Yarwood A, et al. Combined genetic analysis of juvenile idiopathic arthritis clinical subtypes identifies novel risk loci, target genes and key regulatory mechanisms. *Ann Rheum Dis*. (2021) 80:321–8. doi: 10.1136/annrheumdis-2020-218481

8. Choi SW, Mak TS, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc*. (2020) 15:2759–72. doi: 10.1038/s41596-020-0353-1

9. Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. *Genome Med*. (2020) 12:44. doi: 10.1186/s13073-020-00742-5

10. Hestetun S, Rudsari HK, Jaholkowski PP, Shadrin AA, Haftorn KL, Andersen S, et al. Incidence and genetic risk of juvenile idiopathic arthritis in Norway by latitude. *Arthritis Rheumatol*. (2025) 77:458–67. doi: 10.1002/art.43040

11. Collister JA, Liu X, Clifton L. Calculating polygenic risk scores (PRS) in UK biobank: A practical guide for epidemiologists. *Front Genet*. (2022) 13:818574. doi: 10.3389/fgene.2022.818574

12. Mostafavi H, Harpak A, Agarwal I, Conley D, Pritchard JK, Przeworski M. Variable prediction accuracy of polygenic scores within an ancestry group. *Elife*. (2020) 9. doi: 10.7554/eLife.48376

13. Piccini P, Montagnani C, de Martino M. Gender disparity in pediatrics: a review of the current literature. *Ital J Pediatr*. (2018) 44:1. doi: 10.1186/s13052-017-0437-x

14. Saleh R, Sundberg E, Olsson M, Tengvall K, Alfredsson L, Kockum I, et al. Genetic association of antinuclear antibodies with HLA in JIA patients: a Swedish cohort study. *Pediatr Rheumatol Online J*. (2024) 22:79. doi: 10.1186/s12969-024-01017-8

15. Goulielmos GN, Chiaroni-Clarke RC, Dimopoulou DG, Zervou MI, Trachana M, Pratsidou-Gertsi P, et al. Association of juvenile idiopathic arthritis with PTPN22 rs2476601 is specific to females in a Greek population. *Pediatr Rheumatol Online J*. (2016) 14:25. doi: 10.1186/s12969-016-0087-3

16. Chiaroni-Clarke RC, Li YR, Munro JE, Chavez RA, Scurrah KJ, Pezic A, et al. The association of PTPN22 rs2476601 with juvenile idiopathic arthritis is specific to females. *Genes Immun*. (2015) 16:495–8. doi: 10.1038/gene.2015.32

17. Sjakste T, Paramonova N, Rumba-Rozenfelde I, Trapina I, Sugoka O, Sjakste N. Juvenile idiopathic arthritis subtype- and sex-specific associations with genetic variants in the PSMA6/PSMC6/PSMA3 gene cluster. *Pediatr Neonatol*. (2014) 55:393–403. doi: 10.1016/j.pedneo.2014.01.007

18. Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, et al. Cohort profile update: the norwegian mother and child cohort study (MoBa). *Int J Epidemiol*. (2016) 45:382–8. doi: 10.1093/ije/dyw029

19. Paltiel L, Anita H, Skjerden T, Harbak K, Bækken S, Nina Kristin S, et al. The biobank of the Norwegian Mother and Child Cohort Study – present status. *Norsk Epidemiol*. (2014) 24:1–2. doi: 10.5324/nje.v24i1-2.1755

20. Corfield EC, Shadrin AA, Frei O, Rahman Z, Lin A, Athanasiu L, et al. The Norwegian Mother, Father, and Child cohort study (MoBa) genotyping data resource: MoBaPsychGen pipeline v.1. *bioRxiv*. (2024). doi: 10.1101/2022.06.23.496289

21. Bakken IJ, Ariansen AMS, Knudsen GP, Johansen KI, Vollset SE. The Norwegian Patient Registry and the Norwegian Registry for Primary Health Care: Research potential of two nationwide health-care registries. *Scand J Public Health*. (2020) 48:49–55. doi: 10.1177/1403494819859737

22. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience*. (2019) 8. doi: 10.1093/gigascience/giz082

23. Coombes BJ, Ploner A, Bergen SE, Biernacka JM. A principal component approach to improve association testing with polygenic risk scores. *Genet Epidemiol*. (2020) 44:676–86. doi: 10.1002/gepi.22339

24. Sedaghati-Khayat B, Boer CG, Runhaar J, Bierma-Zeinstra SMA, Broer L, Ikram MA, et al. Risk assessment for hip and knee osteoarthritis using polygenic risk scores. *Arthritis Rheumatol*. (2022) 74:1488–96. doi: 10.1002/art.42246

25. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing (2023). Available at: https://www.R-project.org/.

26. Wood SN. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *J Am Stat Assoc*. (2004) 99:673–86. doi: 10.1198/016214504000000980

27. Wood S. (2017). Generalized Additive Models: An Introduction with R, Second Edition (2nd ed.). Chapman and Hall/CRC. doi: 10.1201/9781315370279

28. Koch E, Nyberg L, Lundquist A, Pudas S, Adolfsson R, Kauppi K. Sex-specific effects of polygenic risk for schizophrenia on lifespan cognitive functioning in healthy individuals. *Trans Psychiatry*. (2021) 11:520. doi: 10.1038/s41398-021-01649-4

29. Paruchuri K, Bhukar R, Urbut S, Ruan Y, Ganesh S, Postupaka D, et al. Using sex-specific polygenic risk to prognosticate coronary artery disease in women. *J Am Heart Assoc*. (2024) 13:e034946. doi: 10.1161/JAHA.123.034946

30. Surakka I, Wolford BN, Ritchie SC, Hornsby WE, Sutton NR, Elvenstad Gabrielsen M, et al. Sex-specific survival bias and interaction modeling in coronary artery disease risk prediction. *Circ: Genom Precis Med*. (2023) 16:e003542. doi: 10.1161/CIRCGEN.121.003542

31. Hughes T, Adler A, Merrill JT, Kelly JA, Kaufman KM, Williams A, et al. Analysis of autosomal genes reveals gene-sex interactions and higher total genetic risk in men with systemic lupus erythematosus. *Ann Rheum Dis*. (2012) 71:694–9. doi: 10.1136/annrheumdis-2011-200385

32. Zhang C, Ye Y, Zhao H. Comparison of methods utilizing sex-specific PRSs derived from GWAS summary statistics. *Front Genet*. (2022) 13:892950. doi: 10.3389/fgene.2022.892950

33. Sakkas LI, Chikanza IC. Sex bias in immune response: it is time to include the sex variable in studies of autoimmune rheumatic diseases. *Rheumatol Int*. (2024) 44:203–9. doi: 10.1007/s00296-023-05446-8

34. Lahita RG. Sex and gender influence on immunity and autoimmunity. *Front Immunol*. (2023) 14. doi: 10.3389/fimmu.2023.1142723

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 15 August 2025
DOI 10.3389/fimmu.2025.1539659

OPEN ACCESS

EDITED BY
Maria Giovanna Danieli,
Università Politecnica delle Marche, Italy

REVIEWED BY
Silvia Sánchez-Ramón,
Complutense University of Madrid, Spain
Claire Deakin,
University College London, United Kingdom

*CORRESPONDENCE
Frederick W. Miller
millerf@mail.nih.gov

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

‡These authors have contributed equally to this work and share senior authorship

RECEIVED 04 December 2024
ACCEPTED 29 July 2025
PUBLISHED 15 August 2025

CITATION
Alhassan E, Patnaik A, Shamim EA, Pandey JP, Rider LG and Miller FW (2025) A possible role for immunogenetic factors in myositis developing after vaccination in the pre-covid-19 era. *Front. Immunol.* 16:1539659. doi: 10.3389/fimmu.2025.1539659

A possible role for immunogenetic factors in myositis developing after vaccination in the pre-covid-19 era

Eaman Alhassan[1†], Anna Patnaik[2,3†], Ejaz A. Shamim[3,4,5†], Janardan P. Pandey[6], Lisa G. Rider[3‡] and Frederick W. Miller[3,7*‡]

[1]Division of Rheumatology and Clinical Immunology, Department of Medicine, University of Pittsburgh Medical Center, Pittsburgh, PA, United States, [2]Department of Neurology, Johns Hopkins University, Baltimore, MD, United States, [3]Environmental Autoimmunity Group, Clinical Research Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Bethesda, MD, United States, [4]Mid-Atlantic Permanente Medical Group and Mid-Atlantic Permanente Research Institute, Washington DC, United States, [5]National Institutes of Health, National Institute of Neurological Disorders and Stroke, Bethesda, MD, United States, [6]Department of Pharmacology and Immunology, Medical University of South Carolina, Charleston, SC, United States, [7]Environmental Autoimmunity Group, Clinical Research Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Durham, NC, United States

Introduction: Vaccinations have had a transformative impact on public health, reducing the incidence of many infectious diseases and increasing survival. However, there remains uncertainty about the potential of vaccines to trigger autoimmune diseases such as the idiopathic inflammatory myopathies (IIM). Myositis after vaccination (MAV) is a rare clinical entity, but given immunogenetic associations with other adverse events, we explored genetic risk factors, particularly human leukocyte antigen (HLA) alleles and GM/KM immunoglobulin allotypes, that may predispose individuals to develop MAV.

Methods: We examined clinical characteristics, vaccination history, autoantibodies, HLA alleles and GM/KM allotypes from 56 patients who developed MAV, 133 myositis cases with no documented vaccination within 6 months of onset (non-MAV), and 527 healthy controls from the pre-COVID-19 era. Genotyping for HLA and GM/KM allotypes was performed by standard assays. Differences in allele frequencies in race-matched groups were evaluated using chi-square tests, odds ratios (OR) and 95% confidence intervals (CI). Multivariate logistic regression adjusted for age, sex, and vaccination type. Statistical significance was defined as a Holms corrected p-value of less than 0.05.

Results: No clinical or serologic differences were found between MAV and non-MAV patients. However, the HLA-DQA1*03:03 allele was a unique risk factor for MAV in Caucasians (OR=3.87, 95% CI=1.56-9.54, p=0.002), while the known myositis risk factor, HLA-DRB1*03:01, was a protective factor for MAV (OR=0.41, 95% CI=0,18-0.94, p= 0.033). GM2, GM13, and KM1 allotypes were more frequently observed in MAV patients than healthy controls, and other HLA alleles were risk or protective factors for specific vaccines given in patients who developed MAV.

Conclusion: Immunogenetic factors may influence the likelihood of developing MAV. Further studies of larger, deeply phenotyped populations are needed to confirm these associations and could inform personalized risk assessments and targeted interventions, thereby enhancing vaccine safety.

KEYWORDS

polymyositis, dermatomyositis, vaccination, adverse events, HLA, GM/KM, juvenile dermatomyositis

1 Introduction

The idiopathic inflammatory myopathies (IIM) are a group of rare systemic autoimmune conditions characterized by muscle inflammation and weakness that arise from chronic immune activation in genetically predisposed individuals in response to certain environmental triggers (1). Major strides have been made in defining the genetic risks for IIM and other autoimmune conditions (2), but identifying the even more important environmental risk factors has been hampered by the lack of validated measures and the constantly changing mixtures of exposures that occur over a lifetime (3). Vaccines, while highly beneficial, can in rare cases, cause chronic immune activation followed by the development of a number of autoimmune diseases, including myositis (4, 5).

Certain polymorphic immune response genes have been associated with IIM. One of the strongest genetic associations for autoimmune diseases is located on chromosome 6p21.3 that includes the human leukocyte antigen (HLA) locus in addition to other immune system-modulating genes (6). Alleles of the 8.1 ancestral haplotype (8.1 AH), *HLA-DRB1*03:01* and *HLA-B*08:01*, show the strongest association with IIM in Caucasians (7, 8). Other polymorphic genes associated with autoimmune diseases, including IIM, are the immune response genes that encode immunoglobulin gamma heavy chains (GM) and immunoglobulin kappa light chains (KM) (9). These have also been identified as genetic susceptibility factors across different ages and ethnicities for various clinical and serological IIM phenotypes (10).

There is no doubt that vaccines have significantly improved global public health by boosting immune responses to many infectious agents, preventing infections, and minimizing morbidity and mortality. However, it is plausible that vaccines, often given intramuscularly, could cause initial immune activation in muscles to progress to a chronic systemic inflammatory response in those with certain immunogenetic backgrounds. While many patients develop myositis without any documented recent vaccination, the concept that vaccinations may be linked to the onset of some cases of myositis has been previously suggested in case reports (11–14). The first identified cases of myositis following vaccination (MAV) included myositis developing in a temporally related way to diphtheria-tetanus-pertussis vaccines (11, 15, 16) and smallpox vaccines (17) in adult and pediatric patients. Additional reports of vaccine constituents, including aluminum hydroxide, and not the immunization antigens themselves, have led to macrophagic myofasciitis (18).

Certain adverse events to drugs, medical implants and vaccines have previously been associated with clinical, serologic or immunogenetic features (19–21). Based on our observation that some myositis cases were temporally associated with vaccinations, we systematically compared those patients who developed myositis within 6 months of a documented vaccination to those who had no documented vaccinations within 6 months of myositis disease onset and to healthy controls (HC) from the pre-COVID-19 era to assess possible clinical, serological, and immunogenetic differences.

2 Materials and methods

2.1 Study participants

Myositis patients and HC were enrolled into investigational review board-approved clinical protocols at the National Institutes of Health (NIH) Warren Grant Magnuson Clinical Center and the United States Food and Drug Administration from 1983 to 2002. These protocols studied the natural history of myositis and twins and siblings discordant for myositis.

Per our protocol criteria, all patients met Bohan and Peter criteria for definite or probable myositis (22, 23). They were all diagnosed with IIM, including dermatomyositis (DM), juvenile dermatomyositis (JDM), polymyositis (PM), juvenile polymyositis (JPM), and inclusion body myositis (IBM) based on the accepted criteria at the time of enrollment. Patients with myositis and another connective tissue disease (CTM) were also included. IIM patients who received a documented vaccination within six months prior to first myositis symptom onset were included in the myositis after vaccination (MAV) group (n=56), while those who did not receive vaccination during this time interval (documented by history and review of medical records) were categorized as non-MAV (n=133). All patients underwent a comprehensive medical history and physical examination, which included detailed protocol

questionnaires completed by the patients and their enrolling physicians.

The clinical data included age, self-classified race, gender, and signs and symptoms. Since gene frequencies differ by race, the HLA and GM/KM data were assessed in Caucasian patients, which was the largest cohort and the only one adequate for reliable statistical analysis. The HC groups were race-matched.

2.2 HLA typing

HLA allele typing was performed using purified genomic DNA, using laboratory-designed and commercial reagents (Genovision, West Chester, PA; Dynal Biotech, Lafayette Hill, PA) and PCR-mediated sequence-specific oligonucleotide probe hybridization and sequence-specific priming technique via standard techniques (24).

Allele frequencies per patient (carriage rates) were determined by the number of allele-positive subjects divided by the total number of subjects for which complete HLA data were available at a given locus. All patients in the HLA allele analysis were self-identified as Caucasians and divided into MAV (n=48) and non-MAV (n=93) groups. For comparison, the HC data (n=527), who did not have myositis, were obtained through the NIH HLA laboratory.

2.3 GM and KM allotyping

Immunoglobulin gamma heavy chain (GM) and immunoglobulin kappa light chain (KM) allotyping was performed using standard hemagglutination inhibition methods to type for IgG1m, IgG2m, and IgG3m and for IgKM1 and IgKM3 (25). Allotype and phenotype frequencies were determined by the number of allotype-positive subjects divided by the total number of subjects for which data were available at a given locus. All patients in the GM and KM allotype analysis were Caucasian and divided into MAV (n=19) and non-MAV (n=34) cases. Race-matched HC (n=266) were used for comparison.

