



OPEN ACCESS

EDITED BY

Shisan (Bob) Bao,
The University of Sydney, Australia

REVIEWED BY

Zhongheng Zhang,
Sir Run Run Shaw Hospital, China
Dragos Cretoiu,
Carol Davila University of Medicine and
Pharmacy, Romania

*CORRESPONDENCE

Tamara Hernandez-Beefink
✉ tamarahdez7@gmail.com
Fernando Vaquerizo-Villar
✉ fernando.vaquerizo@uva.es

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

RECEIVED 10 June 2025

REVISED 12 November 2025

ACCEPTED 13 November 2025

PUBLISHED 15 December 2025

CITATION

Vaquerizo-Villar F, Hernandez-Beefink T,
Heredia-Rodríguez M, Gómez-Sánchez E,
Lorenzo-López M, López-Herrero R,
Bardaji-Carrillo M, Tamayo-Velasco Á,
Martín-Fernández M, Sánchez-de-Prada L,
Álvarez-Escudero J, Veiras S, Baluja A,
Gonzalo-Benito H, Martínez-Paz P,
García-Concejo A, Fernández-Rodríguez A,
Jiménez-Sousa MA, Resino S,
Martínez-Campelo L, Suárez-Pajés E,
Quintela I, Cruz R, Carracedo Á, Villar J,
Flores C, Hornero R and Tamayo E (2025)
Identifying sepsis susceptibility genes in
post-surgical patients using an artificial
intelligence approach.
Front. Med. 12:1644800.
doi: 10.3389/fmed.2025.1644800

COPYRIGHT

© 2025 Vaquerizo-Villar, Hernandez-Beefink,
Heredia-Rodríguez, Gómez-Sánchez,
Lorenzo-López, López-Herrero,
Bardaji-Carrillo, Tamayo-Velasco,
Martín-Fernández, Sánchez-de-Prada,
Álvarez-Escudero, Veiras, Baluja,
Gonzalo-Benito, Martínez-Paz,
García-Concejo, Fernández-Rodríguez,
Jiménez-Sousa, Resino, Martínez-Campelo,
Suárez-Pajés, Quintela, Cruz, Carracedo,
Villar, Flores, Hornero and Tamayo. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Identifying sepsis susceptibility genes in post-surgical patients using an artificial intelligence approach

Fernando Vaquerizo-Villar^{1,2,3*}, Tamara Hernandez-Beefink^{4*},
María Heredia-Rodríguez^{5,6,7}, Esther Gómez-Sánchez^{1,5,6,7},
Mario Lorenzo-López^{1,5,6,7}, Rocío López-Herrero^{1,5,6,7},
Miguel Bardaji-Carrillo^{1,5,7}, Álvaro Tamayo-Velasco^{5,6,8},
Marta Martín-Fernández^{5,6,9}, Laura Sánchez-de-Prada⁵,
Julián Álvarez-Escudero¹⁰, Sonia Veiras^{10,11}, Aurora Baluja^{11,12},
Hugo Gonzalo-Benito^{5,6,13}, Pedro Martínez-Paz^{5,6,14},
Adrián García-Concejo^{5,6}, Amanda Fernández-Rodríguez^{5,6,15},
María A. Jiménez-Sousa^{5,6,15}, Salvador Resino^{5,6,15},
Laura Martínez-Campelo^{16,17}, Eva Suárez-Pajés¹⁸,
Inés Quintela^{11,16,17}, Raquel Cruz^{16,17}, Ángel Carracedo^{12,16,17},
Jesús Villar^{19,20,21,22}, Carlos Flores^{18,19,23,24}, Roberto Hornero^{2,3} and
Eduardo Tamayo^{1,5,6,7}

¹Department of Anaesthesiology, Hospital Clínico Universitario de Valladolid, Valladolid, Spain,

²Biomedical Engineering Group, University of Valladolid, Valladolid, Spain, ³Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Instituto de Salud Carlos III, Valladolid, Spain, ⁴Division of Public Health and Epidemiology, School of Medical Sciences, University of Leicester, Leicester, United Kingdom, ⁵BioCritic, Group for Biomedical Research in Critical Care Medicine, Valladolid, Spain, ⁶Centro de Investigación Biomédica en Red de Enfermedades Infecciosas (CIBERINFEC), Instituto de Salud Carlos III, Madrid, Spain, ⁷Department of Surgery, University of Valladolid, Valladolid, Spain, ⁸Department of Haematology and Hemotherapy, Hospital Clínico Universitario de Valladolid, Valladolid, Spain, ⁹Department of Cell Biology, Genetics, Histology and Pharmacology, Universidad de Valladolid, Valladolid, Spain, ¹⁰Department of Anaesthesiology and Intensive Care Medicine, Clinical University Hospital of Santiago, Santiago de Compostela, Spain, ¹¹Sanitary Research Institute of Santiago (IDIS), Santiago de Compostela, Spain, ¹²Servicio de Anestesiología, Hospital Virxe Da Xunqueira, A Coruña, Spain, ¹³Research Support Unit, Hospital Clínico Universitario de Valladolid, Valladolid, Spain, ¹⁴Centre for Experimental Medicine and Rheumatology, William Harvey Research Institute, Queen Mary University of London, London, United Kingdom, ¹⁵Unidad de Infección Viral e Inmunidad, Centro Nacional de Microbiología (CNM), Instituto de Salud Carlos III, Majadahonda, Spain, ¹⁶Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Universidad de Santiago de Compostela, Santiago de Compostela, Spain, ¹⁷Fundación Pública Galega de Medicina Xenómica, Servizo Galego de Saúde, Santiago de Compostela, Spain, ¹⁸Research Unit, Hospital Universitario Nuestra Señora de Candelaria, Instituto de Investigación Sanitaria de Canarias (IISC), Santa Cruz de Tenerife, Spain, ¹⁹CIBER de Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, Madrid, Spain, ²⁰Research Unit, Hospital Universitario Dr. Negrín, Las Palmas de Gran Canaria, Spain, ²¹Li Ka Shing Knowledge Institute at St. Michael's Hospital, Toronto, ON, Canada, ²²Faculty of Health Sciences, Universidad del Atlántico Medio, Tafira Baja, Las Palmas, Spain, ²³Genomics Division, Instituto Tecnológico y de Energías Renovables (ITER), Santa Cruz de Tenerife, Spain, ²⁴Facultad de Ciencias de la Salud, Universidad Fernando de Pessoa Canarias, Las Palmas de Gran Canaria, Spain

Background: Early detection of sepsis is essential for its successful management. Although genome-wide association studies (GWAS) have shown potential in identifying sepsis-related genetic variants, they often involve heterogeneous patient groups and use single-locus analysis methods. Here, we aim to identify

new sepsis susceptibility loci in post-surgical patients using an explainable artificial intelligence (XAI) approach applied to GWAS data.

Methods: GWAS was performed in 750 post-operative patients with sepsis and 3,500 population controls. We applied a novel XAI-based methodology to GWAS-derived single nucleotide polymorphisms (SNPs) to predict sepsis and prioritize new genetic variants associated with post-operative sepsis susceptibility. We also assessed functional and enrichment effects using empirical data from integrated software tools and datasets, with the top-ranked variants and associated genes.

Results: Our XAI-GWAS approach showed a notable performance in predicting post-surgical sepsis and prioritized SNPs (such as rs17653532, rs1575081785, and rs74707084) with higher contribution to post-operative sepsis prediction. It also facilitated the discovery of post-operative sepsis risk loci with important functional implications related to gene expression regulation, DNA replication, cyclic nucleotide signaling, cell proliferation, and cardiac dysfunction.

Conclusion: The combination of GWAS and XAI prioritized loci associated with post-operative sepsis susceptibility. The determination of key genes, such as *PRIM2*, *SYNPR*, and *RBSN*, through pre-operative blood tests could enhance risk stratification, enable early detection of post-operative sepsis, and guide targeted interventions to improve patient outcomes. Further research with additional and ethnically diverse cohorts comprising sepsis and non-sepsis patients undergoing major surgery is needed to validate these exploratory findings.

KEYWORDS

explainable artificial intelligence (XAI), genome-wide association study (GWAS), sepsis, personalized medicine, surgical patients

1 Introduction

Sepsis, a global health priority, is defined as a severe host response to a systemic infection, leading to a life-threatening organ dysfunction (1), with an incidence of approximately 189 adult cases per 100,000 population/annually (2, 3). The global mortality rate was around 17% in 2017, with significantly higher rates observed in patients with septic shock (4). Sepsis survivors may face significant functional and cognitive long-term disability (5) with important health and socioeconomic consequences. The use of sepsis bundles could improve survival (6, 7) although early identification of sepsis is mandatory, which is challenging even by experienced clinicians (6).