2.4 Autoantibody identification

Myositis-specific autoantibodies (anti-synthetases, anti-signal recognition particle (anti-SRP), anti–Mi-2 and myositis-associated autoantibodies (anti-Ku, anti-La, anti-Ro, anti–URNP, and anti–PM-Scl), were identified from frozen serum samples using previously validated methods of protein and RNA immunoprecipitation (IPP) and double immunodiffusion (10). The NXP2 and TIF1 autoantibodies were identified with IPP, followed by immunoblotting (26).

2.5 Statistical analysis

Analyses were performed using GraphPad Prism (GraphPad, Inc., La Jolla, CA). For both the HLA allele analysis and the GM/KM allotype analysis, the allele or allotype frequencies were compared by chi-square test or Fisher's exact test for counts below 5, for 2x2 contingency tables between MAV and controls, MAV and non-MAV, or non-MAV and controls. The odds ratios (OR), 95% confidence intervals (CI) were determined. The MAV group was also divided and compared to non-MAV and HC by the four most frequent vaccines: Hepatitis B, Influenza, Tetanus, and Mumps-Measles-Rubella (MMR).

A p-value was considered significant if below 0.05 using the Holm procedure to adjust for multiple comparisons (27). The U-test, or Mann-Whitney test, was used to compare non-parametric variables, such as the months from vaccine to first symptom, calculations between children and adults, and between the different vaccines.

Chi-square tests were performed to examine differences in the frequency distributions between the MAV and non-MAV groups. An analysis in which the distribution of clinical subgroups significantly differed between the MAV and non-MAV groups led to performing a sensitivity analysis, in which a random sample of patients were selected in similar clinical subgroups. This was also performed with the MAV group within 6 months and 3 months from vaccination. If the genetic results differed from the primary analysis, the difference in clinical subgroup distribution was interpreted to have affected the result, however, if the genetic results remained the same, the difference in clinical subgroup distribution was interpreted as not affecting the genetic results.

3 Results

3.1 Clinical findings

There were 56 patients, including 28 females, in the MAV group, 48 of whom were Caucasian, three African American, and five of mixed race, and 133 patients, including 92 females, in the non-MAV group, of which 98 were Caucasian, 12 African American, six Asian or Hispanic, and 17 of mixed race. Of these, 48 MAV patients and 95 non-MAV patients were Caucasian and HLA-typed, while 19 MAV patients and 34 non-MAV patients were Caucasian and also underwent GM/KM typing. The clinical and autoantibody subgroup, race, gender, and signs and symptom distributions were similar in the MAV and non-MAV groups for all patients included in the study (Table 1), as well as for the HLA-analyzed groups. The patients in which GM/KM was examined had a lower frequency of JDM in the MAV group (21.1%) and a higher frequency of JDM in the non-MAV group (73.5%) ($p = 0.0004$). The median age of disease onset for the MAV group was 5.4 years in children and 43.8 years in adults, which was similar to the non-MAV group (6.7 and 45.4 years, respectively).

Of the 56 MAV patients, 17 received a form of the tetanus vaccine, 15 received a Hepatitis B vaccine, 15 received an influenza vaccine, and 13 received a MMR vaccine (Table 2). The median time to myositis symptoms after vaccination was 2.2 months with a range of 0–6 months and an IQR of 3.5 months, while the median time to diagnosis of myositis after vaccination was 7.0 months.

TABLE 1 Distribution of clinical and autoantibody subgroups, and signs and symptoms of myositis patients developing symptoms of myositis within 6 months of vaccination (MAV) and those without documented vaccination within 6 months of symptom onset (non-MAV).

Clinical Groups*	MAV (n=56)	Non-MAV (n=133)
	N (%)	N (%)
JDM	26 (46.4)	84 (63.2)
DM	10 (17.8)	17 (12.8)
PM	13 (23.2)	15 (11.3)
CTM	3 (5.4)	7 (5.3)
IBM	2 (3.6)	6 (4.5)
JPM	2 (3.6)	4 (3.0)
Myositis-Autoantibody Groups*+		
MSA and MAA Negative	36 (64.3)	89 (66.9)
p155 (TIF1)	14 (25.0)	40 (30.1)
Mi-2	5 (8.9)	6 (4.5)
MJ (NXP2)	3 (5.4)	23 (17.3)
SRP	3 (5.4)	6 (4.5)
Aminoacyl tRNA-Synthetases	4 (7.2)	11 (8.3)
Ro60	5 (8.9)	12 (9.0)
PM-Scl	2 (3.6)	4 (3.0)
U1RNP	1 (1.8)	6 (4.5)
Clinical Features*		
Myalgia	35 (63.6)	88 (67.2)
Distal muscle weakness	29 (51.8)	62 (47.0)
Muscle atrophy	27 (50.0)	51 (38.6)
Falling	25 (46.3)	62 (47.3)
Dysphagia	24 (44.4)	61 (46.2)
Cuticular overgrowth	22 (40.7)	41 (31.3)
Fever	20 (35.7)	42 (31.8)
Arthritis	17 (30.4)	58 (43.9)
V-sign rash	17 (31.5)	38 (29.0)
Asymmetric weakness	12 (22.2)	21 (16.2)
Raynaud's Phenomenon	12 (22.2)	18 (13.6)
Shawl-sign rash	9 (16.7)	25 (19.1)
Mechanic's hands	7 (13.0)	13 (9.9)
Palpitations	4 (7.4)	12 (9.2)
Carpal Tunnel Syndrome	4 (7.4)	7 (5.3)
Interstitial lung disease	2 (3.8)	8 (6.1)

#MAV, myositis symptoms developing within 6 months of documented vaccination; non-MAV, no documented immunization within 6 months of onset of myositis; JDM, juvenile dermatomyositis; DM, dermatomyositis; PM, polymyositis; CTM, connective tissue disease overlap with myositis; IBM, inclusion body myositis; JPM, juvenile polymyositis; MSA, myositis-specific autoantibody; MAA, myositis-associated autoantibody; p155 (TIF1), anti-transcription intermediary factor 1 autoantibodies; MJ (NXP2), anti-nuclear matrix protein autoantibodies 2; SRP, anti-signal recognition particle autoantibodies; Ro60, autoantibodies to the 60kD protein of the heterogeneous antigenic complex; PM-Scl, autoantibodies to the 75kD and 100kD proteins seen in the polymyositis/scleroderma complex; U1RNP, autoantibodies to the U1 ribonucleoprotein complex.
+ Sum is > 100%, as some patients have both MSA and MAA.
*No significant differences were detected between the MAV and Non-MAV groups.

Tetanus, influenza, and MMR had a similar period from vaccination to first myositis symptom. However, for those who received Hepatitis B vaccine, there was a significantly shorter latency period, with a median of 1 month from vaccination to first myositis symptom ($p = 0.045$). In the cases where vaccines were given in a series, there was a median of 2.2 months from the time of first vaccination to first myositis symptom, a median of 3.0 months after the second vaccine, and a median of 3.5 months after the third vaccine.

In total, 98 vaccines were administered to the 56 patients (Table 2). Sixteen patients received multiple vaccines on different days within the 6-month period and nine patients received 2 or 3 doses of Hepatitis B vaccine. Among 16 Hepatitis B patients, five developed MAV after the 1st dose, five developed MAV after the 2nddose, and six developed MAV after the 3rd dose.

3.2 HLA analysis

The frequency of DQA1*02:01 was significantly higher in the MAV group compared to non-MAV (OR = 3.80, 95% CI = 1.36-10.58, $p = 0.007$), however, it was protective for non-MAV versus HC (OR = 0.25, 95% CI = 0.11-0.55, $p = 0.0004$) (Table 3). The frequency of DRB1*03:01 was significantly lower for MAV compared to non-MAV (OR = 0.41, 95% CI = 0.18-0.94, $p = 0.033$) but it was a risk factor for the non-MAV versus HC (OR = 3.42, 95% CI = 2.14-5.48, $p < 0.0001$), but not for MAV vs. HC. DRB1*15 was a protective factor for the non-MAV group compared to HC (OR = 0.44, 95% CI = 0.22-0.88, $p = 0.017$). DQA1*05 was a risk factor for the non-MAV group (OR = 2.25, 95% CI = 1.40-3.45, $p = 0.004$). Adult and juvenile data were similar in the overall HLA analysis and showed no significant differences.

Several risk and protective alleles for the non-MAV group were also shared by the MAV group, including DRB1*10:01 (OR = 6.29, 95% CI = 1.78-22.20, $p = 0.001$) and DQA1*03:01 (OR = 3.43, 95% CI = 1.92-6.13, $p < 0.0001$) as risk factors. DRB1*02 (OR = 0.06, 95% CI = 0.01-0.25, $p < 0.0008$) was a protective factor for the non-MAV and MAV groups (Table 3). Homozygosity of HLA alleles did not show a significant impact for either risk or protective factors for the MAV or non-MAV groups.

Several HLA alleles demonstrated significant associations in the MAV versus HC groups (Table 3). The DRB1*10:01 allele was significantly associated with MAV (OR = 8.95, 95% CI = 2.05-39.00, $p = 0.012$) compared to HC. The DQA1 03:01 allele (OR = 4.23, 95% CI = 1.92-9.32, $p = 0.007$) and DQA1*03:03 (OR = 3.86, 95% CI = 1.56-9.54, $p = 0.002$) were also risk factors for MAV when compared to HC. HLA DQA1*03:03 was the only unique risk factor allele for MAV that was not also a risk for the non-MAV group when compared to HC (Table 3). However, the frequency of DRB1*02 (OR = 0.03, 95% CI = 0.01-0.46, $p < 0.0001$) was lower in MAV, indicating a lower likelihood of MAV in individuals with this allele. A sensitivity analysis of HLA alleles of MAV cases developing within three months of vaccination resulted in the same findings.

The frequencies of the linked alleles DQA1*02:01 and DRB1*07:01 were significantly higher in the MAV group receiving the Hepatitis B or influenza vaccines compared to non-

TABLE 2 Distribution of the number of patients receiving vaccines and the number of vaccines administered prior to first symptoms in 56 patients who developed myositis within 6 months after vaccination#.

Vaccine	Patients receiving a vaccine within 6 months of onset (% of all 56 patients)+	Vaccinations administered within 6 months of onset (% of all 98 vaccinations)
Any Tetanus (DPT, DTaP, or Td)	17 (30.4)	18 (18.4)
Hepatitis B	15 (26.8)	27 (27.6)
Influenza A/B	15 (26.8)	15 (15.3)
MMR or Measles	13 (23.2)	13 (13.3)
OPV or IPV	6 (10.7)	6 (6.1)
Prevnar or Pneumococcal	3 (5.4)	3 (3.1)
Hepatitis A, Hemophilus influenzae type B, Varicella, Meningococcal, Typhoid, or Yellow fever *	12 (21.4)	12 (12.2)
Rabies, Japanese Encephalitis, Influenza A virus subtype H1N1, or Lyme **	4 (7.2)	4 (4.0)

*2 patients each received one of these 6 vaccines, and 2 vaccinations were administered for each vaccine listed.
**1 patient each received one of these 4 vaccines, and 1 vaccination was administered for each vaccine listed.
DPT, diphtheria pertussis tetanus vaccine; DTaP, diphtheria tetanus acellular pertussis vaccine; Td, tetanus booster; MMR, measles mumps rubella vaccine; OPV, oral polio vaccine; IPV, inactivated polio vaccine.
+Ten patients received more than 1 vaccine at the same time and the combinations of vaccines given within 6 months of developing myositis were: Patient 1 - 1st HepB, influenza; Patient 2 - 1st HepB, MMR; Patient 3 – OPV, MMR; Patient 4 - DTP, Haemophilus influenzae type B, 3rd HepB, pneumococcal conjugate vaccine; Patient 5 – DTaP, HIB, 3rd HepB; Patient 6 – DTaP, IPV, MMR; Patient 7 – DTaP, OPV; Patient 8 – Td, MMR; Patient 9 – Varicella, MMR; and Patient 10 – Td, HepA.

10.3389/fimmu.2025.1539659

TABLE 3 Differences in HLA types in Caucasian myositis after vaccination (MAV) patients, non-MAV patients, and healthy controls*.

HLA Alleles	MAV % (n=48)	Non-MAV % (n=93)	Control % (n=527)	MAV vs. Non-MAV		MAV vs. Healthy Control		Non-MAV vs. Healthy Control	
				P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)
HLA-DRB1									
*02	0.0	2.2	26.8	0.543	1.05 (0.09-11.93)	<0.0001	0.03 (0.01-0.46)	<0.0008	0.06 (0.01-0.25)
*03:01	26.3	45.6	20.3	0.033	0.41 (0.18-0.94)	0.498	0.71 (0.33-1.51)	<0.0001	3.42 (2.14-5.48)
*07:01	26.3	12.5	23.9	0.099	0.40 (0.15-1.04)	0.888	0.88 (0.42-1.86)	0.173	0.45 (0.23-0.88)
*10:01	7.9	5.7	0.9	0.697	0.70 (0.16-3.10)	0.0125	8.95 (2.05-39.0)	0.0012	6.29 (1.78-22.20)
*15	12.5	10.8	21.4	0.976	0.84 (0.28-2.47)	0.202	1.91 (0.79-4.61)	0.0172	0.44 (0.22-0.88)
HLA-DQA1									
*02:01	23.4	7.4	24.3	0.007	3.80 (1.36-10.58)	0.920	1.04 (0.51-2.11)	0.0004	0.25 (0.11-0.55)
*03:01	19.1	22.3	7.7	0.823	1.21 (0.51-2.91)	0.007	4.23 (1.92-9.32)	<0.0001	3.43 (1.92-6.13)
*03:03	14.9	7.4	4.3	0.231	0.46 (0.15-1.40)	0.002	3.86 (1.56-9.54)	0.293	0.56 (0.23-1.35)
*05	47.9	62.1	42.6	0.105	1.78 (0.88-3.59)	0.544	0.81 (0.45-1.46)	0.004	2.2 (1.40-3.45)

*Carriage rates were determined by the number of allele-positive subjects divided by the number of subjects for whom complete HLA data were available at a given locus. Abbreviations per prior tables.

MAV (Table 4). The DQA1*03:03 allele was a risk factor for MAV patients who received influenza vaccines compared to HC (Table 4).

3.3 GM/KM analysis

The GM phenotype 1, 2, 3, 5, 13, 17, 21, 23 and allotypes GM 2 (OR = 3.17, 95% CI = 1.24-8.13, p = 0.012) and GM13 (OR = 12.5, 95% CI = 1.64-95.05, p = 0.001) were risk factors for MAV compared to HC, but were not risk factors for the non-MAV group (Table 5). KM1 (OR = 3.43, 95% CI = 1.30-9.03, p = 0.009), and KM1,3 (OR = 5.19, 95% CI = 1.47-18.29, p = 0.008) were also risk factors for MAV.

The allotypes GM 2 (OR = 3.61, 95% CI = 1.09-11.99, p = 0.0319), KM 1 (OR 5.57, 95% CI = 1.64-18.94, p = 0.004), and the phenotype KM1,3 (OR 5.19, 95% CI = 1.47-18.29, p = 0.0078) were risk factors for MAV compared to Non-MAV. Because the JDM subgroup was more frequent in the non-MAV than MAV groups, we performed a sensitivity analysis with the MAV group that received their last vaccination within three months and selecting a random sample of JDM patients to create a similar proportion of myositis clinical subgroups in the non-MAV group as in the MAV group in the three month window. In this analysis, the MAV group's GM/KM risk alleles remained unchanged comparing the MAV and non-MAV groups.