Multiple studies have focused on early detection of sepsis using information from patient demographics, vital signs, laboratory results, biosensors, and/or genetics (8–11). Since the host immune response to microbial agents is influenced by genetic variation (8), recent genome-wide association studies (GWAS) showed a potential to identify genomic variants associated with sepsis susceptibility (12), sepsis-associated acute respiratory distress syndrome (13), and sepsis-associated mortality (4, 14–17) in adults.

The genes identified in the few GWAS on sepsis patients differ among studies (4, 12–17), which could be mainly attributed to the heterogeneity of the patient populations involved. In this context, post-operative sepsis represents up to one-third of all sepsis patients (18) and has mortality rates that can reach up to 50% (3, 19). Its management involves high economic costs due to the need for mechanical ventilation and prolonged hospital stays (19, 20). Early diagnosis is complicated by the similarity of its symptoms to normal post-operative inflammatory responses (21), which delays timely intervention. Additionally, antibiotic resistance and the need for additional surgical interventions

increase the complexity of treatment, especially in immunocompromised patients or those with multiple risk factors, where available biomarkers lack sufficient specificity for an accurate diagnosis (20). Postoperative sepsis is highly relevant for genetic studies, because the time of the precipitating insult is known (surgery). Hence, a preoperative assessment of genetic susceptibility to sepsis could enable risk stratification, guiding intensified monitoring or preventive interventions to better protect high-risk surgical patients (12, 22).

The existing GWAS of sepsis are characterized by the use of a single-locus analysis to select significant single-nucleotide polymorphisms (SNPs) (4, 12–17). In this respect, conventional methods may fail to capture complex interactions of SNPs with intermediate specificity that may have a higher contribution to the heritability of phenotypes (23, 24). Conversely, explainable artificial intelligence (XAI) approaches, which provide an ability to explain the decisions taken by complex artificial intelligence (AI)-based algorithms (25), have shown their usefulness to prioritize disease-associated genes (26–29). Specifically, these approaches have been previously applied for accurate phenotype prediction and susceptibility genes identification from GWAS-derived SNPs in complex diseases, such as atrial fibrillation (26), attention deficit hyperactivity disorder (27), hypertension, or diabetes (28).

In this exploratory study, we have analyzed genotype data from a large and homogeneous cohort of post-operative patients with sepsis and population controls to accurately predict sepsis and prioritize novel genetic variants associated with sepsis susceptibility. We hypothesized that XAI-based analysis of GWAS data could lead to the identification of new post-surgical sepsis susceptibility genes. Accordingly, our main objective was to obtain an XAI model that allows us to identify SNPs contributing to accurate prediction of

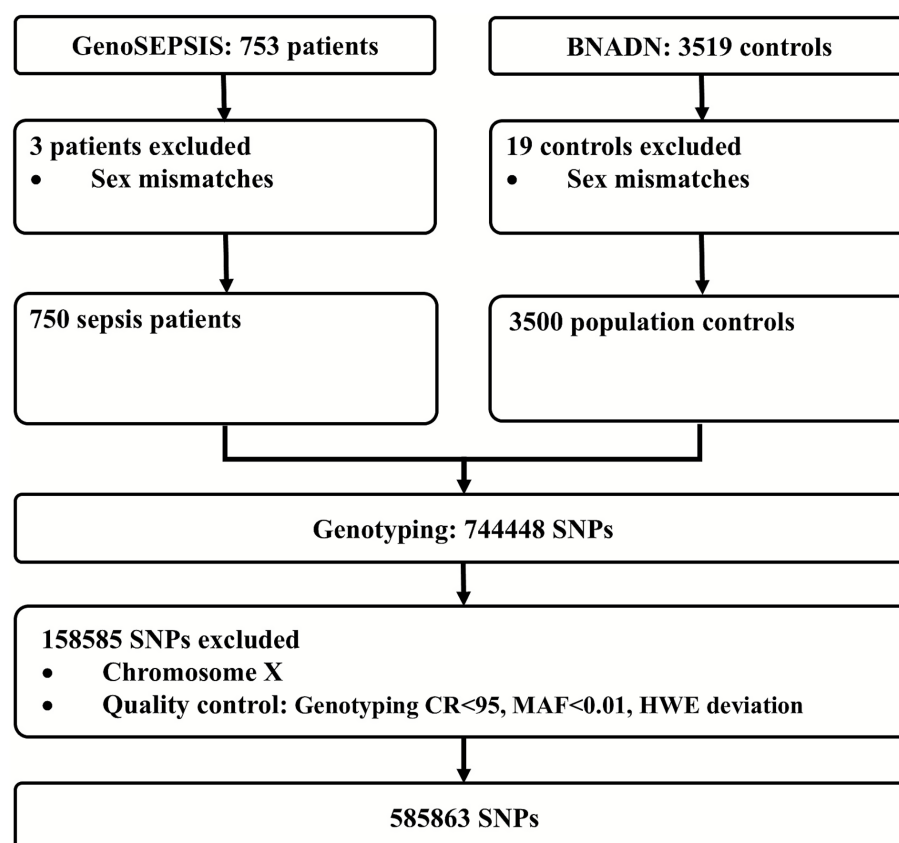


FIGURE 1
Study profile.

post-operative sepsis and analyze their biological and functional implications, thereby facilitating their interpretation.

2 Materials and methods

2.1 Study design and participants

We performed a prospective cohort study using a sepsis patient cohort (GenoSEPSIS) and a population cohort from the Spanish National DNA Bank (BNADN) (Figure 1). The patient cohort GenoSEPSIS included 753 adult patients who underwent major surgery, admitted to two intensive care units (ICUs) in Spain from Hospital Clínico Universitario de Valladolid (HCUV) and Hospital Clínico Universitario de Santiago (CHUS) from November 2004 to December 2016. All patients were on mechanical ventilation and did not have any infection prior to surgery. Patients fulfilled the diagnosis of sepsis or septic shock according to SEPSIS-3 criteria (1) and the DNA extraction was performed within the first 24 h after the diagnosis of sepsis. Details of patient management and treatment, as well as data collection and follow-up, are included in the [Supplementary methods](#) and were also described for HCUV patients by Martín-Fernández et al. (30).

This study followed current Spanish legislation on biomedical research and the Declaration of Helsinki. Written informed consent was obtained from all participants or their representatives. The study was

approved by the Ethics Committees for Clinical Research at participating centers (#No. PI 20–2070). The control cohort included genetic and demographic (sex and age) data available from 3,519 subjects from the BNADN, University of Salamanca, Spain¹ and have been used in recent GWAS (31, 32). Subjects from the BNADN were unrelated individuals, uniformly distributed throughout different geographical areas of Spain, and lacking personal or family history of clinical conditions such as infectious diseases, cancer, circulatory disorders, endocrine issues, mental or behavioral disorders, as well as diseases affecting the nervous, visual, auditory, respiratory, and immune systems, among others. This cohort allows us to identify genetic variants that may be exclusively associated with sepsis predisposition, without the confounding influence of other comorbidities linked to critical illness.

2.2 SNP genotyping and preprocessing

A general scheme of the methods used in this study is presented in the [Supplementary Figure S1](#). DNA samples from GenoSEPSIS and BNADN were genotyped at Centro Nacional de Genotipado-Universidad de Santiago de Compostela (CeGen-USC) using the Axiom Spain Biobank Array (Thermo Fisher Scientific).

¹ <http://www.bancoadn.org>

Genotyping quality control (QC) and filtering procedures are described in the [Supplementary methods](#). A total of 585,863 SNPs from 750 sepsis patients and 3,500 population controls were obtained after QC analyses. Association analysis between SNP genotypes and sepsis was performed by PLINK 1.9, adjusting by age, sex, and the first two principal components (33) ([Supplementary Figure S2](#)). Different subsets of relevant SNPs were selected according to several thresholds of p -value (5×10^{-2} , 5×10^{-3} , 5×10^{-4} , 5×10^{-5} , 5×10^{-6} , 5×10^{-7} , and 5×10^{-8}). See [Supplementary methods](#) for details.

2.3 XAI analysis and biological interpretation

XAI analysis was performed in two steps ([Supplementary Figure S1](#)). In the first step, a deep-learning model was designed to accurately predict sepsis from each subset of relevant SNPs. Specifically, a convolutional neural network (CNN) architecture, which has previously shown its usefulness in analyzing GWAS data (26, 27), was trained to automatically detect sepsis using previously subsets of selected SNPs ([Supplementary Figure S3](#)). To train the CNN, the whole dataset (4,250 samples) was randomly divided into training (50%), validation (25%), and test (25%) sets. The ratios of sepsis/control cases in these cohorts remained similar. See [Supplementary methods](#) for further details.

After obtaining the deep-learning model for sepsis prediction, the second step was the prioritization of sepsis-related SNPs. We applied the Deep SHAP XAI technique to obtain the SHAP values (34), which measure, for each patient, the contribution of each SNP to the prediction of sepsis. To find the most important SNPs contributing to sepsis, we took the average SHAP values (in absolute magnitude) from all patients accurately predicted as sepsis in the test set for each SNP. See [Supplementary Figure S4](#) for further details.

For the top SNPs with the highest SHAP values for sepsis prediction, we assessed the functional *in silico* effects based on empirical data from different integrated software tools and datasets, and their association with the clinical characteristics of the sepsis patient cohort. Finally, we also performed a gene enrichment analysis querying different databases for the top SNPs and related genes. See [Supplementary methods](#) for further details.

2.4 Statistical analysis

The performance of the trained CNN models to detect sepsis (i.e., population control vs. sepsis) was assessed by the sensitivity (Se, proportion of sepsis subjects rightly classified), specificity (Sp, proportion of control subjects rightly classified), accuracy (Acc, proportion of subjects rightly classified), area under the receiver operating characteristic (ROC) curve (AUC), and odd ratio (OR). PLINK 1.9 was used to perform association analysis between SNPs and sepsis phenotype, as well as between the top SNPs identified by the proposed XAI methodology and various clinical characteristics (including comorbidities, diagnostic measurements, sources of infection, disease progression and hospital outcomes). For the gene enrichment analysis, the z -score of the deviation from the expected rank by the Fisher exact test was computed to assess statistically significant associations.

3 Results

3.1 Patient's baseline characteristics

[Table 1](#) shows the demographics of the sepsis patients and controls, where sepsis cases represented 17.6% (750 patients) with a median age of 72 (61–78) years and 65.9% of the proportion of males. Regarding sepsis patients, 83.9% ($n = 629$) had septic shock and their associated 90-day mortality rate was 42.7% ($n = 320$). Median SOFA and APACHE II scores were 9 (IQR 7–11) and 18 (IQR 15–22), respectively. A total of 561 patients (74.8%) had one or several associated comorbidities, including chronic cardiovascular disease (257 cases, 34.3%), chronic respiratory disease (156 cases, 20.8%), arterial hypertension (318 cases, 42.4%), chronic renal failure (89 cases, 11.8%), chronic liver failure (43 cases, 5.7%), diabetes mellitus (166 cases, 22.1%), obesity (109 cases, 14.5%), and immunosuppression (73 cases, 9.7%). Peritonitis (228 cases, 30.4%), pneumonia (185 cases, 24.7%), catheter (62 cases, 8.3%), and surgical wound (20 cases, 2.7%) are the main causes of infection.

3.2 Identification of the most important SNPs for sepsis prediction

The sepsis prediction performance in the training, validation, and test sets of CNN models obtained with each subset of relevant SNPs are shown in [Supplementary Tables S1–S3](#) and [Supplementary Figure S5](#). Of note, the trained CNN model using SNPs with a p -value lower than 5×10^{-3} (3,761 SNPs) was considered the best model, achieving the highest accuracy in the validation (94.8%) set compared to CNN models derived using SNP subsets with p -values $< 5 \times 10^{-2}$, 5×10^{-4} , 5×10^{-5} , 5×10^{-6} , 5×10^{-7} , and 5×10^{-8} ([Supplementary Tables S1–S3](#)). Notably, this model also achieved the highest accuracy on the test set, with an accuracy of 96.4%, an AUC of 0.985, a sensitivity of 85.6%, a specificity of 98.7%, and an odds ratio of 465.99.

[Figure 2](#) shows the top 20 SNPs with the highest impact in the automatic prediction of sepsis in the whole test cohort, as determined using the contribution score (i.e., mean |SHAP value|). Notably, the sepsis prediction performance remains high using the top 20 SNPs (AUC = 0.951) or the top 3 SNPs (AUC = 0.886). The top 3 SNPs (rs17653532, rs1575081785, and rs74707084) had a high contribution to the detection of sepsis, with a considerably higher SHAP value (SHAP value > 0.04) compared to the other top-ranked SNPs. Further details of the contribution score for each SNP, as well as sepsis prediction performance for different subsets of top SNPs in the test set are shown in the [Supplementary Figure S6](#) and [Supplementary Table S4](#).

3.3 *In silico* functional, clinical, and biological interpretation

We analyzed the *in silico* functional effect of the top 20 SNPs identified by the proposed XAI approach. [Table 2](#) shows our findings from the 20 SNPs with the highest SHAP contribution value to the accurate prediction of sepsis. Among these variants, 11 were located within genes (two were missense, eight intronic, and one 3' untranslated region) ([Table 2](#)). The intronic variant (rs17653532) with

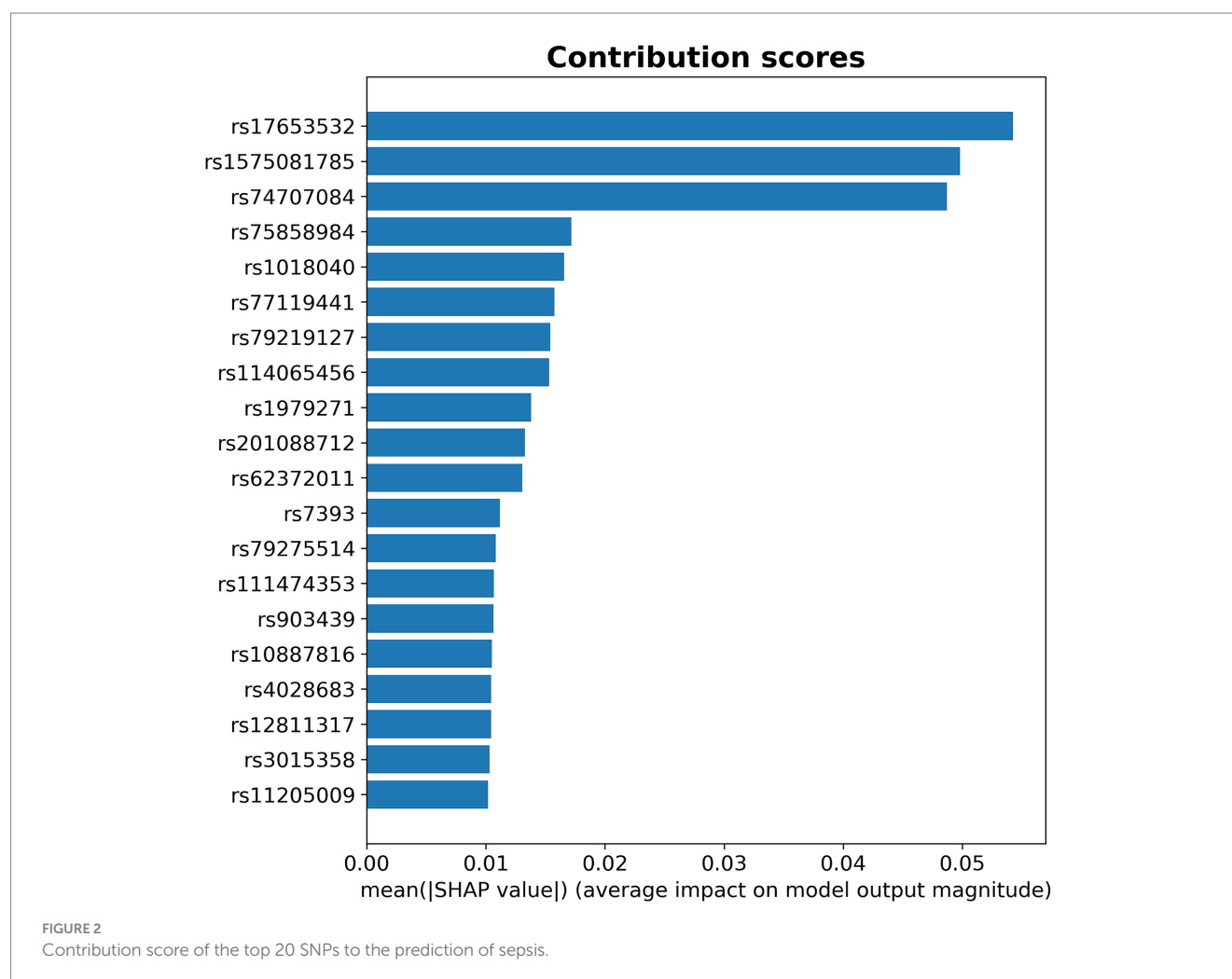
TABLE 1 Baseline and clinical characteristics of patients with sepsis and population controls.