4 Discussion

Gene-environment interactions appear to play an important role in the development of autoimmune diseases (28). Immunogenetic factors are critical for immune responses to vaccines and have been proposed to modulate risk for the development of vaccine adverse reactions (21). This study suggests possible genetic associations with the development of myositis after vaccinations. HLA alleles have been associated with the development of many autoimmune diseases, including multiple sclerosis, systemic lupus erythematosus, type 1 diabetes mellitus, Sjogren disease and IIM (29–34), as well as possible risk factors for some vaccine adverse events (21).

Our study identified HLA-DQA1*03:03 as a unique risk factor for MAV versus HC, as this allele is not known to be associated with any other IIM groups. This unique risk factor for MAV suggests a different immune response pathway leading to myositis after vaccinations. Interestingly, the known myositis risk factor DRB1*03:01 was present in lower frequency in the MAV group compared to non-MAV group.

The frequency of HLA-DQA1*02:01, a known risk factor in Caucasians for anti-Mi-2 autoantibodies, was significantly higher in patients with MAV, particularly after the Hepatitis B and influenza vaccines, compared to non-MAV, but no association of MAV was seen with anti-Mi-2 autoantibodies. However, HLA-DQA1*02:01

TABLE 4 Differences in HLA types in Caucasian myositis after vaccination (MAV) patients, non-MAV patients, and healthy controls by vaccine types*.

Vaccine	HLA Alleles	MAV vs. Non-MAV		MAV vs. Control	
		P-value	OR (95% CI)	P-value	OR (95% CI)
Hepatitis B (n=7)	DRB1*07:01	0.006	14.00 (2.84-76.39)	0.018	7.92 (1.81-41.83)
	DQA1*02:01	0.002	16.57 (3.63-71.83)	0.037	5.23 (1.38-20.89)
Influenza (n=14)	DQA1*01	0.038	0.28 (0.01-0.93)	0.021	0.27 (0.10-0.79)
	DQA1*02:01	0.001	7.77 (2.05-26.21)	0.323	1.96 (0.71-6.26)
	DQA1*03:01	0.497	1.54 (0.20-2.06)	0.017	5.30 (1.73-17.83))
	DQA1*03:03	0.102	3.72 (0.92-14.26)	0.020	6.61 (1.84-25.68))
Tetanus (n=10)	DRB1*16	0.030	7.46 (1.64-36.91)	0.011	8.98 (2.38-35.76)

*Carriage rates were determined by the number of allele-positive subjects divided by the number of subjects for whom complete HLA data were available at a given locus; MAV patients in each group were compared to 93 non-MAV and 527 controls.

appeared to be a protective factor for the non-MAV group compared to the HC. Although DRB1*07 had previously been described to be associated with myositis in certain racial populations (1), we found this allele to be significantly more frequent in Caucasians with MAV after Hepatitis vaccines compared to the non-MAV group. These findings highlight the complex gene-environment interactions involved in MAV and suggest potential areas for future research and interventions.

The results of comparing both MAV and non-MAV to HC revealed significant associations between specific HLA alleles and risk of myositis, showing further alleles of interest in the immunogenetic profiles of these patients. DRB1*10:01 and DQA1*03:01 were linked to an elevated risk of MAV, indicating a genetic predisposition to myositis following immunization. The protective association with DRB1*02 suggests a reduced likelihood of developing myositis in carriers of this allele, potentially due to its

TABLE 5 Differences in GM/KM allotypes and phenotypes in Caucasian myositis after vaccination (MAV), Non-MAV, and control groups*.

GM/ KM Markers	MAV % (n=19)	Non-MAV % (n=34)	Control % (n=266)	MAV vs. Non- MAV		MAV vs. Control		Non-MAV vs. Control	
				P-values	OR (95% CI)	P-values	OR (95% CI)	P-values	OR (95% CI)
Allotypes									
GM 2	52.6	23.5	25.9	0.0319	3.61 (1.09-11.99)	0.012	3.17 (1.24-8.13)	0.7642	1.14 (0.49-2.63)
GM 13	94.7	73.5	59.0	0.0756	0.15 (0.02-1.33)	0.0012	12.5 (1.64-95.05)	0.1483	0.52 (0.23-1.15)
KM 1	62.3	23.5	33.3	0.0043	5.57 (1.64-18.94)	0.0087	3.43 (1.30-9.03)	0.3173	1.53 (0.66-3.51)
Phenotypes									
GM 1, 2, 3, 5, 13, 17, 21, 21, 23	26.3	8.8	3.4	0.1181	0.27 (0.06-1.30)	0.001	10.2 (3.01-34.50)	0.1434	0.36 (0.09-1.41)
KM 1, 3	52.6	17.6	25.9	0.0078	5.19 (1.47-18.29)	0.012	3.17 (91.23-8.13)	0.3994	1.63 (0.65-4.11)
KM 3, 3	42.1	76.5	63.9	0.0124	0.22 (0.07-0.75)	0.0984	2.43 (0.95-6.26)	0.2087	0.54 (0.23-1.25)

*Conventions and abbreviations per prior Tables.

role in modulating immune responses. Previous literature has not elucidated any association of these alleles with myositis, warranting further investigations.

Immunoglobulin genes are important risk and protective factors for many autoimmune diseases, and GM13, KM1 and KM3 allotypes have been described as risk factors for myositis (9, 10). The GM/KM analysis identified GM2 and GM13 as risk factors for MAV compared to HC, but not for the non-MAV group. Similarly, KM1 and the KM1,3 phenotype were also linked to increased MAV risk. These findings suggest that specific GM/KM allotypes may serve as additional non-HLA genetic markers for MAV risk, warranting further research into their potential for personalized risk assessment.

Among the 56 MAV patients, there was a median interval of 2.2 months from vaccination to the first myositis symptom. Previous case reports showed the interval between vaccination and the development of symptoms of myositis ranged from 24 hours to 2 months, which generally aligns with our observations (13, 35–37). It has been postulated that when patients develop myositis after repeated vaccine exposure, it is likely due to an amplified immune response triggered by the repeated doses. While our data showed a delayed onset of myositis symptoms following the influenza vaccine, this contrasts with previous case reports that reported a shorter latency period of less than a month after receiving the influenza vaccine (36, 38). As there was a significantly shorter latency period, with a median of one month from vaccination to first myositis symptom for those developing MAV after Hepatitis B vaccine ($p = 0.045$), it is possible that a different mechanism of immune activation may be at work in these cases.

Our study has several limitations. First, our cohort was relatively small and was collected before the onset of the COVID-19 pandemic, and as a result, it does not include patients who developed myositis after receiving COVID-19 vaccinations. This a notable limitation, particularly in light of numerous case reports that have been published during and after the pandemic documenting the onset of autoimmune diseases, including myositis and specifically anti-melanoma differentiation-associated protein 5 (MDA5) autoantibody-positive DM following COVID-19 vaccination (5, 39–41). Other recently approved vaccines, including those to rotavirus, human papillomavirus, and herpes zoster were also not included in our study. It is interesting that so many different vaccine antigens might be associated with myositis, which suggests a single mechanistic explanation is not likely, and also raises the question of the role of the various adjuvants used in these many vaccines. However, given the small numbers of cases and variations in adjuvants from vaccine to vaccine, from manufacturer to manufacturer, and over time, it was not possible to carefully evaluate this. Furthermore, our investigation did not include certain recently identified myositis autoantibodies, including anti-MDA5, and did not include the most recent genotyping methods. And some non-MAV cases may have received vaccinations that were not recalled or documented, potentially biasing the comparisons. Nevertheless, our study lays the groundwork for future research on MAV. We hope that future research will build on this foundation, incorporating more recent methods and including all vaccines and phenotypes of myositis to provide a more comprehensive understanding of MAV.

5 Conclusion

Our study highlights the complex relationship between vaccinations and the onset of myositis. Our findings are generally consistent with previous studies and reports of MAV, although our data showed a somewhat more delayed onset of myositis symptoms after vaccination, particularly following the influenza vaccine. The novel identification of the HLA-DQA1*03:03 allele as a unique risk factor for MAV and the protective factor of HLA-DRB1*03:01 suggests the role of a genetic predisposition in the MAV group that differs from non-MAV myositis patients. GM/KM associations and other HLA genes were noted among specific vaccines and MAV. These genetic associations could provide insights into the pathogenesis of myositis, suggesting that specific gene-environment interactions may influence the susceptibility of developing MAV. Studies in larger populations exploring greater numbers of deeply clinically, immunologically, and genetically phenotyped subjects, and including all currently available vaccines, are needed to understand possible associations among vaccines and myositis and the genetic risk and protective factors involved. A larger study population would also be instrumental in determining the possible epistatic or interactive effects of HLA, GM, and KM alleles on MAV.

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 humans were approved by National Institutes of Health Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants or their legal guardians/next of kin.

Author contributions

EA: Data curation, Writing – original draft, Writing – review & editing. AP: Data curation, Writing – original draft, Writing – review & editing. ES: Data curation, Resources, Writing – review & editing, Conceptualization. JP: Conceptualization, Investigation, Methodology, Resources, Writing – review & editing, Data curation. LR: Data curation, Conceptualization, Investigation, Project administration, Resources, Supervision, Writing – review & editing, Funding acquisition. FM: Conceptualization, Funding

acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was supported in part by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences.

Acknowledgments

The authors are indebted to Drs. Terrance O'Hanlon for laboratory assistance and Ira Targoff for autoantibody identification, and the NIH HLA laboratory for genotyping assistance.

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. The authors declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

1. Miller FW, Lamb JA, Schmidt J, Nagaraju K. Risk factors and disease mechanisms in myositis. *Nat Rev Rheumatol.* (2018) 14:255–68. doi: 10.1038/nrrheum.2018.48

2. Caliskan M, Brown CD, Maranville JC. A catalog of GWAS fine-mapping efforts in autoimmune disease. *Am J Hum Genet.* (2021) 108:549–63. doi: 10.1016/j.ajhg.2021.03.009

3. Miller FW. Environment, lifestyles, and climate change: the many nongenetic contributors to the long and winding road to autoimmune diseases. *Arthritis Care Res (Hoboken).* (2024) 77(1):3–11. doi: 10.1002/acr.25423

4. Toussirot É, Bereau M. Vaccination and induction of autoimmune diseases. *Inflammation Allergy Drug Targets.* (2015) 14:94–8. doi: 10.2174/1871528114666160105113046

5. Shumnalieva R, Ravichandran N, Hannah J, Javaid M, Darooka N, Roy D, et al. Characteristics of emerging new autoimmune diseases after COVID-19 vaccination: A sub-study by the COVAD group. *Int J Rheum Dis.* (2024) 27:e15178. doi: 10.1111/1756-185X.15178

6. Ricaño-Ponce I, Wijmenga C. Mapping of immune-mediated disease genes. *Annu Rev Genomics Hum Genet.* (2013) 14:325–53. doi: 10.1146/annurev-genom-091212-153450

7. Rothwell S, Cooper RG, Lundberg IE, Miller FW, Gregersen PK, Bowes J, et al. Dense genotyping of immune-related loci in idiopathic inflammatory myopathies confirms HLA alleles as the strongest genetic risk factor and suggests different genetic background for major clinical subgroups. *Ann Rheum Dis.* (2016) 75:1558–66. doi: 10.1136/annrheumdis-2015-208119

8. Miller FW, Chen W, O'Hanlon TP, Cooper RG, Vencovsky J, Rider LG, et al. Genome-wide association study identifies HLA 8.1 ancestral haplotype alleles as major genetic risk factors for myositis phenotypes. *Genes Immun.* (2015) 16:470–80. doi: 10.1038/gene.2015.28

9. Dugoujon JM, Cambon-Thomsen A. Immunoglobulin allotypes (GM and KM) and their interactions with HLA antigens in autoimmune diseases: a review. *Autoimmunity.* (1995) 22:245–60. doi: 10.3109/08916939508995322

10. O'Hanlon TP, Rider LG, Schiffenbauer A, Targoff IN, Malley K, Pandey JP, et al. Immunoglobulin gene polymorphisms are susceptibility factors in clinical and autoantibody subgroups of the idiopathic inflammatory myopathies. *Arthritis Rheumatol.* (2008) 58:3239–46. doi: 10.1002/art.23899

11. Ehrengut W. Dermatomyositis and vaccination. *Lancet.* (1978) 1:1040–1. doi: 10.1016/S0140-6736(78)90761-4

12. Cotterill JA, Shapiro H. Dermatomyositis after immunisation. *Lancet.* (1978) 2:1158–9. doi: 10.1016/S0140-6736(78)92325-5

13. Kåss E, Straume S, Munthe E. Dermatomyositis after B.C.G. vaccination. *Lancet.* (1978) 1:772. doi: 10.1016/S0140-6736(78)90885-1

14. Kåss E, Straume S, Mellbye OJ, Munthe E, Solheim BG. Dermatomyositis associated with BCG vaccination. *Scand J Rheumatol.* (1979) 8:187–91. doi: 10.3109/03009747909114454

15. Gotoff SP, Smith RD, Sugar O. Dermatomyositis with cerebral vasculitis in a patient with agammaglobulinemia. *Am J Dis Child.* (1972) 123:53–6. doi: 10.1001/archpedi.1972.02110070103015

16. Thieffry S, Arthuis M, Martin C, Sorrel-Dejerine J, Benhamida M. Infantile dermatomyositis. Study of 8 personal cases. *Ann Pediatr (Paris).* (1967) 14:554–78.

17. Bitnum S, Daeschner CW Jr., Travis LB, Dodge WF, Hopps HC. DERMATOMYOSITIS. *J Pediatr.* (1964) 64:101–31. doi: 10.1016/S0022-3476(64)80325-5

18. Gherardi RK, Authier FJ. Macrophagic myofasciitis: characterization and pathophysiology. *Lupus.* (2012) 21:184–9. doi: 10.1177/0961203311429557

19. Deshpande P, Hertzman RJ, Palubinsky AM, Giles JB, Karnes JH, Gibson A, et al. Immunopharmacogenomics: mechanisms of HLA-associated drug reactions. *Clin Pharmacol Ther.* (2021) 110:607–15. doi: 10.1002/cpt.2343

20. O'Hanlon T, Koneru B, Bayat E, Love L, Targoff I, Malley J, et al. Immunogenetic differences between Caucasian women with and those without silicone implants in whom myositis develops. *Arthritis Rheumatol.* (2004) 50:3646–50. doi: 10.1002/art.20587

21. Bandinelli F, Pagano M, Vallecoccia MS. Post-COVID-19 and post-COVID-19 vaccine arthritis, polymyalgia rheumatica and horton's arteritis: A single-center assessment of clinical, serological, genetic, and ultrasonographic biomarkers. *J Clin Med.* (2023) 12. doi: 10.3390/jcm12247563

22. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). *N Engl J Med.* (1975) 292:344–7. doi: 10.1056/NEJM197502132920706

23. Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). *N Engl J Med.* (1975) 292:403–7. doi: 10.1056/NEJM197502202920807

24. O'Hanlon TP, Carrick DM, Targoff IN, Arnett FC, Reveille JD, Carrington M, et al. Immunogenetic risk and protective factors for the idiopathic inflammatory myopathies: distinct HLA-A, -B, -Cw, -DRB1, and -DQA1 allelic profiles distinguish European American patients with different myositis autoantibodies. *Med (Baltimore)*. (2006) 85:111–27. doi: 10.1097/01.md.0000217525.82287.eb

25. Vyas GN, Fudenberg HH, Pretty HM, Gold ER. A new rapid method for genetic typing of human immunoglobulins. *J Immunol*. (1968) 100:274–9. doi: 10.4049/jimmunol.100.2.274

26. Targoff IN, Mamyrova G, Trieu EP, Perurena O, Koneru B, O'Hanlon TP, et al. A novel autoantibody to a 155-kd protein is associated with dermatomyositis. *Arthritis Rheumatol*. (2006) 54:3682–9. doi: 10.1002/art.22164

27. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat*. (1979) 6:65–70.