	Sepsis cohort (GenoSEPSIS)	Control cohort (BNADN)
Demographics		
Sex, males [<i>n</i> (%)] ^a	494 (65.9)	1,903 (54.4)
Age at sample intake [years, median (IQR)] ^a	72 (61–78)	47 (41–54)
Hospital site [<i>n</i> (%)]		
Number of subjects [<i>n</i> (%)]	750 (17.6)	3,500 (82.4)
HCUV	389 (51.9)	NA
CHUS	361 (48.1)	NA
Comorbidities [<i>n</i> (%)]		
Chronic cardiovascular disease	257 (34.3)	NA
Chronic respiratory disease	156 (20.7)	NA
High blood pressure	318 (42.4)	NA
Chronic renal failure	89 (11.9)	NA
Chronic hepatic failure	43 (5.7)	NA
Diabetes mellitus	166 (22.1)	NA
Obesity	109 (14.5)	NA
Immunosuppression	73 (9.7)	NA
Measurements at diagnosis median [IQR]		
Creatinine (mg/dl)	1.7 (1.1–3.1)	NA
White blood cells (cells/mm ³)	12,685 (7,630–18,290)	NA
Lymphocytes (cells/mm ³)	7.1 (4.5–11.6)	NA
Neutrophils (cells/mm ³)	87.5 (81.6–91.3)	NA
SOFA score	9 (7–11)	NA
APACHE II score	18 (15–22)	NA
Source of infection [<i>n</i> (%)]		
Pneumonia	185 (24.7)	NA
Peritonitis	228 (30.4)	NA
Catheter	62 (8.3)	NA
Surgical site	20 (2.7)	NA
Others	245 (32.7)	NA
Time course and hospital outcomes		
Length of hospital stay [days, median (IQR)]	31 (19–44)	NA
Length of ICU stay [days, median (IQR)]	14 (7–21)	NA
Length of mechanical ventilation [days, median (IQR)]	9 (2–16)	NA
Severe sepsis [<i>n</i> (%)]	121 (16.1)	NA
Septic shock [<i>n</i> (%)]	629 (83.9)	NA
Mortality at 90 days [<i>n</i> (%)]	320 (42.7)	NA

^a*p*-value < 0.05 between GenoSEPSIS and BNADN cohorts. IQR, interquartile range (25 percentile–75 percentile); SOFA, Sequential Organ Failure Assessment; APACHE II, Acute Physiology And Chronic Health Evaluation II; ICU, intensive care unit; HCUV, Hospital Clínico Universitario de Valladolid; CHUS, Hospital Clínico Universitario de Santiago; NA, Not applicable.

the highest SHAP contribution score for the accurate prediction of sepsis (SHAP value = 0.054) was located in the gene encoding the DNA Primase Subunit 2 (*PRIM2*). The second variant with the highest contribution (SHAP value = 0.050) was a missense variant (rs1575081785) within the gene encoding the Rabenosyn RAB Effector (*RBSN*). The third variant with a SHAP value of 0.049 (rs74707084) was an intronic SNP located within the gene encoding the Synaptoporin (*SYNPR*). Among the top 20 SNPs, and after Bonferroni correction for multiple-tests (*p*-value < 0.0025),

rs79219127, intronic to the *FAM155A* gene, showed statistically significant associations with the length of hospital (*p*-value = 2.6×10^{-8}) and ICU stay (*p*-value = 7.7×10^{-4}) in sepsis patients, whereas rs79275514, intronic to the gene encoding the Parkin protein (*PARK2*), was statistically significant related with high blood pressure (*p*-value = 2.1×10^{-4}) and chronic hepatic failure (*p*-value = 3.2×10^{-5}) comorbidities (Supplementary Table S6). We also found that most top-ranked 20 SNPs showed evidence of biological and regulatory effects on multiple elements related to chromatin state,



changes in regulatory motifs, DNase I sensitivity, and expression quantitative trait loci (eQTLs) in different cell lines and tissues. More functional, clinical, and biological details are reported in the [Supplementary Table S5](#).

Based on the Gene Ontology annotation analysis, the most relevant and significantly enriched biological process was the negative regulation of heart contraction ($p_{\text{adjusted}} = 0.02$), involving two genes: Renalase, FAD Dependent Amine Oxidase (*RNLS*) and Phosphodiesterase 4D (*PDE4D*) ([Supplementary Figure S7](#)). We also found that the cyclic-nucleotide phosphodiesterase activity ($p_{\text{adjusted}} = 1.55 \times 10^{-3}$), the cyclic adenosine monophosphate (cAMP) binding ($p_{\text{adjusted}} = 1.55 \times 10^{-3}$), and the cyclic nucleotide binding ($p_{\text{adjusted}} = 3.91 \times 10^{-3}$) were the most significant enriched molecular process, which involved two genes (*PDE10A* and *PDE4D*) encoding phosphodiesterase proteins ([Supplementary Figure S8](#)). In terms of enrichment in the Jessen disease database, acrodysostosis ($p_{\text{adjusted}} = 0.068$) and dementia ($p_{\text{adjusted}} = 0.068$) were the most relevant diseases, involving *PDE4D*, Sortilin Related Receptor 1 (*SORL1*) and Parkin RBR E3 Ubiquitin Protein Ligase (*PARK2*) genes ([Supplementary Figure S9](#)). In the expression heatmap across different tissues, we observed that *RBSN*, Lysophospholipase Like 1 (*LYPLAL1*), and Zinc Finger Protein 775 (*ZNF775*) genes were highly expressed in

almost all tissues ([Supplementary Figure S10](#)). More details are reported in the [Supplementary results](#).

4 Discussion

To the best of our knowledge, this is the first XAI approach applied to GWAS-derived SNPs for patients with post-operative sepsis. We identified top-ranked SNPs with higher contribution to sepsis prediction involved in chromatin regulation, regulatory motifs, DNase I sensitivity, and significant eQTLs in different cell lines and tissues (including blood, fibroblasts, and immune response cells). Among the top20-ranked SNPs, we found SNPs associated with inflammation, blood cell count, and sepsis traits, and in sensitivity analysis of the sepsis cohort, we observed SNPs associated with hospital and ICU stay length, chronic liver failure, and hypertension. This exploratory study allowed the prioritization of three potential biomarkers with the highest contribution to post-surgical sepsis prediction (rs17653532, rs1575081785, and rs74707084) located in *PRIM2*, *RBSN*, and *SYNPR* genes, which are involved in integrative processes, such as gene expression regulation, DNA replication, and cell proliferation.

TABLE 2 Top 20 SNPs with the highest SHAP contribution to the accurate prediction of sepsis.

Chromosome	Position (Hg37)	SNP	Located	Nearest genes	p-value [^]	SHAP values*
1	218705814	rs1018040	Downstream	<i>TGFB2/</i> <i>LYPLAL1</i>	1.69E-05	0.017
1	152453338	rs11205009	Upstream	<i>LCE5A</i>	2.87E-06	0.010
3	15115369	rs1575081785	Missense	<i>RBSN</i>	4.98E-40	0.050
3	63389846	rs74707084	Intron	<i>SYNPR</i>	1.28E-65	0.049
3	29607405	rs1979271	Intron	<i>RBMS3</i>	2.44E-03	0.014
5	59757497	rs62372011	Intron	<i>PDE4D</i>	2.28E-10	0.013
6	57189167	rs17653532	Intron	<i>PRIM2</i>	1.18E-03	0.054
6	162119549	rs79275514	Intron	<i>PARK2</i>	8.30E-04	0.011
6	30901543	rs111474353	Upstream	<i>DPCR1</i>	6.08E-59	0.011
6	166147479	rs903439	Intron	<i>PDE10A</i>	1.31E-05	0.011
7	150095271	rs7393	UTR-3	<i>ZNF775</i>	9.72E-04	0.011
10	90178820	rs10887816	Intron	<i>RNLS</i>	1.72E-04	0.011
11	121787956	rs77119441	Downstream	<i>SORL1</i>	1.47E-16	0.016
12	33598920	rs12811317	Upstream	<i>SYT10</i>	1.66E-20	0.010
13	108050599	rs79219127	Intron	<i>FAM155A</i>	4.09E-33	0.015
13	25480913	rs201088712	Missense	<i>CENPJ</i>	3.49E-29	0.013
13	105888343	rs3015358	Downstream	<i>DAOA-AS1</i>	1.78E-03	0.010
14	87312206	rs114065456	Downstream	<i>FLRT2/</i> <i>LINC01148</i>	4.96E-32	0.015
15	25037506	rs4028683	Downstream	<i>NPAP1</i>	4.36E-03	0.010
18	55016814	rs75858984	Upstream	<i>ST8SIA3</i>	1.09E-24	0.017

*Based on the contribution on sepsis. [^]p-value resulted from the GWAS analysis. See information from 3,761 SNPs in the [Supplementary Table S2](#) extended.

The SNP with the highest SHAP contribution was an intronic variant located in the *PRIM2* gene involved in DNA replication (35), which is critical in sepsis for its role in apoptosis, oxidative stress, and metabolic changes. *PRIM2* has been related to consecutive trauma-induced sepsis based on an expression profiling analysis (36). The second SNP with the highest SHAP contribution score is in the *RBSN* gene, which encodes a protein from the FYVE zinc finger family and is involved in vesicle trafficking. Zinc finger proteins may contribute to inflammation, immune cell function, and tissue repair by modulating gene expression and regulating key immune-related genes (37–39). In the context of sepsis, these proteins may also affect cellular dysfunction, apoptosis, and DNA repair mechanisms, thereby affecting cell survival and tissue integrity (39). Similarly, a gene co-expression network analysis identified a zinc finger family gene (*ZNF721*) in a gene cluster for septic shock patients (40). Finally, a third gene, *SYNPR*, which encodes Synaptoporin, a protein found in the central nervous system, is involved in synaptic vesicle trafficking and neurotransmitter release (41).

In addition, our analyses revealed two genes encoding phosphodiesterases (*PDE10A* and *PDE4D*), involved in cyclic nucleotide signaling (42, 43). Cyclic nucleotide signaling is involved in several cellular processes, including immune response and inflammation (44), and phosphodiesterase inhibitors have shown potential therapeutic effects in experimental models of sepsis and lung inflammation (45, 46). In fact, cyclic-nucleotide phosphodiesterase activity and the cAMP binding were the most enriched relevant molecular processes in our study. Furthermore, the most significantly

enriched biological process was negative regulation of heart contraction, involving *RNLS* and *PDE4D* genes. Cardiac dysfunction is an important consequence of sepsis, caused by increased inflammation or suppression of fatty acid and glucose oxidation, or due to adenosine triphosphate (ATP) depletion (47, 48). Although we did not find any previously significant associated variants with sepsis in our study, some of the ranked SNPs (rs201088712, rs3015358, and rs114065456) were associated with white blood cell count, sepsis, and sepsis-associated death in the UK Biobank. Thus, we have identified new possible candidate genes associated with post-operative sepsis susceptibility, highlighting the role of genes related to gene expression, DNA replication, cyclic nucleotide signaling, cell proliferation, and cardiac dysfunction.

Comprehensive XAI approaches have identified clinical features (vital signs, laboratory values, or demographics, among others) from electronic health records (EHR) contributing to early detection of sepsis (49–51) but they have not been applied to sepsis-related genomic data until now. Several GWAS studies have identified SNPs and evaluated polygenic risk scores (PRS) associated with sepsis susceptibility and mortality (4, 12–17), offering insights that could inform early prevention and treatment strategies targeting sepsis-related complications. However, only Engoren et al. (12) described a polygenic risk score for sepsis susceptibility, achieving an AUC of 0.752. This sepsis prediction performance in Engoren et al. (12) could be attributed to: (i) including both sepsis-2 and sepsis-3 adult perioperative patients identified from EHR data; (ii) using perioperative controls identified from a university EHR, who may present

confounding comorbidities; (iii) applying a different sepsis prediction model than in our study (logistic regression *vs.* CNN), which may not capture complex SNP interactions contributing to sepsis predisposition. Regarding the methodology in existing GWAS for sepsis (4, 12–17), it is noteworthy that the SNPs most strongly associated with the phenotype (i.e., those with the lowest *p*-values) could not contribute to the most predictive genetic signature. Instead, the most predictive signature often consists of SNPs that provide complementary information (23). In this respect, our XAI methodology relying on a *p*-value $< 5 \times 10^{-3}$ threshold prioritized those SNPs that have a higher influence for phenotype prediction. Among the top 20-ranked SNPs, we found both SNPs with high (i.e., *p*-value from GWAS $< 5 \times 10^{-8}$) and intermediate specificity (i.e., *p*-value $> 5 \times 10^{-8}$), being one of the SNPs with intermediate specificity also in the top 3 SNPs (rs17653532 (*PRIM2*), GWAS *p*-value = 1.18×10^{-3}), thus confirming the limitations of using the standard threshold of 5×10^{-8} for statistical significance in GWAS studies (23). Importantly, our XAI approach prioritized three SNPs (rs17653532, rs1575081785, and rs74707084) with a high contribution to sepsis prediction (i.e., highest SHAP values) and good performance when validated in an independent test subset (AUC = 0.886), highlighting the potential role of these variants and related genes in sepsis risk in post-surgical patients. In clinical practice, the output probability threshold of the sepsis prediction models could be adjusted to identify patients at higher or lower risk of sepsis, prioritizing either sensitivity (to capture more high-risk patients) or specificity (to reduce false positives), depending on the intended clinical application and available resources (52). Our results agree with previous XAI approaches applied to analyze GWAS-derived SNPs in complex diseases, such as atrial fibrillation (26), attention deficit hyperactivity disorder (27), hypertension, or diabetes (28).

Although the *in silico* functional findings suggest potential implications for early and personalized prevention and treatment for post-operative sepsis, further biological validation of these exploratory findings through *in vitro* and *in vivo* analyses is required for clinical relevance and generalizability. XAI provides a powerful approach to prioritize SNPs for sepsis prediction, providing insights into cohort-wide and individual-specific genetic predispositions. We also highlighted the clinical and biological relevance of genes associated with the top 3 SNPs (*PRIM2*, *RBSN*, and *SYNPR*). Identifying these genetic variants in a preoperative blood test could help in the early detection of sepsis in surgical patients, may enable risk stratification, and allow prompt pharmacological treatment, thus reducing mortality and long-term disability. This concept is in line with the recent clinical study by Liesenfeld et al. (53), who clinically validated an AI-driven blood test using host mRNAs expression data to predict acute infection and sepsis. However, our findings remain exploratory and further external clinical validation is crucial to confirm their reproducibility and generalizability of our findings. Studies in ethnically diverse cohorts are also warranted, given that most existing sepsis GWAS have been conducted in populations of European ancestry and genetic associations often fail to replicate across ancestral group (54, 55).