28. Woo JMP, Parks CG, Jacobsen S, Costenbader KH, Bernatsky S. The role of environmental exposures and gene-environment interactions in the etiology of systemic lupus erythematous. *J Intern Med*. (2022) 291:755–78. doi: 10.1111/joim.13448

29. Brassat D, Salemi G, Barcellos LF, McNeill G, Proia P, Hauser SL, et al. The HLA locus and multiple sclerosis in Sicily. *Neurology*. (2005) 64:361–3. doi: 10.1212/01.WNL.0000149765.71212.0A

30. Fernando MM, Stevens CR, Sabeti PC, Walsh EC, McWhinnie AJ, Shah A, et al. Identification of two independent risk factors for lupus within the MHC in United Kingdom families. *PloS Genet*. (2007) 3:e192. doi: 10.1371/journal.pgen.0030192

31. Bronson PG, Komorowski LK, Ramsay PP, May SL, Noble J, Lane JA, et al. Analysis of maternal-offspring HLA compatibility, parent-of-origin effects, and noninherited maternal antigen effects for HLA-DRB1 in systemic lupus erythematosus. *Arthritis Rheumatol*. (2010) 62:1712–7. doi: 10.1002/art.27426

32. Nerup J, Platz P, Andersen OO, Christy M, Lyngsoe J, Poulsen JE, et al. HL-A antigens and diabetes mellitus. *Lancet*. (1974) 2:864–6. doi: 10.1016/S0140-6736(74)91201-X

33. Erlich HA, Griffith RL, Bugawan TL, Ziegler R, Alper C, Eisenbarth G. Implication of specific DQB1 alleles in genetic susceptibility and resistance by identification of IDDM siblings with novel HLA-DQB1 allele and unusual DR2 and DR1 haplotypes. *Diabetes*. (1991) 40:478–81. doi: 10.2337/diab.40.4.478

34. Cruz-Tapias P, Rojas-Villarraga A, Maier-Moore S, Anaya JM. HLA and Sjögren's syndrome susceptibility. A meta-analysis of worldwide studies. *Autoimmun Rev*. (2012) 11:281–7. doi: 10.1016/j.autrev.2011.10.002

35. Hanissian AS, Martinez AJ, Jabbour JT, Duenas DA. Vasculitis and myositis secondary to rubella vaccination. *Arch Neurol*. (1973) 28:202–4. doi: 10.1001/archneur.1973.00490210082014

36. Jani FM, Gray JP, Lanham J. Influenza vaccine and dermatomyositis. *Vaccine*. (1994) 12:1484. doi: 10.1016/0264-410X(94)90166-X

37. Altman A, Szyper-Kravitz M, Shoenfeld Y. HBV vaccine and dermatomyositis: is there an association? *Rheumatol Int*. (2008) 28:609–12. doi: 10.1007/s00296-007-0485-4

38. Ferri C, Colaci M, Manzini CU, Sebastiani M, Giuggioli D, Brugioni L. Polymyositis following pandemic influenza A (H1N1) and 2009–10 seasonal trivalent vaccines. *Case Rep Rheumatol*. (2012) 2012:836930. doi: 10.1155/2012/836930

39. Watson RA, Ye W, Taylor CA, Jungkurth E, Cooper R, Tong O, et al. Severe acute myositis and myocarditis on initiation of 6-weekly pembrolizumab post-COVID-19 mRNA vaccination. *J Immunother Cancer*. (2024) 12. doi: 10.1136/jitc-2023-008151

40. Diaz-Menindez M, Sullivan MM, Wang B, Majithia V, Abril A, Butendieck RR Jr., et al. Dermatomyositis in association with SARS-coV-2 infection or COVID-19 vaccine. *Arthritis Care Res (Hoboken)*. (2024) 76:98–104. doi: 10.1002/acr.25236

41. David P, Sinha S, Iqbal K, De Marco G, Taheri S, McLaren E, et al. MDA5-autoimmunity and interstitial pneumonitis contemporaneous with the COVID-19 pandemic (MIP-C). *EBioMedicine*. (2024) 104:105136. doi: 10.1016/j.ebiom.2024.105136

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 27 August 2025
DOI 10.3389/fimmu.2025.1537043

OPEN ACCESS

EDITED BY
Roberto Paganelli,
Institute for Advanced Biologic Therapies, Italy

REVIEWED BY
Gerson D. Keppeke,
Universidad Católica del Norte, Chile
Andrei Ivashynka,
Moriggia Pelascini Hospital, Italy

*CORRESPONDENCE
Frederick W. Miller
millerf@mail.nih.gov

RECEIVED 29 November 2024
ACCEPTED 04 August 2025
PUBLISHED 27 August 2025

CITATION
Dinse GE, Weinberg CR, Parks CG, Co CA, Priest JT, Chan EKL and Miller FW (2025) Decreased cigarette smoking may partially explain the increased prevalence of antinuclear antibodies in the United States. *Front. Immunol.* 16:1537043. doi: 10.3389/fimmu.2025.1537043

Decreased cigarette smoking may partially explain the increased prevalence of antinuclear antibodies in the United States

Gregg E. Dinse[1], Clarice R. Weinberg[2], Christine G. Parks[3], Caroll A. Co[1], Jessica T. Priest[1], Edward K. L. Chan[4] and Frederick W. Miller[5]*

[1]Public Health & Scientific Research, DLH, LLC, Bethesda, MD, United States, [2]Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, United States, [3]Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, United States, [4]Department of Oral Biology, University of Florida Health Science Center, Gainesville, FL, United States, [5]Clinical Research Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, United States

Introduction: Despite well-known harmful health effects of smoking, research supports an inverse association with some autoimmune diseases. High-titer antinuclear antibodies (ANA) are associated with autoimmune diseases, and ANA prevalence in the US increased between 1988 and 2012. Tobacco smoking decreased during those years while vaping of electronic cigarettes (e-cigarettes) increased after their introduction in 2007. Carbon monoxide (CO) may ameliorate autoimmunity, and e-cigarettes deliver much less CO than regular cigarettes. We explored interdependencies among ANA, smoking, and time.

Methods: We analyzed cross-sectional data on ANA and the primary nicotine metabolite, cotinine, in 13,288 participants ≥12 years old from three time periods (1988-1991, 1999-2004, 2011-2012) of the US National Health and Nutrition Examination Survey. Smoking exposure (none, passive, active) was inferred from serum cotinine. We used logistic regression to analyze ANA prevalence, adjusted for sex, age, and race/ethnicity.

Results: Over the study periods, ANA prevalence was highest (13.3-19.2%) for nonsmokers but non-trending; lower (11.1-15.5%) for "passive" smokers but steadily increasing; and even lower for active smokers but increasing from 7.4% in 1999–2004 to 13.3% in 2011-2012. The increases in ANA among passive and active smokers were mainly in adolescents (ages 12–19 years). Smokers had reduced odds of ANA in 1999-2004, with an odds ratio (OR) of 0.65 and a 95% confidence interval (CI) of 0.45-0.93, but this association was weaker in 1988-1991 (OR=0.80; 95% CI:0.52-1.22) and 2011-2012 (OR=0.82; 95% CI:0.56-1.21).

Discussion: Although smoking causes harmful health effects, ANA data are consistent with smoking playing a role in decreasing autoimmunity. Recent vaping among adolescents may partially explain their large increase in ANA prevalence. The inverse ANA association with smoking strengthened between 1988-1991 and 1999-2004 but then weakened by 2011-2012. The initial

10.3389/fimmu.2025.1537043

strengthening was potentially because nonsmokers were exposed to progressively less CO (and/or other components of secondhand smoke), due to tightened smoking restrictions, while the potential nicotine-associated protection against ANA may have weakened after e-cigarettes became a source. Smoking should not be recommended given its negative health impacts. However, further studies could elucidate new mechanisms, perhaps involving components of tobacco smoke or vaping, possibly enabling development of novel preventative or treatment measures.

KEYWORDS

antinuclear antibodies (ANA), autoimmune diseases, carbon monoxide (CO), cotinine, e-cigarettes, National Health and Nutrition Examination Survey (NHANES), tobacco smoking, vaping

1 Introduction

High-titer antinuclear antibodies (ANA) are biomarkers associated with many autoimmune diseases (1–6), some of which have increased in incidence over recent decades for unknown reasons. Previously, based on data from the US National Health and Nutrition Examination Survey (NHANES), we reported an increasing ANA time trend (7) and investigated possible ANA associations with 253 xenobiotics (8). Our initial goal was to explore whether temporal changes in the levels of any xenobiotics associated with ANA could help explain the increase in ANA prevalence over time. However, many xenobiotics were evaluated at only one point in time or had mostly undetectable levels. We ultimately focused on serum cotinine, which was measured in nearly all participants.

Smoking tobacco is a major cause of preventable deaths, illnesses, and health care costs worldwide (9, 10), but despite overwhelming evidence of harmful effects of smoking in general, smoking has appeared to be inversely associated with ANA (7). Cotinine has often been used as a biomarker for tobacco smoke exposure (11–13), and as the primary metabolite of nicotine, cotinine has long been regarded as the most reliable indicator of active and passive exposure to tobacco smoke (11, 14). However, cotinine can also signal other nicotine exposures such as nicotine gum, chewing tobacco, snuff, and snus. Recently, an increasingly popular nicotine-delivering alternative to regular cigarettes, electronic cigarettes (e-cigarettes), has expanded the opportunities for smokeless exposure to nicotine (15).

In this article, we explore whether the decrease in cigarette smoking over the past few decades (16, 17) could plausibly account for some of the increase in ANA. We assess associations seen in the large NHANES database, some of which were observed previously (7), and postulate a potentially protective (or immunosuppressive) effect of carbon monoxide (CO) that might help explain the apparent inverse correlation between cigarette smoking and ANA.

The effects of smoking and CO on autoimmune diseases depend on individual variability, exposure levels, and the disease in question. Perricone et al. (18) discuss numerous studies of the relationship between smoking and autoimmune diseases. While smoking is a risk factor for many autoimmune diseases, smoking appears to have a protective effect for others, including ulcerative colitis, celiac disease, Behcet's disease, type 1 diabetes, and autoimmune hypothyroidism. Epidemiologic studies have suggested that smoking may protect against ulcerative colitis (19–21), Behcet's disease (21), autoimmune hypothyroidism (22–24), and Sjogren's syndrome (19, 20), and that CO may protect against discoid lupus erythematosus (25). Rodent studies have suggested that CO may have therapeutic effects for various autoimmune diseases, including multiple sclerosis (26, 27), collagen-induced arthritis (28), systemic lupus erythematosus (29), type 1 diabetes (30), uveitis (31), and autoimmune hepatitis (32).

Starting early this century, many smokers began using e-cigarettes, either in addition to or instead of regular cigarettes (15, 33–36). Among 116 adult e-cigarette users in one study (37), 68% self-reported as current smokers, 24% as former smokers, and 8% as never smokers. The use of e-cigarettes doubled between 2010 and 2013 among US adults, with over 20 million having tried them (34), and the use among high school students increased from 1.5% in 2011 to 16% in 2015 (15). E-cigarettes supply nicotine and thus cotinine (15, 38) but produce much less CO than regular cigarettes (39, 40). We hypothesize that a decrease in CO and/or other possibly "protective" smoking byproducts, either due to quitting all forms of smoking or switching fully or partially from regular cigarettes to e-cigarettes (or other nicotine delivery systems, such as chewing tobacco, snuff, or snus), may have contributed to the recent increase in ANA in the US.

2 Subjects and methods

2.1 Study participants

Data on ANA were available for 13,519 participants from five NHANES cycles: 1988-1991, 1999-2000, 2001-2002, 2003-2004, and 2011-2012. The NHANES sampled representative members

of the noninstitutionalized civilian US population and provided sampling weights to adjust for selection probabilities and nonresponse (41), which enables inference that generalizes to most of the US population. All participants signed informed consent documents and completed questionnaires, and most were physically examined and provided blood and urine specimens. Available data included demographic characteristics, health covariates, measured factors, and constructed variables. The NHANES protocol was approved by the Human Subjects Institutional Review Board of the US Centers for Disease Control and Prevention (CDC).

2.2 ANA assessment

All serum samples were evaluated for ANA in the laboratory of Dr. Edward K.L. Chan between 2016 and 2017 by indirect immunofluorescence at a 1:80 dilution using the NOVA Lite HEp-2 ANA slide with DAPI kit (Inova Diagnostics, San Diego, California, USA), with a highly specific fluorescein isothiocyanate-conjugated secondary antibody (goat anti-human IgG). Immunofluorescence staining intensities were graded 0–4 compared to standard references (42). Grades 1–4 were considered positive for ANA and grade 0 was considered negative. For more assay details see Dinse et al. (7).

2.3 Data on ANA, cotinine, and smoking

For cost and other practical reasons, ANA were only assayed in a subset of participants ≥12 years old in each of the five cycles. The ANA subsamples from 1999-2000, 2001-2002, and 2003–2004 were each roughly one-third the size of those from 1988–1991 and 2011-2012. Thus, as in our earlier studies (7, 8), we combined the three middle cycles to create three time periods with similar sample sizes: 1988-1991 (N=4,727), 1999-2004 (N=4,527), and 2011-2012 (N=4,265). As before, we focused on these three periods rather than the five cycles.

All analyses were restricted to the 13,519 participants with ANA data. The CDC adjusted the sampling weights to account for analyzing this ANA subsample. Data were available on cotinine, and thus smoking exposure as defined by cotinine concentration, for 13,288 participants; on self-reported smoking history for 12,278 participants; and on both smoking exposure and smoking history for 12,063 participants. Supplementary Table S1 shows the numbers of participants in each time period (and overall) with data on ANA, smoking exposure, and smoking history. Throughout this article, "cotinine" refers to serum cotinine and not urinary cotinine.

2.4 Model variables

The ANA outcome variable was a binary indicator of ANA positivity/negativity. Cotinine concentration (ng/mL) was a quantitative variable and was used to classify smoking exposure as none (≤0.05), passive (>0.05 to 10), or active (>10), as recommended by the CDC and the US Environmental Protection Agency (EPA) (13), though a sensitivity analysis applied a more recent recommendation of >3 ng/mL for defining active smoking exposure. Combining the first two exposure categories produced an indicator of smoking status: nonsmoker (none or passive exposure) versus smoker (active exposure). Smoking history was based on questionnaire data, with individuals self-reporting as never, former, or current smokers.