We acknowledge some strengths and limitations of our study. The existing GWAS in sepsis patients include patients with sepsis of any etiology (4, 12–17), which leads to the identification of genes that differ among studies (4, 12–17). In contrast, we focus exclusively on cases based on patients with postsurgical sepsis, making our

population highly homogeneous and with applicable results within this context. Nevertheless, patient heterogeneity persists due to differences in surgical procedures, infection sources, or host responses (56–58). Thus, future studies are warranted to investigate whether specific post-operative patient subgroups exhibit distinct genetic associations or clinical outcomes. Apart from this, the study samples were collected from two hospitals, and the sample size was not very large due to the challenging nature of data collection, and rare variants in or near identified regions may go undetected due to technological limitations. In this respect, the use of a test set derived from the same underlying cohort represents a potential limitation regarding external generalizability. However, the adoption of a hold-out validation scheme (with independent training, validation, and test subsets) provides a robust internal assessment of model performance in the absence of an available external or independent postsurgical sepsis cohort, that reduces the risk of overfitting. This validation approach has been shown to yield generalizable, reproducible and biologically meaningful results in recent AI-based genomic studies (26–28). Nevertheless, the validation in a large, independent, geographically, and ancestrally distinct dataset of post-operative patients with diverse ancestries would be a necessary future step to confirm our findings. This would also allow consideration of additional factors such as comorbidities, measurements at diagnosis, and sources of infection in sepsis prediction. Genome and whole exome sequencing analyses would also provide a better resolution to achieve this goal. Regarding the proposed XAI methodology, we used a CNN for sepsis prediction, originally designed for image analysis. Recent studies have shown that CNNs are suitable to analyze GWAS data (26, 27). Another limitation is the use of population controls who were not clinically evaluated for sepsis. As such, we cannot entirely exclude the possibility that some controls could have been cases exposed to relevant environment factors or survivors of a previous sepsis episode. However, the use of population controls provided access to a well-characterized and substantial sample size cohort, thereby increasing the statistical power of the study, and they are commonly utilized in large-scale GWAS studies of infectious diseases (31, 32, 59–61). Notably, the COVID-19 Host Genetics Initiative consortium studies demonstrated that genetic analyses using various comparisons of controls with COVID-19 severity (e.g., infection, hospitalization, critical illness) yield overlapping results, supporting the validity of using population controls in genetic research on infectious diseases (59). Similarly, several GWAS of infectious conditions have successfully relied on population-based biobanks as control sources when disease-specific non-affected cohorts were unavailable, highlighting the practical and methodological acceptance of this approach in complex traits (31, 32, 60, 61). Conversely, although this has been observed in other studies, and we adjusted for sex and age in the association analysis to reduce bias due to age imbalance between cases and controls, young patients may eventually undergo surgery and develop sepsis. Moreover, the genotype by environment interactions may show challenges in this kind of study due to the dynamic influence of the environment on gene expression. As a result, the effects of genetic variants could be masked or modified by environmental factors. Thus, future studies should include clinical validation in postsurgical or critically ill patients without sepsis as controls once reliable genetic data is available, to better delineate genetic variants specifically associated with sepsis susceptibility within the context of critical illness.

In conclusion, our XAI approach applied to GWAS-derived SNPs enabled the identification of significant risk loci associated with post-surgical sepsis that could be implemented in clinical practice for improving patient outcomes. We found variants with functional, regulatory and clinical implications, as well as genes related to gene expression regulation, DNA replication, cyclic nucleotide signaling and cell proliferation, and cardiac dysfunction, among other biological processes. We also identified three potential biomarkers with the highest contribution to sepsis prediction (rs17653532, rs1575081785, and rs74707084), located in *PRIM2*, *RBSN* and *SYNPR* genes, which could be determined in a preoperative blood test, allowing targeted and precise interventions to prevent and treat sepsis in patients undergoing surgery. Further investigations, including *in vitro* and *in vivo* analyses, as well as complementary studies in cohorts comprising sepsis and non-sepsis patients undergoing major surgery will be needed to optimally evaluate the genetic factors contributing to sepsis predisposition and to provide an external validation of our exploratory findings.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: GWAS are not publicly available, but it can be obtained upon reasonable request from the authors. Requests to access these datasets should be directed to Eduardo Tamayo, eduardo.tamayo@uva.es.

Ethics statement

The studies involving humans were approved by the Ethics Committee for Clinical Research, Hospital Clínico Universitario in Valladolid, Spain (approval no. PI 20-2070) and the Ethics Committee for Clinical Research, Hospital Clínico Universitario de Santiago, Spain (approval No. PI 20-2070). Written informed consent was obtained from patients, patients' relatives, or their legal representatives before their enrolment. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

FV-V: Validation, Methodology, Writing – review & editing, Visualization, Software, Investigation, Writing – original draft, Formal analysis, Funding acquisition. TH-B: Software, Writing – review & editing, Investigation, Writing – original draft, Visualization, Data curation, Methodology, Validation. MH-R: Writing – review & editing, Funding acquisition, Writing – original draft, Data curation. EG-S: Writing – review & editing, Funding acquisition, Writing – original draft, Data curation. ML-L: Writing – review & editing, Writing – original draft, Data curation. RL-H: Writing – original draft, Data curation, Writing – review & editing. MB-C: Data curation, Writing – review & editing, Writing – original draft. AT-V: Writing – review & editing, Writing – original draft, Data curation. MM-F: Data curation, Writing – review & editing, Writing – original draft. LS-de-P: Writing – original draft, Data curation, Writing – review & editing. JÁ-E: Writing – review & editing, Writing – original draft, Data curation. SV: Data curation, Writing – review & editing, Writing

– original draft. AB: Writing – original draft, Writing – review & editing, Data curation. HG-B: Writing – review & editing, Data curation, Writing – original draft. PM-P: Visualization, Data curation, Writing – review & editing, Writing – original draft. AG-C: Writing – review & editing, Writing – original draft, Data curation. AF-R: Writing – review & editing, Formal analysis, Funding acquisition, Visualization, Writing – original draft. MJ-S: Formal analysis, Writing – original draft, Writing – review & editing, Visualization. SR: Writing – review & editing, Visualization, Formal analysis, Writing – original draft, Funding acquisition. LM-C: Writing – original draft, Formal analysis, Writing – review & editing, Methodology. ES-P: Writing – original draft, Formal analysis, Methodology, Writing – review & editing. IQ: Writing – review & editing, Writing – original draft, Formal analysis, Methodology. RC: Formal analysis, Writing – review & editing, Writing – original draft, Methodology. ÁC: Methodology, Writing – original draft, Writing – review & editing, Funding acquisition, Formal analysis. JV: Funding acquisition, Visualization, Writing – original draft, Writing – review & editing. CF: Writing – original draft, Visualization, Funding acquisition, Methodology, Writing – review & editing. RH: Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Visualization, Investigation, Conceptualization, Resources. ET: Investigation, Conceptualization, Visualization, Supervision, Funding acquisition, Resources, Project administration, Data curation, Writing – review & editing, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by 'Instituto de Salud Carlos III (ISCIII)' PI18/01238, PI19/00141, PI20/00876, PI23/00980, and PI23CIII/00010, by 'Consorcio Centro de Investigación Biomédica en Red (CIBER) en Enfermedades Respiratorias (CIBERES)' (CB06/06/1088 and AC_212/00039), by 'CIBER en Enfermedades Infecciosas (CIBERINFEC)' (CB21/13/00051, CB21/13/00044, and IM23/INFEC/1), by 'CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)' (CB19/01/00012), by 'CIBER en Enfermedades Raras (CIBERER)' (CB06/07/0088), by 'Junta de Castilla y León' (VA321P18, GRS 1922/A/19, GRS 2057/A/19, GRS 2425/A/21), by 'Fundación Ramón Areces' (CIVP19A5953), by ERA PerMed (JTC_2021) by the contract AC21_2/00039 with Instituto de Salud Carlos III and funds from Next Generation EU as part of the actions of the Recovery Mechanism and Resilience (MRR), by ITER agreements (OA17/008 and OA23/043), and by 'Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación/10.13039/501100011033/, ERDF A way of making Europe, and NextGenerationEU/PRTR (PID2023-148895OB-I00). FV-V is supported by a 'Sara Borrell' grant (CD23/00031) from ISCIII cofounded by the 'Fondo Social Europeo Plus (FSE+)'. ES-P was supported by 'Agencia Canaria de Investigación, Innovación y Sociedad de la Información de la Consejería de Economía, Conocimiento y Empleo y por el Fondo Social Europeo (FSE) Programa Operativo Integrado de Canarias 2014–2020, Eje 3 Tema Prioritario 74 (85%) Gobierno de Canarias, Social European Fund "Canarias Avanza con Europa" (TESIS202201004)." JV is supported by the European Regional Development Funds, Fundación Canaria Instituto de Investigación Sanitaria de Canarias, Spain (PIFIISC24/22) and Asociación Científica Pulmón y Ventilación Mecánica, Las Palmas de Gran Canaria, Spain.

Acknowledgments

This study has been possible thanks to the collaboration of all patients and their relatives, and the medical and nursery staff of participating clinical departments for their continuous support. The authors thank data managers who have taken part in the project.

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 authors declare that no Gen 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

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/fmed.2025.1644800/full#supplementary-material>

References

- Singer, M, Deutschman, CS, Seymour, C, Shankar-Hari, M, Annane, D, Bauer, M, et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA*. (2016) 315:801–10. doi: 10.1001/jama.2016.0287
- Fleischmann, C, Scherag, A, Adhikari, NKJ, Hartog, CS, Tsaganos, T, Schlattmann, P, et al. Assessment of global incidence and mortality of hospital-treated sepsis current estimates and limitations. *Am J Respir Crit Care Med*. (2016) 193:259–72. doi: 10.1164/rccm.201504-0781OC
- Fleischmann-Struzek, C, Mellhammar, L, Rose, N, Cassini, A, Rudd, KE, Schlattmann, P, et al. Incidence and mortality of hospital- and ICU-treated sepsis: results from an updated and expanded systematic review and meta-analysis. *Intensive Care Med*. (2020) 46:1552–62. doi: 10.1007/s00134-020-06151-x
- Hernandez-Beetink, T, Guillen-Guio, B, Lorenzo-Salazar, JM, Corrales, A, Suarez-Pajes, E, Feng, R, et al. A genome-wide association study of survival in patients with sepsis. *Crit Care*. (2022) 26:341–10. doi: 10.1186/s13054-022-04208-5
- Iwashyna, TJ, Ely, EW, Smith, DM, and Langa, KM. Long-term cognitive impairment and functional disability among survivors of severe sepsis. *JAMA*. (2010) 304:1787–94. doi: 10.1001/jama.2010.1553
- Levy, MM, Evans, LE, and Rhodes, A. The surviving Sepsis campaign bundle: 2018 update. *Intensive Care Med*. (2018) 44:925–8. doi: 10.1007/s00134-018-5085-0
- Herran-Monge, R, Muriel-Bombin, A, Garcia-Garcia, MM, Merino-Garcia, PA, Martinez-Barrios, M, Andaluz, D, et al. Epidemiology and changes in mortality of sepsis after the implementation of surviving sepsis campaign guidelines. *J Intensive Care Med*. (2019) 34:740–50. doi: 10.1177/0885066617711882
- Skibsted, S, Bhasin, MK, Aird, WC, and Shapiro, NI. Bench-to-bedside review: future novel diagnostics for sepsis – a systems biology approach. *Crit Care*. (2013) 17:231. doi: 10.1186/cc12693
- Kumar, S, Tripathy, S, Jyoti, A, and Singh, SG. Recent advances in biosensors for diagnosis and detection of sepsis: a comprehensive review. *Biosens Bioelectron*. (2019) 124:125:205–15. doi: 10.1016/j.bios.2018.10.034
- Schinkel, M, Paranjape, K, RSN, P, Skyttberg, N, and PWB, N. Clinical applications of artificial intelligence in sepsis: a narrative review. *Comput Biol Med*. (2019) 115:103488. doi: 10.1016/j.combiomed.2019.103488
- Wu, M, Du, X, Gu, R, and Wei, J. Artificial intelligence for clinical decision support in sepsis. *Front Med*. (2021) 8:1–9. doi: 10.3389/fmed.2021.665464
- Engoren, M, Jewell, ES, Douville, N, Moser, S, Maile, MD, and Bauer, ME. Genetic variants associated with sepsis. *PLoS One*. (2022) 17:e0265052. doi: 10.1371/journal.pone.0265052
- Guillen-Guio, B, Lorenzo-Salazar, JM, Ma, SF, Hou, PC, Hernandez-Beetink, T, Corrales, A, et al. Sepsis-associated acute respiratory distress syndrome in individuals of European ancestry: a genome-wide association study. *Lancet Respir Med*. (2020) 8:258–66. doi: 10.1016/S2213-2600(19)30368-6
- Scherag, A, Schöneweck, F, Kesselmeier, M, Taudien, S, Platzer, M, Felder, M, et al. Genetic factors of the disease course after Sepsis: a genome-wide study for 28 day mortality. *EBioMedicine*. (2016) 12:239–46. doi: 10.1016/j.ebiom.2016.08.043
- Rautanen, A, Mills, TC, Gordon, AC, Hutton, P, Steffens, M, Nuamah, R, et al. Genome-wide association study of survival from sepsis due to pneumonia: an observational cohort study. *Lancet Respir Med*. (2015) 3:53–60. doi: 10.1016/S2213-2600(14)70290-5
- Rosier, F, Brisebarre, A, Dupuis, C, Baaklini, S, Puthier, D, Brun, C, et al. Genetic predisposition to the mortality in septic shock patients: from GWAS to the identification of a regulatory variant modulating the activity of a CISH enhancer. *Int J Mol Sci*. (2021) 22:5852. doi: 10.3390/ijms22115852
- D'Urso, S, Rajbhandari, D, Peach, E, De Guzman, E, Li, Q, Medland, SE, et al. Septic shock: a genomewide association study and polygenic risk score analysis. *Twin Res Hum Genet*. (2020) 23:204–13. doi: 10.1017/thg.2020.60
- Anderson, RN, and Smith, BL. Deaths: leading causes for 2002. *Natl Vital Stat Rep*. (2005) 53:1–89.
- van den Berg, M, van Beuningen, FE, ter Maaten, JC, and Bouma, HR. Hospital-related costs of sepsis around the world: a systematic review exploring the economic burden of sepsis. *J Crit Care*. (2022) 71:154096. doi: 10.1016/j.jccr.2022.154096
- Higgins, AM, Brooker, JE, Mackie, M, Cooper, DJ, and Harris, AH. Health economic evaluations of sepsis interventions in critically ill adult patients: a systematic review. *J Intensive Care*. (2020) 8:5. doi: 10.1186/s40560-019-0412-2
- Vogel, TR, Dombrovskiy, VY, Carson, JL, Graham, AM, and Lowry, SF. Postoperative Sepsis in the United States. *Ann Surg*. (2010) 252:1065–71. doi: 10.1097/SLA.0b013e3181dcf36e
- Villar, J, Maca-meyer, N, Pérez-méndez, L, and Flores, C. Bench-to-bedside review: understanding genetic predisposition to sepsis. *Crit Care*. (2004) 8:180–9. doi: 10.1186/cc2863
- Lakiotaki, K, Papadovasilakis, Z, Lagani, V, Fafalios, S, Charonyktakis, P, Tsagris, M, et al. Automated machine learning for genome wide association studies. *Bioinformatics*. (2023) 39:1–12. doi: 10.1093/bioinformatics/btad545
- Sandoval-Motta, S, Aldana, M, Martínez-Romero, E, and Frank, A. The human microbiome and the missing heritability problem. *Front Genet*. (2017) 8:1–12. doi: 10.3389/fgene.2017.00080
- Yang, G, Ye, Q, and Xia, J. Unbox the black-box for the medical explainable AI via multi-modal and multi-Centre data fusion: a mini-review, two showcases and beyond. *Inf Fusion*. (2022) 77:29–52. doi: 10.1016/j.inffus.2021.07.016
- Kwon, OS, Hong, M, Kim, TH, Hwang, I, Shim, J, Choi, EK, et al. Genome-wide association study-based prediction of atrial fibrillation using artificial intelligence. *Open Hear*. (2022) 9:1–10. doi: 10.1136/openhrt-2021-001898
- Liu, L, Feng, X, Li, H, Cheng, Li, S, Qian, Q, and Wang, Y. Deep learning model reveals potential risk genes for ADHD, especially Ephrin receptor gene EPHA5. *Brief Bioinform*. (2021) 22:1–11. doi: 10.1093/bib/bbab207
- Mieth, B, Rozier, A, Rodriguez, JA, Höhne, MMC, Görnitz, N, and Müller, KR. DeepCOMBI: explainable artificial intelligence for the analysis and discovery in genome-wide association studies. *NAR Genom Bioinform*. (2021) 3:1–21. doi: 10.1093/nargab/lqab065