Except where otherwise noted, our primary analyses adjusted for sex, age, race/ethnicity, and the survey design variables (i.e., strata, clusters, and weights proportional to the inverse probability of sampling), each of which was available for all participants. Age was measured in years and categorized by decade (12-19, 20-29,…, 70-79, ≥80), though sensitivity analyses explored the use of fewer age categories, a quantitative age variable, or a restricted cubic spline in age. Self-reported race/ethnicity was categorized as non-Hispanic White, non-Hispanic Black, Mexican American, or Other. Secondary analyses adjusted for body mass index (BMI), alcohol intake, poverty income ratio (PIR), and education, as defined previously (42). Secondary analyses also investigated CO content in cigarettes, pack-years of smoking, lifetime years of smoking, and years since former smokers quit smoking, though these data were very limited.

2.5 Statistical analysis

When analyzing ANA prevalence, we used logistic regression models to allow the probability of ANA positivity to depend on explanatory variables. All models adjusted for the survey design variables. The basic model for estimating overall ANA prevalence and its 95% confidence interval (CI) did not include adjustment covariates, but we did include a categorical covariate for period when estimating ANA prevalence in each of the three time periods. When assessing ANA time trends, we adjusted for sex, age, and race/ethnicity and calculated an ANA prevalence odds ratio (OR) and its 95% CI for each period relative to the first period. The statistical significance of an ANA time trend was evaluated by replacing the categorical period covariate with a quantitative time covariate and then inspecting its p-value, where time was defined as the number of years between the midpoints of the participant's period and the first period.

When analyzing the cotinine data, we calculated the geometric mean cotinine concentration for each time period. We also derived a trend line by using linear regression to model individual log-transformed cotinine concentration as a function of the number of years between the midpoints of the participant's period and the first period. Any concentration below the limit of detection (LOD) was replaced by an imputed value of LOD/$\sqrt{2}$ (43, 44). The cotinine LOD was initially 0.05 ng/mL but was lowered to 0.015 ng/mL during the second period due to an improvement in the assay; the corresponding imputed values were 0.035 and 0.011 ng/mL. We also evaluated mean cotinine concentrations over time (and estimated trend lines) within subgroups of self-reported never,

former, and current smokers, and we used kernel density plots to assess the full cotinine concentration distribution for each smoking-history subgroup and time period.

When analyzing smoking time trends, we used logistic regression to estimate the prevalence of smokers in each time period. Overall prevalence estimates were adjusted for the survey design variables but not for any covariates. Also, after further adjusting for sex, age, and race/ethnicity, we estimated a prevalence OR (and a 95% CI) for each period relative to the first period.

When investigating the relationship between ANA and smoking, we performed logistic regression analyses similar to those described above for ANA prevalence. First, we stratified by smoking and analyzed ANA prevalence and time trends separately in each stratum. Second, we stratified by both age and smoking to see whether the ANA association with smoking depended on age. Third, we added a smoking covariate (instead of stratifying) and assessed whether that smoking covariate affected the ANA association with time or whether removing the period covariate altered the ANA association with smoking. Fourth, we also added a smoking-by-period interaction to evaluate whether smoking modified the ANA time trend. Fifth, we stratified by period and compared ANA prevalence for smokers versus nonsmokers to gauge how the ANA association with smoking changed over time.

Finally, we conducted sensitivity analyses to assess whether our results changed when using an alternative age covariate (fewer categories, quantitative, or restricted cubic spline) or when only considering adults (ages ≥20 years). We also explored the use of a more recent recommendation of >3 ng/mL for the cotinine cutpoint when defining active smoking exposure. In addition, we investigated several other covariates (BMI, alcohol intake, PIR, and education) and the limited data on cigarette CO content, pack-years, years of smoking, and years since quitting.

All analyses were performed with SAS software (version 9.4, SAS Institute, Cary, NC) and accounted for the survey design variables by using special survey procedures. Domain statements were used to properly handle the sampling weights in subgroup analyses. Variance estimates for the 95% CIs were obtained using the Taylor series method. Reported p-values were 2-sided. All plots were constructed in SAS except the kernel density plot, which was created in R (version 4.4.0, R Foundation, Vienna, Austria).

3 Results

3.1 ANA time trend

The prevalence of ANA rose over the 25-year span for which NHANES data on ANA were available, with most of the increase occurring between the second and third time periods. Accounting only for time period and the survey design variables, the weighted estimates of ANA prevalence were 11.0% (95% CI: 9.7-12.5%) in Period 1 (1988–1991), 11.4% (95% CI: 10.2-12.8%) in Period 2 (1999-2004), and 16.1% (95% CI: 14.5-17.9%) in Period 3 (2011-2012). These overall estimates, along with sample sizes and numbers of ANA-positive participants, are shown in the last row of Table 1. Relative to Period 1 and after further adjustment for sex, age, and race/ethnicity, the ANA prevalence OR was 1.02 (95% CI: 0.84-1.24) for Period 2 and 1.49 (95% CI: 1.23-1.82) for Period 3 (Table 2), and there was strong statistical evidence of a positive trend in ANA prevalence over time (p=0.0001). We reported these results earlier (7), with slight discrepancies due to minor differences in analysis, but repeat them here for context.

3.2 Cotinine time trend

Supplementary Figure S1 shows the geometric mean cotinine concentration and its 95% CI for each period, along with the best-fitting trend line. There was strong statistical evidence (p<0.0001) of a steady decrease over time. When stratified by self-reported smoking history, the mean cotinine levels ranged from 0.04 to 0.27 ng/mL for never smokers, 0.08 to 0.59 ng/mL for former smokers, and 104.2 to 158.5 ng/mL for current smokers (top half of Supplementary Table S2). The best-fitting trend line had a negative

TABLE 1 Sample sizes, ANA-positive counts, and ANA prevalence estimates by time period and smoking exposure.

Smoking	Period 1: 1988-1991		Period 2: 1999-2004		Period 3: 2011-2012		All Periods Combined	
Exposure [a]	N+/N	Prev (95% CI) [b]	N+/N	Prev (95% CI) [b]	N+/N	Prev (95% CI) [b]	N+/N	Prev (95% CI) [b]
None	93/429	19.2 (13.6-26.3)	264/1,884	13.3 (11.3-15.7)	401/2,379	17.4 (14.7-20.4)	758/4,692	16.0 (14.2-18.0)
Passive	343/2,739	11.1 (9.6-12.8)	190/1,581	12.7 (10.5-15.4)	141/1,001	15.5 (13.2-18.1)	674/5,321	12.5 (11.3-13.7)
Active	168/1,357	8.6 (6.4-11.4)	89/1,034	7.4 (5.6- 9.7)	127/884	13.3 (11.0-15.9)	384/3,275	9.5 (8.2-10.9)
Total	643/4,727	11.0 (9.7-12.5)	545/4,527	11.4 (10.2-12.8)	669/4,265	16.1 (14.5-17.9)	1,857/13,519	13.0 (12.1-13.9)

ANA, antinuclear antibodies; CI, confidence interval; LOD, limit of detection; N, total number of participants (sample size); N+, number of ANA-positive participants; Prev, ANA prevalence (as a percent).

[a]Smoking exposure categories were based on serum cotinine concentrations (None, ≤0.05 ng/mL; Passive, >0.05 to 10 ng/mL; and Active, >10 ng/mL). Due to a technical improvement in the cotinine assay, the cotinine LOD decreased from 0.05 to 0.015 ng/mL during Period 2. The number of participants with a missing cotinine value also decreased over time from 202 in Period 1 to 28 in Period 2, and then to 1 in Period 3.

[b]ANA prevalence was estimated under two logistic regression models for ANA positivity (yes/no), adjusted for the survey design variables (sampling weights, strata, and clusters). One model included only an intercept, which produced an overall estimate for all time periods combined. The other model included a categorical covariate for time period, which produced a separate estimate for each period. Both models were applied initially to all participants with data on ANA regardless of data on smoking exposure (Total) and then within subgroups with data on both ANA and smoking exposure (None, Passive, and Active). The subgroup counts sum to less than the total sample size because some participants were missing data on smoking exposure (i.e., cotinine).

TABLE 2 Covariate-adjusted assessments of ANA time trends by smoking exposure.

Smoking	ANA Prevalence Odds Ratio for Time Period (95% CI) [b]			Time Trend
Exposure [a]	Period 1: 1988-1991	Period 2: 1999-2004	Period 3: 2011-2012	p-value [b]
None	1.00 (reference)	0.70 (0.45-1.09)	0.99 (0.62-1.57)	0.2139
Passive	1.00 (reference)	1.28 (0.96-1.69)	1.72 (1.33-2.22)	0.0001
Active	1.00 (reference)	0.81 (0.53-1.23)	1.45 (1.01-2.08)	0.0661
Total	1.00 (reference)	1.02 (0.84-1.24)	1.49 (1.23-1.82)	0.0001

ANA, antinuclear antibodies; CI, confidence interval.
[a]Smoking exposure categories were based on serum cotinine concentrations (None, ≤0.05 ng/mL; Passive, >0.05 to 10 ng/mL; and Active, >10 ng/mL).
[b]ANA time trend assessments were based on two logistic regression models for ANA positivity (yes/no). Each model adjusted for the survey design variables (sampling weights, strata, and clusters) and categorical covariates for sex, age, and race/ethnicity. One model added a categorical covariate for time period, which allowed estimates of the ANA prevalence odds ratio for each period relative to the first period. The other model instead added a continuous covariate for time, as measured by the number of years between period midpoints relative to the first period, and produced a p-value from a t-test to assess a linear ANA time trend. Both models were applied initially to all participants with data on ANA regardless of data on smoking exposure (Total) and then within subgroups with data on both ANA and smoking exposure (None, Passive, and Active).

slope in all three subgroups but was steeper for never and former smokers than for current smokers (Figure 1). Also, the decrease over time was statistically significant for both never and former smokers ($p<0.0001$), but not for current smokers ($p=0.08$). Thus, on average, current smokers had cotinine levels that were high and fairly constant over time, while former and never smokers had levels that were low and decreasing, likely due to steady reductions in secondhand smoke exposure. Similar results were obtained when excluding participants under age 20 years (bottom half of Supplementary Table S2) to account for smoking history data being available for different age ranges across time periods (≥17 years in Period 1, ≥12 years in Period 2, and ≥20 years in Period 3).

Rather than focusing on means, Figure 2 displays kernel density estimates of the entire distribution of cotinine concentrations by time period and smoking history. These plots clearly show the differences in cotinine levels for never and former smokers (low) versus current smokers (high), as well as the consistency over time for current smokers. The cotinine distributions for never and

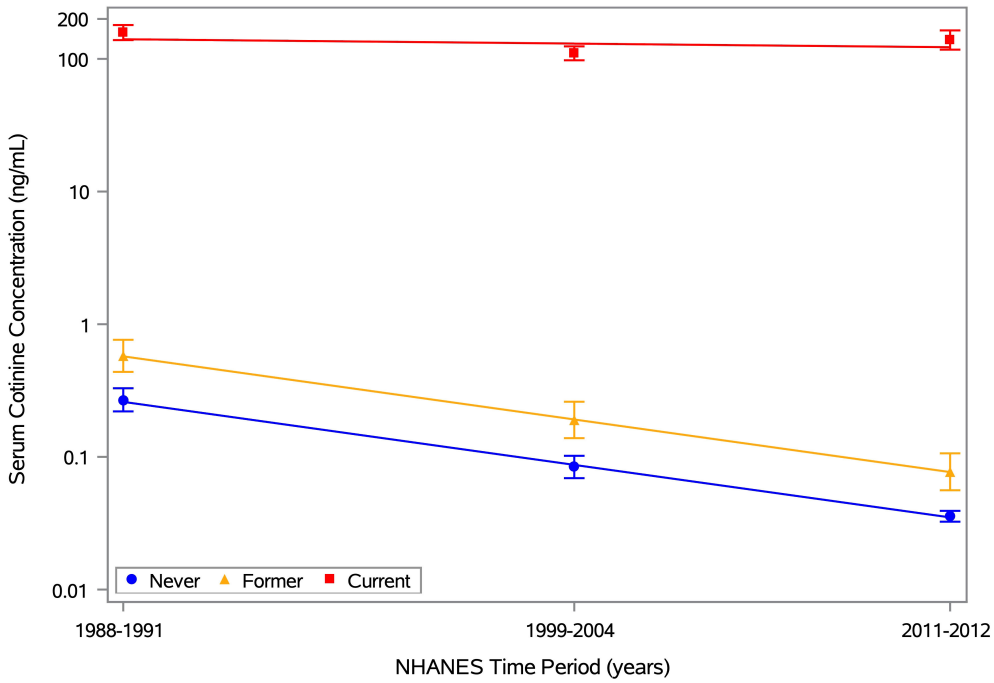


FIGURE 1
Mean serum cotinine concentration by time period and smoking history. Estimates of the geometric mean serum cotinine concentration and its 95% CI are plotted for each of 3 time periods (1988-1991, 1999-2004, and 2011-2012), along with the best-fitting trend line. Separate estimates are shown for self-reported never, former, and current smokers, based on the 12,063 NHANES participants aged ≥12 years with data on ANA, serum cotinine, and smoking history. The means for never, former, and current smokers are depicted by blue circles, yellow triangles, and red squares, respectively, with the same colors used for the 95% CI error bars and trend lines. Any concentration below the limit of detection (LOD) was replaced by an imputed value equal to LOD/ $\sqrt{2}$. The horizontal axis is linear in time, defined as the number of years between the midpoints of the participant's period and the first period, and the vertical axis is logarithmic in serum cotinine concentration (ng/mL).

former smokers were less consistent, with a notable shift toward lower values as time progressed. Much of this shift was likely due to many never and former smokers having cotinine levels below the LOD, which decreased from 0.05 to 0.015 ng/mL in the second time period. Nondetectable levels were replaced by imputed values of 0.035 and 0.011 ng/mL, respectively, which match well with the peaks of the period-specific cotinine distributions for never smokers. The cotinine distributions were more spread out for former smokers than for never smokers, perhaps due to a larger proportion of former smokers interacting with a current smoker.

3.3 Smoking time trend

Cigarette smoking in the US has decreased for a half-century (16, 17). We confirmed this downward trend in the NHANES data by examining the proportions of active smokers (defined by cotinine levels) and current smokers (based on self-reports), both of which clearly decreased over time. Unadjusted period-specific estimates of smoking prevalence for both classifications demonstrated similar decreases among all participants and among adults only (Supplementary Table S3), as did covariate-adjusted estimates of the smoking prevalence ORs for time period (Supplementary Table S4).

3.4 ANA time trends by smoking exposure

Estimates of ANA prevalence exhibited different temporal patterns in the three smoking exposure subgroups. For individuals with no exposure, these estimates were highest but did not show a clear trend; for passive exposure, they were intermediate and increased steadily across all periods from 11.1% (95% CI: 9.6-12.8%) to 12.7% (95% CI: 10.5-15.4%) to 15.5% (95% CI: 13.2-18.1%); and for active exposure, they were lowest and initially flat but then rose markedly from 7.4% (95% CI: 5.6-9.7%) in Period 2 to 13.3% (95% CI: 11.0-15.9%) in Period 3 (Table 1). Covariate-adjusted estimates of the ANA prevalence OR for Period 3 relative to Period 1 were 0.99 (95% CI: 0.62-1.57) for no exposure, 1.72 (95% CI: 1.33-2.22) for passive exposure, and 1.45 (95% CI: 1.01-2.08) for active exposure (Table 2). When assessing a linear trend in ANA prevalence across all three periods, the p-values for the three exposure subgroups were 0.2139, 0.0001, and 0.0661, respectively (Table 2). We reported similar estimates previously (7), but with smoking exposure categories defined by slightly different cutpoints for cotinine concentration.