29. Huang, K, Zeng, T, Koc, S, Pettet, A, Zhou, J, Jain, M, et al. Small-cohort GWAS discovery with AI over massive functional genomics knowledge graph. *medRxiv*. (2024). doi: 10.1101/2024.12.03.24318375
30. Martín-Fernández, M, Heredia-Rodríguez, M, González-Jiménez, I, Lorenzo-López, M, Gómez-Pesquera, E, Poves-Álvarez, R, et al. Hyperoxemia in postsurgical sepsis/septic shock patients is associated with reduced mortality. *Crit Care*. (2022) 26:4–9. doi: 10.1186/s13054-021-03875-0
31. Cruz, R, Diz-de Almeida, S, López de Heredia, M, Quintela, I, Ceballos, FC, Pita, G, et al. Novel genes and sex differences in COVID-19 severity. *Hum Mol Genet*. (2022) 31:3789–806. doi: 10.1093/hmg/ddac132
32. Suarez-Pajes, E, Marcelino-Rodríguez, I, Hernández Brito, E, Gonzalez-Barbuzano, S, Ramirez-Falcon, M, Tosco-Herrera, E, et al. A genome-wide association study of adults with community-acquired pneumonia. *Respir Res*. (2024) 25:374. doi: 10.1186/s12931-024-03009-4
33. Chang, CC, Chow, CC, Tellier, LCAM, Vattikuti, S, Purcell, SM, and Lee, JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. (2015) 4:s13742–015. doi: 10.1186/s13742-015-0047-8
34. Lundberg, SM, and Lee, S-I. A unified approach to interpreting model predictions. *31st Conference on Neural Information Processing Systems (NIPS 2017)*. Long Beach, CA, USA (2017). p. 1208–1217
35. Chung, J, Tsai, S, James, AH, Thames, BH, Shytle, S, and Piedrahita, JA. Lack of genomic imprinting of DNA primase, polypeptide 2 (PRIM2) in human term placenta and white blood cells. *Epigenetics*. (2012) 7:429–31. doi: 10.4161/epi.19777
36. Dong, L, Li, H, Zhang, S, and Su, L. Identification of genes related to consecutive trauma-induced sepsis via gene expression profiling analysis. *Medicine (Baltimore)*. (2018) 97:e0362. doi: 10.1097/MD.00000000000010362
37. Maruyama, K, Kidoya, H, Takemura, N, Sugisawa, E, Takeuchi, O, Kondo, T, et al. Zinc finger protein St18 protects against septic death by inhibiting VEGF-A from macrophages. *Cell Rep*. (2020) 32:107906. doi: 10.1016/j.celrep.2020.107906
38. Cassandri, M, Smirnov, A, Novelli, F, Pitolli, C, Agostini, M, Malewicz, M, et al. Zinc-finger proteins in health and disease. *Cell death Discov*. (2017) 3:1–12. doi: 10.1038/cddiscovery.2017.71
39. Rakhra, G, and Rakhra, G. Zinc finger proteins: insights into the transcriptional and post transcriptional regulation of immune response. *Mol Biol Rep*. (2021) 48:5735–43. doi: 10.1007/s11033-021-06556-x
40. Martínez-Paz, P, Gomez-Pilar, J, Martín-Fernández, M, Ceballos, FC, Gómez-Sánchez, E, Hornero, R, et al. Gene co-expression networks offer new perspectives on sepsis pathophysiology. *IEEE ACM Trans Comput Biol Bioinforma*. (2023) 20:3660–8. doi: 10.1109/TCBB.2023.3309998
41. Knaus, P, Marquèze-Pouey, B, Scherer, H, and Betzt, H. Synaptoporin, a novel putative channel protein of synaptic vesicles. *Neuron*. (1990) 5:453–62. doi: 10.1016/0896-6273(90)90084-S
42. Maurice, DH, Ke, H, Ahmad, F, Wang, Y, Chung, J, and Manganiello, VC. Advances in targeting cyclic nucleotide phosphodiesterases. *Nat Rev Drug Discov*. (2014) 13:290–314. doi: 10.1038/nrd4228
43. Bender, AT, and Beavo, JA. Cyclic nucleotide phosphodiesterases: molecular regulation to clinical use. *Pharmacol Rev*. (2006) 58:488–520. doi: 10.1124/pr.58.3.5
44. Raker, VK, Becker, C, and Steinbrink, K. The cAMP pathway as therapeutic target in autoimmune and inflammatory diseases. *Front Immunol*. (2016) 7:123. doi: 10.3389/fimmu.2016.00123
45. Hsu, CG, Fazal, F, Rahman, A, Berk, BC, and Yan, C. Phosphodiesterase 10A is a key mediator of lung inflammation. *J Immunol*. (2021) 206:3010–20. doi: 10.4049/jimmunol.2001026
46. Kazmi, I, Al-Abbasi, FA, Afzal, M, Nadeem, MS, Altayb, HN, and Gupta, G. Phosphodiesterase-4 inhibitor Roflumilast-mediated protective effect in Sepsis-induced late-phase event of acute kidney injury: a narrative review. *Pharmaceuticals*. (2022) 15:899. doi: 10.3390/ph15070899
47. Zaky, A, Deem, S, Bendjelid, K, and Treggiari, MM. Characterization of cardiac dysfunction in sepsis: an ongoing challenge. *Shock*. (2014) 41:12–24. doi: 10.1097/SHK.0000000000000065
48. Drosatos, K, Lymperopoulos, A, Kennel, PJ, Pollak, N, Schulze, PC, and Goldberg, IJ. Pathophysiology of sepsis-related cardiac dysfunction: driven by inflammation, energy mismanagement, or both? *Curr Heart Fail Rep*. (2015) 12:130–40. doi: 10.1007/s11897-014-0247-z
49. Yang, M, Liu, C, Wang, X, Li, Y, Gao, H, Liu, X, et al. An explainable artificial intelligence predictor for early detection of Sepsis. *Crit Care Med*. (2020) 48:E1091–6. doi: 10.1097/CCM.0000000000004550
50. Strickler, EAT, Thomas, J, Thomas, JP, Benjamin, B, and Shamsuddin, R. Exploring a global interpretation mechanism for deep learning networks when predicting sepsis. *Sci Rep*. (2023) 13:3067–17. doi: 10.1038/s41598-023-30091-3
51. Rosnati, M, and Fortuin, V. MGP-AttTCN: an interpretable machine learning model for the prediction of sepsis. *PLoS One*. (2021) 16:e0251248. doi: 10.1371/journal.pone.0251248
52. Buda, M, Maki, A, and Mazurowski, MA. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw*. (2018) 106:249–59. doi: 10.1016/j.neunet.2018.07.011
53. Liesenfeld, O, Arora, S, Aufderheide, TP, Clements, CM, DeVos, E, Fischer, M, et al. Clinical validation of an AI-based blood testing device for diagnosis and prognosis of acute infection and sepsis. *Nat Med*. (2025):1–11. doi: 10.1038/s41591-025-03933-y
54. Brinkworth, JF, and Shaw, JG. On race, human variation, and who gets and dies of sepsis. *Am J Biol Anthropol*. (2022) 178:230–55. doi: 10.1002/ajpa.24527
55. Minejima, E, and Wong-beringer, A. Impact of socioeconomic status and race on Sepsis epidemiology and outcomes. *J Appl Lab Med*. (2021) 6:194–209. doi: 10.1093/jalm/jfaa151
56. Stortz, JA, Cox, MC, Hawkins, RB, Ghita, GL, Brumback, BA, Mohr, AM, et al. Phenotypic heterogeneity by site of infection in surgical sepsis: a prospective longitudinal study. *Crit Care*. (2020) 24:203. doi: 10.1186/s13054-020-02917-3
57. Leligdowicz, A, and Matthay, MA. Heterogeneity in sepsis: new biological evidence with clinical applications. *Crit Care*. (2019) 23:80. doi: 10.1186/s13054-019-2372-2
58. Yang, J, Zhang, B, Hu, C, Jiang, X, Shui, P, Huang, J, et al. Laparoscopic, endoscopic and robotic surgery Identification of clinical subphenotypes of sepsis after laparoscopic surgery. *Laparosc Endosc Robot Surg*. (2024) 7:16–26. doi: 10.1016/j.lers.2024.02.001
59. Niemi, MEK, Karjalainen, J, Liao, RG, Neale, BM, Daly, M, Ganna, A, et al. Mapping the human genetic architecture of COVID-19. *Nature*. (2021) 600:472–7. doi: 10.1038/s41586-021-03767-x
60. Wu, P, Ding, L, Li, X, Liu, S, Cheng, F, He, Q, et al. Trans-ethnic genome-wide association study of severe COVID-19. *Commun Biol*. (2021) 4:1034. doi: 10.1038/s42003-021-02549-5
61. Degenhardt, F, Ellinghaus, D, Juzenas, S, Lerga-Jaso, J, Wendorff, M, Maya-Miles, D, et al. Detailed stratified GWAS analysis for severe COVID-19 in four European populations. *Hum Mol Genet*. (2022) 31:3945–66. doi: 10.1093/hmg/ddac158