To investigate whether age modified the association between smoking and temporal patterns of ANA, in addition to stratifying by smoking exposure, we further stratified by three age groups (12-19, 20-49, and ≥50 years) instead of including a categorical covariate for

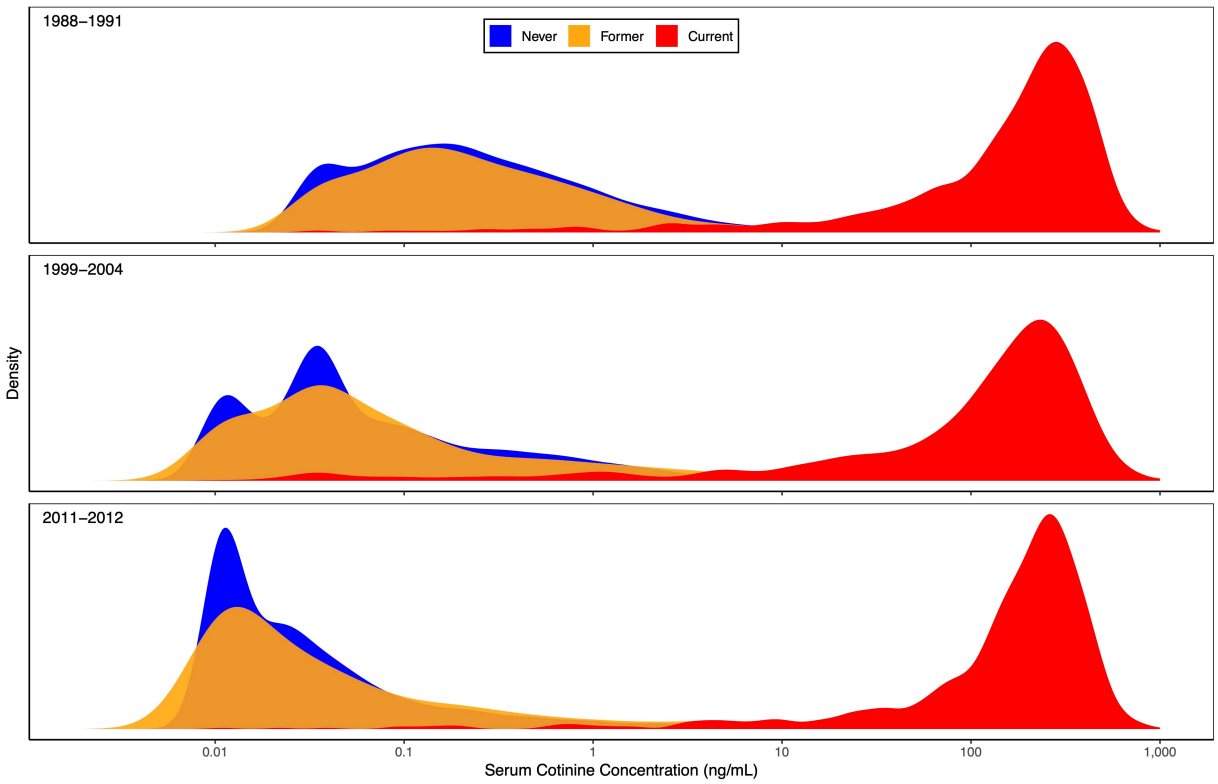


FIGURE 2
Serum cotinine concentration distribution by time period and smoking history. Kernel density estimates of the entire serum cotinine concentration distribution are plotted for each of 3 time periods (1988-1991, 1999-2004, and 2011-2012). Separate estimates are shown for self-reported never (blue), former (yellow), and current (red) smokers, based on the 12,063 NHANES participants aged ≥12 years with data on ANA, serum cotinine, and smoking history. Any concentration below the limit of detection (LOD) was replaced by an imputed value equal to LOD/ $\sqrt{2}$. The horizontal axis is logarithmic in serum cotinine concentration (ng/mL).

age. This approach essentially allowed for interactions between age and the covariates (sex, race/ethnicity, and time period). Despite the larger number of subgroups leading to smaller counts within each, there was statistical evidence that the observed increase in ANA prevalence over time was associated mainly with 12–19 year-olds who were passive (p=0.005) or active (p=0.003) smokers (Table 3). Among adolescents and relative to Period 1, the ORs and 95% CIs for passive smokers were 1.63 (0.83-3.23) in Period 2 and 2.64 (1.37-5.08) in Period 3, and for active smokers they were 3.01 (0.53-17.3) in Period 2 and 9.92 (2.20-44.7) in Period 3. The wide CIs are indicative of the small counts, but the ORs are large, especially for adolescents who were active smokers (which would have included vapers), for whom the odds of being ANA positive were roughly 10 times greater in Period 3 compared with Period 1. The differences across age categories, based on assessing an interaction between age group and time period, were statistically significant (p=0.009).

3.5 ANA associations with smoking by time period

In an alternative covariate-adjusted analysis, we focused on smoking status and assessed the odds of ANA positivity for smokers relative to nonsmokers (Table 4). Overall, smokers were less likely to have ANA than nonsmokers (OR=0.73; 95% CI: 0.58-0.92; p=0.007). When stratified by time period, the ANA prevalence ORs for smoking status varied in magnitude and statistical significance but not in direction. The odds of having ANA were significantly lower for smokers than nonsmokers in Period 2 (OR=0.65; 95% CI: 0.45-0.93; p=0.020), but that inverse association was weaker and not statistically significant in Period 1 (OR=0.80; 95% CI: 0.52-1.22; p=0.297) and Period 3 (OR=0.82; 95% CI: 0.56-1.21; p=0.310). This nonmonotonic temporal pattern is illustrated in Figure 3, where ANA prevalence estimates are smaller for smokers than nonsmokers in all three time periods, but the difference is much greater in Period 2 than in Periods 1 and 3.

3.6 Additional analyses

We performed several sensitivity analyses by adding covariates to a base model that was adjusted for sex, age, race/ethnicity, and time period (Supplementary Table S5). Rather than stratifying by smoking exposure, including it as a categorical covariate led to the same basic pattern of ANA prevalence not changing much between Periods 1 and 2, followed by a marked increase in Period 3. When we also added a smoking-by-period interaction, the main effects of both smoking and period were statistically significant, but the interaction was not. On the other hand, excluding time period significantly worsened the model fit (p<0.0001), suggesting that calendar time was important and that smoking on its own could not fully explain the observed ANA differences.

We also performed secondary analyses that accounted for BMI, which had been shown previously to modify ANA time trends (7). Adding a 3-level categorical covariate for BMI (underweight/healthy, <25; overweight, 25 to <30; or obese, ≥30) to the base

TABLE 3 Covariate-adjusted assessments of ANA time trends by smoking exposure and age group.

Smoking	ANA Prevalence Odds Ratio for Time Period (95% CI) [b]			Time Trend
Exposure [a]	Period 1: 1988-1991	Period 2: 1999-2004	Period 3: 2011-2012	p-value [b]
Age Group 1: 12–19 years old				
None	1.00 (reference)	3.64 (0.99-13.4)	2.84 (0.75-10.8)	0.5002
Passive	1.00 (reference)	1.63 (0.83-3.23)	2.64 (1.37-5.08)	0.0047
Active	1.00 (reference)	3.01 (0.53-17.3)	9.92 (2.20-44.7)	0.0032
Age Group 2: 20–49 years old				
None	1.00 (reference)	0.39 (0.21-0.72)	0.58 (0.32-1.06)	0.9447
Passive	1.00 (reference)	1.09 (0.70-1.70)	1.53 (0.91-2.60)	0.1435
Active	1.00 (reference)	0.73 (0.43-1.24)	1.28 (0.82-2.00)	0.4474
Age Group 3: ≥50 years old				
None	1.00 (reference)	0.87 (0.47-1.60)	1.29 (0.70-2.36)	0.1303
Passive	1.00 (reference)	1.37 (0.91-2.06)	1.41 (0.96-2.06)	0.0452
Active	1.00 (reference)	0.85 (0.48-1.47)	1.43 (0.78-2.63)	0.2353

ANA, antinuclear antibodies; CI, confidence interval.
[a]Smoking exposure categories were based on serum cotinine concentrations (None, ≤0.05 ng/mL; Passive, >0.05 to 10 ng/mL; and Active, >10 ng/mL).
[b]ANA time trend assessments were based on two logistic regression models for ANA positivity (yes/no). Both models stratified by smoking exposure and age group, and both adjusted for the survey design variables (sampling weights, strata, and clusters) and categorical covariates for sex and race/ethnicity. One model added a categorical covariate for time period, which allowed estimates of the ANA prevalence odds ratio for each period relative to the first period. The other model instead added a continuous covariate for time, as measured by the number of years between period midpoints relative to the first period, and produced a p-value from a t-test to assess a linear ANA time trend.

TABLE 4 Covariate-adjusted assessments of ANA associations with smoking status by time period.

Time Period	ANA Prevalence Odds Ratio for Smoking Status (95% CI) [a]		p-value [a]
	Nonsmoker	Smoker	
Period 1: 1988-1991	1.00 (reference)	0.80 (0.52-1.22)	0.297
Period 2: 1999-2004	1.00 (reference)	0.65 (0.45-0.93)	0.020
Period 3: 2011-2012	1.00 (reference)	0.82 (0.56-1.21)	0.310
All Periods Combined	1.00 (reference)	0.73 (0.58-0.92)	0.007

ANA, antinuclear antibodies; CI, confidence interval.
[a]Assessments of the association between ANA and smoking status were based on a logistic regression model for ANA positivity (yes/no) that adjusted for the survey design variables (sampling weights, strata, and clusters) and categorical covariates for sex, age, and race/ethnicity. The model also included a categorical covariate for smoking status, as defined by serum cotinine concentrations (Nonsmoker, ≤10 ng/mL; Smoker, >10 ng/mL), which allowed estimates of the ANA prevalence odds ratio for smokers relative to nonsmokers. The model was applied separately for each time period and also for all periods combined. The p-value for assessing statistical significance was based on a t-test.

model did not change the ANA time trends, nor did also adding a BMI interaction with any factor in the base model or in an expanded model that also included a covariate for smoking exposure (Supplementary Table S5). Similarly, the original ANA time trends within smoking exposure subgroups (as shown in Table 2) did not change much when augmenting the base model with covariates for a BMI main effect and a BMI-by-age interaction (Supplementary Table S6).

Replacing the 8-category age covariate with a 3-category age covariate, a quantitative age covariate, or a restricted cubic spline in age did not alter the basic pattern of ANA prevalence being relatively flat between Periods 1 and 2, and then increasing substantially in Period 3 (Supplementary Table S5). Similarly, that basic ANA pattern also remained consistent when other covariates were added to the base model, such as an individual main effect for alcohol intake, PIR, or education; those same main effects plus a

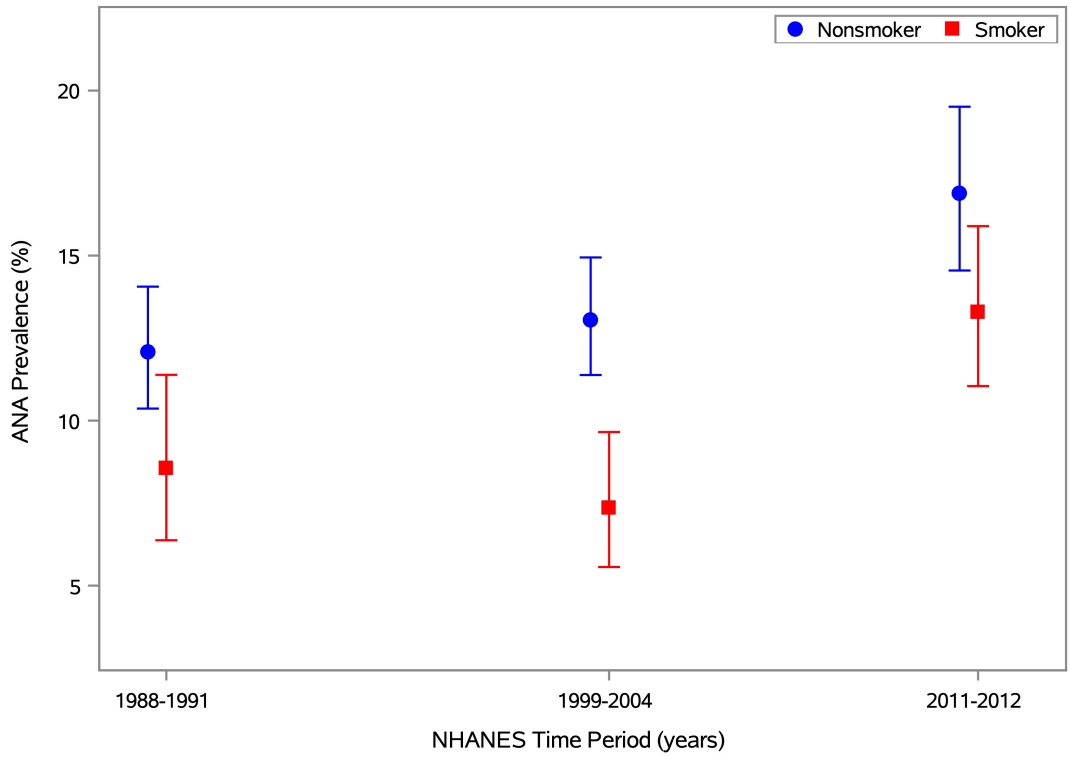


FIGURE 3
ANA prevalence by time period and smoking status. Estimates of ANA prevalence and its 95% CI are plotted for smokers and nonsmokers in Period 1 (1988-1991), Period 2 (1999-2004), and Period 3 (2011-2012), based on the 13,288 NHANES participants aged ≥12 years with data on both ANA and smoking status (i.e., serum cotinine). The prevalence estimates for nonsmokers and smokers are shown by blue circles and red squares, respectively, with the same colors used for the 95% CI error bars. Separately for each smoking status, the prevalence estimates and 95% CIs were derived from a logistic regression model for ANA positivity, adjusted for the survey-design variables and a categorical covariate for time period. The horizontal axis is linear in time, defined as the number of years between the midpoints of the participant's period and the first period, and the vertical axis is linear in ANA prevalence (as a percentage).

main effect for BMI; and those same main effects plus both a main effect for BMI and an interaction between BMI and each of those other covariates.

Data were available on the CO content in the brand of cigarettes used by each of 1,157 current smokers from Periods 2 and 3 aged ≥20 years. After adjusting for sex, age, race/ethnicity, and time period, there was mild evidence that ANA prevalence decreased as CO content per cigarette increased (OR=0.92; 95% CI: 0.85-1.00; p=0.042). We also multiplied CO content by average number of cigarettes smoked per day to estimate total CO, but the covariate-adjusted analysis showed no evidence of an association with ANA (p=0.65). In additional covariate-adjusted analyses of participants of all ages from all time periods, we found no evidence of an ANA association with pack-years among 4,795 ever (former or current) smokers (p=0.91), years of smoking among 5,047 ever smokers (p=0.18), or years since quitting among 2,738 former smokers (p=0.56). We also allowed for various pack-year threshold values, but no ANA associations with pack-years were significant.

4 Discussion

In summary, we assessed representative US data regarding ANA, time, smoking, CO and their interdependencies. Table 5 lists several relevant concepts and results, including the following information. High-titer ANA are associated with autoimmune diseases (1–6). In the US, ANA increased from 1988 to 2012, primarily in the second half of that interval (7); both active and passive exposure to smoke from regular cigarettes decreased during those years (16, 17); and e-cigarette use rapidly increased after being introduced in 2007 (15, 33–36), especially among adolescents (ages 12–19 years). CO may protect against ANA and certain autoimmune diseases (18–20, 25–32). E-cigarettes deliver much less CO than regular cigarettes (39, 40).

Our general observation is that something related to smoking cigarettes appears to have been inversely associated with ANA and any potentially protective effect waned in the later time period, possibly because people were smoking less and vaping more, or because something else about smoking changed. In most cases, our use of the word "protective" refers to a statistical association and not a proven biological protection. We hypothesize that reduced CO from decreased exposure to cigarette smoke may account for some of the overall increase in ANA. This reduction in CO could have come from current smokers cutting back on their cigarette consumption (including some degree of switching to vaping), from former smokers who quit (and possibly switched to vaping), and from never or passive smokers being exposed to less secondhand smoke (due to regulations and social pressure). We also hypothesize that the rapid increase in e-cigarette use after 2007, especially among teenagers, may partially explain why the increase in ANA prevalence was larger during the latter half of the study years and why the increasing ANA time trend was the most pronounced in teenagers (7). Our two-part hypothesis is consistent with what is already known about ANA, smoking, and CO, as well as with the results from our analyses of the NHANES data. Specifically, we assessed how the ANA time trend depended on smoking exposure levels, including within age subgroups, and how the ANA association with smoking depended on calendar time. Both are described below.

TABLE 5 Concepts and results related to the hypothesis that decreased cigarette smoking may partially explain the increased prevalence of antinuclear antibodies in the United States.

1. High-titer antinuclear antibodies (ANA) are associated with some autoimmune diseases, and ANA prevalence estimates increased over time: a little between Period 1 (1988-1991) and Period 2 (1999-2004) and a lot between Period 2 and Period 3 (2011-2012).
2. Viewing ANA time trends by smoking exposure, there was no clear trend over time in ANA prevalence estimates for individuals with no exposure (negligible serum cotinine), a steady increase for individuals with passive exposure (low serum cotinine), and a flat-then-increasing trend for individuals with active exposure (high serum cotinine). The ANA time trends among passive and active smokers were associated mainly with 12–19 year-olds.
3. Viewing ANA associations with smoking by time period, the estimated odds of having ANA were less among active smokers (high serum cotinine) than nonsmokers (negligible or low serum cotinine) in all time periods, but only the difference in Period 2 was statistically significant.
4. Serum cotinine steadily decreased over time, primarily in self-reported never and former smokers, but not in self-reported current smokers.
5. Smoking of regular cigarettes and secondhand exposure to their smoke steadily decreased over time.
6. Vaping of electronic cigarettes (e-cigarettes) began after Period 2 (in 2007) and rapidly increased over time.
7. Both regular cigarettes and e-cigarettes deliver nicotine and hence produce cotinine, but e-cigarettes produce much less carbon monoxide (CO) than regular cigarettes.
8. Some studies suggest that low levels of CO may be protective against ANA and certain autoimmune diseases.
9. In summary, less smoking of regular cigarettes may have led to less low-level CO exposure and more ANA. The hypothesized explanation involving potential CO protection against ANA is consistent with the observed patterns of ANA prevalence estimates, the long-term decreases in secondhand smoke exposure, and the recent increases in vaping, especially among adolescents (12–19 years old).

The ANA time trends across the three smoking-exposure subgroups (as defined by serum cotinine level) are consistent with our hypothesis. Individuals with no smoking exposure had negligible cotinine levels and presumably were not affected by changes in vaping or secondhand smoke. Thus, we infer that their exposure to CO from cigarette smoke was minimal and, consistent with our hypothesis, their ANA prevalence showed no clear time trend. Individuals with passive exposure to smoke had detectable but relatively low cotinine levels, which means they would have been affected by changes in secondhand smoke but probably were not regular vapers. Hence, these individuals might have experienced a small but steady increase in ANA prevalence across all time periods, which we speculate could be due to the steady decrease in their low-level CO "protection" from decreasing secondhand smoke (and possibly also from reduced exposure via air pollution (https://www.epa.gov/air-trends/carbon-monoxide-trends)). Active smokers had high cotinine levels, which could result from either regular cigarettes or e-cigarettes, and would have been affected by

changes in vaping but not secondhand smoke. Thus, these individuals presumably would not have had any change in potential CO protection or ANA prevalence between Periods 1 and 2, since vaping did not begin until 2007, but would have had a decrease in potential CO protection and, consistent with our hypothesis, a corresponding increase in ANA prevalence between Periods 2 and 3, as some of them took up vaping. Therefore, our hypothesis regarding potential smoking-associated CO protection from ANA is consistent with the possibility that the observed ANA patterns could be at least partially explained by the continued decrease in secondhand smoke exposure and the recent increase in vaping. In fact, when viewed by age group, the largest increase in ANA prevalence was between Periods 2 and 3 in teenagers who were active smokers, the timeframe and age group most associated with vaping.

The ANA associations with smoking seen across the three time periods are also consistent with our hypothesis and may relate to events that affected nonsmokers in the early years and smokers in the later years. Active smokers had significantly lower odds of having ANA than nonsmokers in Period 2, as would be expected if CO is protective, but this evident reduction was weaker (and not significant) in Periods 1 and 3. Between the first two periods, secondhand smoke exposure decreased (which would only affect nonsmokers) but vaping had not yet been introduced (which could only affect smokers who later started switching to e-cigarettes). All smokers had active smoking exposure, but nonsmokers were a mix of individuals with no exposure and passive exposure. The proportion of nonsmokers with passive exposure decreased over time, as presumably did their potential CO protection from secondhand smoke, and thus their ANA prevalence would have increased. However, neither potential CO protection nor ANA prevalence would have changed among smokers. Hence, the odds of having ANA for smokers versus nonsmokers would be smaller in Period 2 than in Period 1 (as we observed). Between Periods 2 and 3, secondhand smoke exposure again decreased (which would only affect nonsmokers) while vaping increased rapidly (which would mainly affect cotinine-identified active smokers). As described above, the level of potential CO protection from secondhand smoke among nonsmokers would have decreased, increasing their ANA prevalence. Concurrently, potential CO protection among active smokers (some of whom were vapers) would also have decreased due to increased vaping, and thus their ANA prevalence would have increased. The increase in ANA due to increased vaping among smokers could have more than offset the increase in ANA due to decreased secondhand smoke among nonsmokers, resulting in the ANA prevalences for smokers and nonsmokers to appear more similar in Period 3 than in Period 2 (as we observed).

Although we hypothesize that decreased CO and increased vaping may help explain both the changes in ANA time trends across smoking exposure levels and the changes in ANA associations with smoking across time periods, other factors may also have played a role. For example, cigarette smoke is composed of many chemicals with a wide array of effects on the body and we have an incomplete understanding of their immune impacts that could include both stimulatory and inhibitory elements that may vary from product to product (18, 45). Also, certain components of e-cigarettes, such as flavoring agents (46), may potentially increase the risk of developing ANA in users, and vaping may introduce additional chemical contaminants contributing to bystander health effects from secondhand exposure (47). In addition to vaping, there are other nicotine-delivering alternatives to regular cigarettes, including nicotine gum, chewing tobacco, snuff, and snus (48), that can have immune system effects (49). Another consideration is that some ANA subtypes may be more relevant than others. In a previous study (50), we found that time period and smoking exposure were more strongly associated with anti-dense fine speckled 70 autoantibodies than with total ANA. Miller (51) discussed a wide range of other potentially relevant factors such as elements of the environment, various lifestyles, and even climate change that could impact recent increases in autoimmunity and autoimmune diseases.

Our study had several strengths. The NHANES cohort with data on ANA was very large and spanned 25 years (1988–2012), with all ANA assays performed in the same laboratory, using the same evaluators, methods, and equipment. All statistical analyses were weighted to enable analytic results that generalize to the civilian noninstitutionalized US population ≥12 years old. Many of our analyses of ANA, cotinine, and smoking accounted for sex, age, and race/ethnicity as potential correlates or modifiers, and some analyses also adjusted for BMI, alcohol intake, PIR, or education.

On the other hand, our descriptive findings are subject to certain limitations. There may be concerns about the age of serum samples used for ANA assessment, some of which were nearly three decades old when assayed. However, there were no gross differences in appearance or behavior to suggest degradation, and antibodies are stable over time in frozen storage (52). Some NHANES data were obtained from questionnaires, such as smoking history, but self-reported nicotine product use has been shown to be valid (53). As vaping has increased, high cotinine levels have become less reliable for identifying persons who only smoke regular cigarettes (and thus are exposed to more CO). We considered using self-reported smoking history instead, but that information was often missing and it was not clear whether persons who replaced some or all of their regular cigarettes with e-cigarettes would classify themselves as former or current smokers. Also, we used 10 ng/mL of cotinine to distinguish passive and active smokers, as recommended by the CDC and EPA (13), but some researchers have suggested using a lower cutpoint, such as 3 ng/mL (54). However, our sensitivity analysis found that using the lower cotinine cutpoint had little effect on the results. No participant was followed longitudinally; thus, both cotinine and ANA were assessed cross-sectionally at only one point in time per participant, so measured cotinine levels may poorly reflect the levels when ANA developed. Reported associations, even if confirmed, may not correspond to causal effects. In fact, there could be reverse-causal effects if immune system or other changes associated with ANA influence smoking behavior or the metabolism of nicotine, cotinine, or other byproducts of smoking.

Perhaps the most serious deficiency in our data is the lack of direct information about e-cigarette use. At the time of our analyses, there were limited NHANES data on vaping in the 2013-2014, 2015-2016, and 2017–2018 cycles, but none in cycles with data on ANA. However, despite this absence of direct data, we might assume that most self-reported current smokers in Period 3 with a high cotinine level probably smoked regular cigarettes, whereas most self-reported former smokers with a high cotinine level had probably switched to e-cigarettes. The first group included 89/639 (13.9%) with ANA, while the second group included 18/85 (21.2%) with ANA, a difference that is consistent with our hypothesis of a potentially protective effect of CO derived from smoking regular cigarettes (and also consistent with an effect of something in e-cigarettes on ANA). Also, direct information on individual CO levels would have been helpful, though we found some evidence that lower ANA prevalence was associated with cigarette brands having higher CO content, which provides additional indirect support for our hypothesis.

In conclusion, cigarette smoking decreased over the past several decades and ANA prevalence increased, which we corroborated with analyses of NHANES data. However, the degree to which these two time trends might be causally related is unclear. Cotinine was used to infer exposure to cigarette smoke, and average levels steadily declined between 1988 and 2012 in the NHANES cohorts, with a downward-sloping straight line providing a good fit to log-transformed cotinine concentrations. The prevalence of ANA rose between 1988 and 2012, but this upward trend was not linear, showing a relatively small increase from 1988–1991 to 1999-2004, followed by a much larger increase from 1999–2004 to 2011-2012. The latter time interval coincides with the introduction of vaping, with many smokers replacing at least some of their regular cigarettes with e-cigarettes. That change might not have affected cotinine levels but should have reduced CO levels. We suggest that such a drop in CO levels potentially could be causally associated with the concurrent increase in ANA, as there is evidence that low levels of CO are protective against ANA and certain autoimmune diseases. However, while CO may be one factor in this process, one should keep in mind that there are many additional byproducts of smoking that possibly could play a role. Nonetheless, decreased smoking exposure (active and passive) across all study years could have contributed to a general increase in ANA, which could have been greatly supplemented in the later years by the rapid increase in vaping. Thus, smokers who reduced their use of regular cigarettes in favor of vaping may have lost some of the hypothesized protective effect afforded by CO, which could have increased their risk of developing autoimmunity.

We searched the literature for additional mechanisms and contributing factors that might help explain why decreased smoking could lead to increased ANA and found conflicting data on the complex mixtures that make up tobacco smoke and e-cigarette vapor. One parallel mechanism to CO is nicotine itself. Reduced cigarette smoking, if not replaced by other nicotine sources (48), would decrease the nicotine anti-inflammatory processes, which could then increase inflammation and ANA. For example, despite smoking being an established risk factor for rheumatoid arthritis (RA), several investigators have discussed a possible therapeutic effect of nicotine on RA (55–57). In the end, we concluded that exact mechanisms for why less smoking is associated with more ANA are unclear and further research is needed to identify the causes of the recent dramatic increases in ANA in the US. Hopefully, future studies will collect data on vaping history and CO biomarkers, which could provide direct evidence to assess our hypothesis.

In closing, given the many negative effects of smoking on increasing deaths, illnesses, and health care costs worldwide, we are certainly not recommending that smoking should be considered as an approach to prevent autoimmunity or autoimmune diseases. Rather, we believe that further studies in this area are needed as they may elucidate new mechanisms, perhaps involving certain components of tobacco smoke or e-cigarette vapor, that could allow for the development of novel preventative or treatment measures in the future.

Data availability statement

All data are publicly available. 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 humans were approved by CDC Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The human samples used in this study were acquired from the CDC NHANES study. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

GD: Conceptualization, Writing – review & editing, Formal analysis, Methodology, Resources, Writing – original draft, Software, Visualization. CW: Conceptualization, Methodology, Writing – review & editing. CP: Conceptualization, Methodology, Writing – review & editing. CC: Methodology, Writing – review & editing, Formal analysis, Data curation, Software. JP: Formal analysis, Methodology, Writing – review & editing, Software, Visualization. EC: Writing – review & editing, Conceptualization, Data curation, Resources. FM: Conceptualization, Writing – review & editing, Funding acquisition, Project administration.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This research was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences under project Z01 ES101074 and under contracts HHSN273201600011C GS-00F-173CA/75N96022F00055 to Social & Scientific Systems, Inc (a DLH Holdings Corp. Company). None of these organizations have a financial relationship with any cigarette or vaping manufacturers or distributors.

Acknowledgments

We thank Drs. Richard Weinberg and Katherine Allen for their helpful comments and Dr. Geraldine McQuillan for administrative and regulatory assistance. We also thank the members of the NHANES Autoimmunity Study Group (including Drs. Linda Birnbaum, Richard Cohn, Dori Germolec, Minoru Satoh, Nigel Walker, Irene Whitt, and Darryl Zeldin) for initiating the studies that motivated much of our research.

Conflict of interest

Authors GD, CC, and JP were employed by DLH, LLC.

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 GK declared a past co-authorship with the author EC to the handling editor.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2025.1537043/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Mean Serum Cotinine Concentration by Time Period. Estimates of the geometric mean serum cotinine concentration and its 95% CI are plotted for each of 3 time periods (1988-1991, 1999-2004, and 2011-2012), along with the best-fitting trend line. These estimates are based on the 13,288 NHANES participants aged ≥12 years with data on both ANA and serum cotinine. Any concentration below the limit of detection (LOD) was replaced by an imputed value equal to LOD/ $\sqrt{2}$. The horizontal axis is linear in time, defined as the number of years between the midpoints of the participant's period and the first period, and the vertical axis is logarithmic in serum cotinine concentration (ng/mL).

References

1. Sur LM, Floca E, Sur DG, Colceriu MC, Samasca G, Sur G. Antinuclear antibodies: marker of diagnosis and evolution in autoimmune diseases. *Lab Med.* (2018) 49:e62–73. doi: 10.1093/labmed/lmy024

2. Damoiseaux J, Andrade LEC, Carballo OG, Conrad K, Francescantonio PLC, Fritzler MJ, et al. Clinical relevance of hep-2 indirect immunofluorescent patterns: the international consensus on ana patterns (Icap) perspective. *Ann Rheum Dis.* (2019) 78:879–89. doi: 10.1136/annrheumdis-2018-214436

3. Bossuyt X, De Langhe E, Borghi MO, Meroni PL. Understanding and interpreting antinuclear antibody tests in systemic rheumatic diseases. *Nat Rev Rheumatol.* (2020) 16:715–26. doi: 10.1038/s41584-020-00522-w

4. Damoiseaux JG, Tervaert JW. From ana to ena: how to proceed? *Autoimmun Rev.* (2006) 5:10–7. doi: 10.1016/j.autrev.2005.05.007

5. Dahlberg A, Tydén H, Jöud AS, Kahn F, Berthold E. Screening for comorbid autoimmune disease should be considered in children with ana positive juvenile idiopathic arthritis - results from the south-swedish juvenile idiopathic arthritis cohort. *Pediatr Rheumatol Online J.* (2024) 22:92. doi: 10.1186/s12969-024-01030-x

6. Mohd Razi MS, Sugumaran Y, Mohd Haniz NA, Khilmie K, Osmera AH, Jauhary EJ, et al. Antinuclear antibody titration and pattern are helpful in the diagnosis of systemic autoimmune rheumatic diseases. *Malays J Pathol.* (2024) 46:63–9.

7. Dinse GE, Parks CG, Weinberg CR, Co CA, Wilkerson J, Zeldin DC, et al. Increasing prevalence of antinuclear antibodies in the United States. *Arthritis Rheumatol.* (2022) 74:2032–41. doi: 10.1002/art.42330

8. Dinse GE, Co CA, Parks CG, Weinberg CR, Xie G, Chan EKL, et al. Expanded assessment of xenobiotic associations with antinuclear antibodies in the United States, 1988-2012. *Environ Int.* (2022) 166:107376. doi: 10.1016/j.envint.2022.107376

9. Goodchild M, Nargis N, Tursan d'Espaignet E. Global economic cost of smoking-attributab le diseases. *Tob Control.* (2018) 27:58–64. doi: 10.1136/tobaccocontrol-2016-053305

10. WHO. (World Health Organization). *Who Global Report: Mortality Attributab le to Tobacco* (2012). Available online at: https://www.who.int/publications/i/item/9789241564434 (Accessed December 16, 2024).

11. Benowitz NL. Cotinine as a biomarker of environmental tobacco smoke exposure. *Epidemiol Rev.* (1996) 18:188–204. doi: 10.1093/oxfordjournals.epirev.a017925

12. Mourino N, Pérez-Ríos M, Santiago-Pérez MI, Lanphear B, Yolton K, Braun JM. Secondhand Tobacco Smoke Exposure among Children under 5 Years Old: Questionnaires Versus Cotinine Biomarkers: A Cohort Study. *BMJ Open.* (2021) 11:e044829. doi: 10.1136/bmjopen-2020-044829

13. EPA. (Environmental Protection Agency). *Serum cotinine level* (2022). Available online at: https://cfpub.epa.gov/roe/indicator.cfm?i=26 (Accessed December 16, 2024).

14. CDC. (Centers for Disease Control and Prevention). *Cotinine factsheet* (2016). Available online at: http://medbox.iiab.me/modules/en-cdc/www.cdc.gov/biomonitoring/Cotinine_FactSheet.html (Accessed September 25, 2024).

15. DHHS. (Department of Health and Human Services). *E-cigarette use among youth and young adults: A report of the surgeon general*. Department of Health and Human Services, Public Health Service, Office of the Surgeon General, Rockville (2016).

16. CDC. (Centers for Disease Control and Prevention). *Current Cigarette Smoking among Adults in the United States* (2024). Available online at: https://www.cdc.gov/tobacco/php/data-statistics/adult-data-cigarettes/?CDC_AAref_Val=https://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/index.htm (Accessed September 25, 2024).

17. NCI. (National Cancer Institute). *Adult tobacco use* (2024). Available online at: https://progressreport.cancer.gov/prevention/tobacco/adult_smoking (Accessed September 25, 2024).

18. Perricone C, Versini M, Ben-Ami D, Gertel S, Watad A, Segel MJ, et al. Smoke and autoimmunity: the fire behind the disease. *Autoimmun Rev*. (2016) 15:354–74. doi: 10.1016/j.autrev.2016.01.001

19. Sopori M. Effects of cigarette smoke on the immune system. *Nat Rev Immunol*. (2002) 2:372–7. doi: 10.1038/nri803

20. Servioli L, Maciel G, Nannini C, Crowson CS, Matteson EL, Cornec D, et al. Association of smoking and obesity on the risk of developing primary sjögren syndrome: A population-based cohort study. *J Rheumatol*. (2019) 46:727–30. doi: 10.3899/jrheum.180481

21. Arnson Y, Shoenfeld Y, Amital H. Effects of Tobacco Smoke on Immunity, Inflammation and Autoimmunity. *J Autoimmun*. (2010) 34:J258–65. doi: 10.1016/j.jaut.2009.12.003

22. Effraimidis G, Tijssen JG, Wiersinga WM. Discontinuation of smoking increases the risk for developing thyroid peroxidase antibodies and/or thyroglobulin antibodies: A prospective study. *J Clin Endocrinol Metab*. (2009) 94:1324–8. doi: 10.1210/jc.2008-1548

23. Carlé A, Bülow Pedersen I, Knudsen N, Perrild H, Ovesen L, Banke Rasmussen L, et al. Smoking cessation is followed by a sharp but transient rise in the incidence of overt autoimmune hypothyroidism - a population-based, case-control study. *Clin Endocrinol (Oxf)*. (2012) 77:764–72. doi: 10.1111/j.1365-2265.2012.04455.x

24. Elsherbiny TM. Quitting smoking as a probable trigger for new-onset hypothyroidism after successful medical treatment of graves' Disease: case report. *Ther Adv Endocrinol Metab*. (2024) 15:20420188241256470. doi: 10.1177/20420188241256470

25. Fagone P, Piombino E, Mangano K, De Pasquale R, Nicoletti F, Caltabiano R. Evaluation of the involvement of heme oxygenase-1 expression in discoid lupus erythematosus lesions. *Antioxidants (Basel)*. (2023) 12:1352. doi: 10.3390/antiox12071352

26. Chora AA, Fontoura P, Cunha A, Pais TF, Cardoso S, Ho PP, et al. Heme oxygenase-1 and carbon monoxide suppress autoimmune neuroinflammation. *J Clin Invest*. (2007) 117:438–47. doi: 10.1172/jci28844

27. Fagone P, Mangano K, Coco M, Perciavalle V, Garotta G, Romao CC, et al. Therapeutic potential of carbon monoxide in multiple sclerosis. *Clin Exp Immunol*. (2012) 167:179–87. doi: 10.1111/j.1365-2249.2011.04491.x

28. Ferrándiz ML, Maicas N, Garcia-Arnandis I, Terencio MC, Motterlini R, Devesa I, et al. Treatment with a co-releasing molecule (Corm-3) reduces joint inflammation and erosion in murine collagen-induced arthritis. *Ann Rheum Dis*. (2008) 67:1211–7. doi: 10.1136/ard.2007.082412

29. Mackern-Oberti JP, Llanos C, Carreño LJ, Riquelme SA, Jacobelli SH, Anegon I, et al. Carbon monoxide exposure improves immune function in lupus-prone mice. *Immunology*. (2013) 140:123–32. doi: 10.1111/imm.12124

30. Nikolic I, Saksida T, Mangano K, Vujicic M, Stojanovic I, Nicoletti F, et al. Pharmacological application of carbon monoxide ameliorates islet-directed autoimmunity in mice via anti-inflammatory and anti-apoptotic effects. *Diabetologia*. (2014) 57:980–90. doi: 10.1007/s00125-014-3170-7

31. Fagone P, Mangano K, Mammana S, Cavalli E, Di Marco R, Barcellona ML, et al. Carbon monoxide-releasing molecule-A1 (Corm-A1) improves clinical signs of experimental autoimmune uveoretinitis (Eau) in rats. *Clin Immunol*. (2015) 157:198–204. doi: 10.1016/j.clim.2015.02.002

32. Mangano K, Cavalli E, Mammana S, Basile MS, Caltabiano R, Pesce A, et al. Involvement of the nrf2/ho-1/co axis and therapeutic intervention with the co-releasing molecule corm-A1, in a murine model of autoimmune hepatitis. *J Cell Physiol*. (2018) 233:4156–65. doi: 10.1002/jcp.26223

33. Regan AK, Promoff G, Dube SR, Arrazola R. Electronic nicotine delivery systems: adult use and awareness of the 'E-cigarette' in the USA. *Tob Control*. (2013) 22:19–23. doi: 10.1136/tobaccocontrol-2011-050044

34. King BA, Patel R, Nguyen KH, Dube SR. Trends in awareness and use of electronic cigarettes among us adults, 2010-2013. *Nicotine Tob Res*. (2015) 17:219–27. doi: 10.1093/ntr/ntu191

35. Glasser AM, Collins L, Pearson JL, Abudayyeh H, Niaura RS, Abrams DB, et al. Overview of electronic nicotine delivery systems: A systematic review. *Am J Prev Med*. (2017) 52:e33–66. doi: 10.1016/j.amepre.2016.10.036

36. Bandi P, Cahn Z, Goding Sauer A, Douglas CE, Drope J, Jemal A, et al. Trends in E-cigarette use by age group and combustible cigarette smoking histories, U.S. Adults, 2014-2018. *Am J Prev Med*. (2021) 60:151–8. doi: 10.1016/j.amepre.2020.07.026

37. Jaber RM, Mirbolouk M, DeFilippis AP, Maziak W, Keith R, Payne T, et al. Electronic cigarette use prevalence, associated factors, and pattern by cigarette smoking status in the United States from nhanes (National health and nutrition examination survey) 2013-2014. *J Am Heart Assoc*. (2018) 7:e008178. doi: 10.1161/jaha.117.008178

38. Pulvers K, Nollen NL, Rice M, Schmid CH, Qu K, Benowitz NL, et al. Effect of pod E-cigarettes vs cigarettes on carcinogen exposure among african american and latinx smokers: A randomized clinical trial. *JAMA Netw Open*. (2020) 3:e2026324. doi: 10.1001/jamanetworkopen.2020.26324

39. Wagener TL, Floyd EL, Stepanov I, Driskill LM, Frank SG, Meier E, et al. Have combustible cigarettes met their match? The nicotine delivery profiles and harmful constituent exposures of second-generation and third-generation electronic cigarette users. *Tob Control*. (2017) 26:e23–e8. doi: 10.1136/tobaccocontrol-2016-053041

40. Raffael F, Pandia P, Tarigan AP, Mutiara E, Osakue OE. Comparison of exhaled carbon monoxide levels and its association with nicotine dependence between electronic and tobacco cigarettes: A cross-sectional study among teenage smokers. *Narra J*. (2023) 3:e418. doi: 10.52225/narra.v3i3.418

41. Johnson CL, Paulose-Ram R, Ogden CL, Carroll MD, Kruszon-Moran D, Dohrmann SM, et al. National health and nutrition examination survey: analytic guidelines, 1999-2010. *Vital Health Stat 2*. (2013) 161):1–24.

42. Satoh M, Chan EK, Ho LA, Rose KM, Parks CG, Cohn RD, et al. Prevalence and sociodemographic correlates of antinuclear antibodies in the United States. *Arthritis Rheum*. (2012) 64:2319–27. doi: 10.1002/art.34380

43. Hornung RW, Reed LD. Estimation of average concentration in the presence of nondetecta ble values. *Appl Occup Environ Hygiene*. (1990) 5:46–51. doi: 10.1080/1047322X.1990.10389587

44. CDC. (Centers for Disease Control and Prevention). *Limit of detection* (2024). Available online at: https://www.cdc.gov/environmental-exposure-report/php/about-the-data/limit-of-detection.html (Accessed September 25, 2024).

45. Maiyo AK, Kibet JK, Kengara FO. A review of the characteristic properties of selected tobacco chemicals and their associated etiological risks. *Rev Environ Health*. (2023) 38:479–91. doi: 10.1515/reveh-2022-0013

46. Morris AM, Leonard SS, Fowles JR, Boots TE, Mnatsakanova A, Attfield KR. Effects of E-cigarette flavoring chemicals on human macrophages and bronchial epithelial cells. *Int J Environ Res Public Health*. (2021) 18:11107. doi: 10.3390/ijerph182111107

47. Islam T, Braymiller J, Eckel SP, Liu F, Tackett AP, Rebuli ME, et al. Secondhand nicotine vaping at home and respiratory symptoms in young adults. *Thorax*. (2022) 77:663–8. doi: 10.1136/thoraxjnl-2021-217041

48. ACOG. (American College of Obstetricians and Gynecologists). Tobacco and nicotine cessation during pregnancy: acog committee opinion, number 807. *Obstet Gynecol*. (2020) 135:e221–e9. doi: 10.1097/aog.0000000000003822

49. Malovichko MV, Zeller I, Krivokhizhina TV, Xie Z, Lorkiewicz P, Agarwal A, et al. Systemic toxicity of smokeless tobacco products in mice. *Nicotine Tob Res*. (2019) 21:101–10. doi: 10.1093/ntr/ntx230

50. Dinse GE, Zheng B, Co CA, Parks CG, Weinberg CR, Miller FW, et al. Anti-dense fine speckled 70 (Dfs70) autoantibodies: correlates and increasing prevalence in the United States. *Front Immunol*. (2023) 14:1186439. doi: 10.3389/fimmu.2023.1186439

51. Miller FW. The environment, lifestyles and climate change: the many nongenetic contributors to the long and winding road to autoimmune diseases. *Arthritis Care Res (Hoboken)*. (2024) 77: 3–11. doi: 10.1002/acr.25423

52. Argentieri MC, Pilla D, Vanzati A, Lonardi S, Facchetti F, Doglioni C, et al. Antibodies are forever: A study using 12-26-year-old expired antibodies. *Histopathology*. (2013) 63:869–76. doi: 10.1111/his.12225

53. Yeager DS, Krosnick JA. The validity of self-reported nicotine product use in the 2001–2008 national health and nutrition examination survey. *Med Care*. (2010) 48:1128–32. doi: 10.1097/MLR.0b013e3181ef9948

54. Benowitz NL, Bernert JT, Caraballo RS, Holiday DB, Wang J. Optimal serum cotinine levels for distinguishing cigarette smokers and nonsmokers within different racial/ethnic groups in the United States between 1999 and 2004. *Am J Epidemiol*. (2009) 169:236–48. doi: 10.1093/aje/kwn301

55. Sungwon R. Smoking as a preventa ble risk factor for rheumatoid arthritis: rationale for smoking cessation treatment in patients with rheumatoid arthritis. *J Rheum Dis*. (2019) 26:12–9. doi: 10.4078/jrd.2019.26.1.12

56. Wu S, Luo H, Xiao X, Zhang H, Li T, Zuo X. Attenuation of collagen induced arthritis via suppression on th17 response by activating cholinergic anti-inflammatory pathway with nicotine. *Eur J Pharmacol*. (2014) 735:97–104. doi: 10.1016/j.ejphar.2014.04.019

57. Zhou Y, Zuo X, Li Y, Wang Y, Zhao H, Xiao X. Nicotine inhibits tumor necrosis factor-α Induced il-6 and il-8 secretion in fibroblast-like synoviocytes from patients with rheumatoid arthritis. *Rheumatol Int*. (2012) 32:97–104. doi: 10.1007/s00296-010-1549-4

Frontiers in Immunology

Explores novel approaches and diagnoses to treat immune disorders.

The official journal of the International Union of Immunological Societies (IUIS) and the most cited in its field, leading the way for research across basic, translational and clinical immunology.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact