Infection Biomarkers at Intensive Care Units

Rúben Araújo, Luís Ramalhete, Tiago A. H. Fonseca, Cristiana P. Von Rekowski, Luís Bento and Cecília R.C. Calado

***Abstract*—** It is relevant to discover infection biomarkers, especially for critically ill patients in intensive care units (ICU), as these patients often present non-infectious inflammatory processes that obscure typical infectious markers. This study focused on 20 ICU patients, half of whom had acquired bacterial blood infections (bacteremia). Due to the significance of inflammatory processes in these patients, it was evaluated how 21 serum cytokines could be used to develop predictive models for bacteremia. Feature selection using a Gain Information algorithm allowed for the construction of an excellent Naïve Bayes model, achieving an AUC of 0.950. These promising results strongly support future studies with larger cohorts, to further evaluate these types of platforms for infection diagnosis in such critical populations.

***Keywords*—**Infection biomarkers; Intensive Care Unit; Cytokines.

I. Introduction

The adequate diagnosis of infection can be difficult in critically ill patients, due to the common confound processes, such as non-infectious inflammatory events [1]. Due to the fragile metabolic state of these patients, it is also relevant to avoid burdening the patients with unnecessary antibiotics. It is therefore urgent to discover new biomarkers of infection applicable for these high-risk patients. From the intensive care unit (UCI) acquired infections, bloodstream infections are predicted to occur approximately in 5% of all ICU patients, being a frequent cause of severe sepsis [2,3]. Indeed, acquired bloodstream infections in UCU are associated with high mortality rates, which can lead to 40% increased risk of 30-day mortality [4]. For these reasons, the present work focused on blood-acquired infections in ICU patients caused by bacteria, *i.e.* bacteremia.

Cytokine storms significantly influence systemic inflammatory responses and septic patient outcome, highlighting the critical role of cytokine profiling in understanding and predicting patient's outcome [5]. Due to the relevance of cytokines, as mediators of the immune system response on infectious and non-infectious inflammatory processes [6], the present work evaluated if serum cytokines can be used to identify bacteremia. For that, it was considered the analysis of a set of 21 serum cytokines in 20 critically ill patients, admitted at ICU, with half of them having presented bacteremia as confirmed in the respective hemocultures.

Manuscript received: October 2024.

R. Araújo, C. Rekowski and T. Fonseca are from ISEL- Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa, Lisbon, Portugal; NMS-Nova Medical School, Universidade Nova de Lisboa, Lisbon, Portugal; and CHRC - Comprehensive Health Research Centre, Universidade Nova.

L. Ramalhate is from ISEL, NMS and IPST – Instituto Português do Sangue e da Transplantação, Lisbon, Portugal

L. Bento is from Intensive Care Department, ULSSJ - Unidade Local de Saúde de São José, Lisbon, Portugal and from Integrated Pathophysiological Mechanisms, CHRC.

C. Calado is from ISEL and iBB-Institute for Bioengineering and Biosciences, The Associate Laboratory Institute for Health and Bioeconomy (i4HB), Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal.

II. Materials and Methods

A. Population

Twenty patients admitted at the ICU of *Hospital São José Hospital*, in Lisbon, were considered. The present study is inserted in the PREMO project, approved by the Hospital Ethics Committee, and informed consent was obtained from each patient or their family members for data collection before participation. The patients' clinical information was anonymized.

All patients were men and were critically ill with Coronavirus 2019 disease (COVID-19), caused by the SARS-CoV-2 virus, as confirmed by Polymerase Chain Reaction (PCR). They were all under invasive mechanical ventilation. Half of the patients were diagnosed with bacteremia, as confirmed by microbiological culture.

B. Serum acquisition and analysis

Peripheral blood was collected in a tube without anticoagulant at the ICU between 7 and 9 a.m. and kept at −4ºC for 2 to 4 hours until centrifugation (3500 rpm for 10 minutes in a Mikro220T centrifuge, Hettich, Tuttlingen, Germany). Serum samples were stored at −80ºC until further analysis. The collection dates for cytokine samples were conducted in median on the first 5 days of ICU admission.

Serum cytokines were analyzed using a MILLIPLEX MAP 384-Well Human High Sensitivity T Cell Panel – Immunology Multiplex Assay, to profile a comprehensive set of 21 cytokines (HSTC384-28K, Millipore, Merck, Germany): ITAC, GM-CSF, Fractalkine, IFN-g, IL-10, MIP-3a, IL-12p70, IL-13, IL-17a, IL-1b, IL-2, IL-21, IL-4, IL-23, IL-5, IL-6, IL-7, IL-8, MIP-1a, MIP-1b, and TNF-a.

III. Data analysis

The statistical analysis concerning the patient's demographic and clinical characteristics, were based on the Student's t-test or the non-parametric Mann-Whitney U, if data was not normally distributed, or the Fishers exact test for categorical data. Normality was evaluated by the Shapiro-Wilk test. These tests were conducted with the IBM SPSS Statistics software, version 27 (IBMCorp., NewYork, USA).

The t-distributed stochastic neighbor embedding (t-SNE), as an unsupervised classification method, was conducted. Supervised classification Naïve Bayes models were also developed, with a 5-fold cross-validation. In this, the dataset is split into 5 equal parts, with the model trained on 4 parts and

tested on the remaining part in each iteration, ensuring every patient is tested once. It was represented the model's performance of the average of this 5-fold cross validation. The Information Gain algorithm was implanted for feature selection. These analyses were conducted with Orange: Data Mining Toolbox [7], version 3.36.2 (Bioinformatics Lab, University of Ljubljana, Ljubljana, Slovenia).

IV. Results and Discussion

The two groups of patients (*i.e.*, with and without bacteremia), were not statistically different concerning variables such as age, BMI, presence of comorbidities and if under Extracorporeal Membrane Oxygenation (ECMO) (Table 1). This increases the probability that the predicting model output depends on the bacteremia.

TABLE I. DEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF THE 20 PATIENTS, FROM WHICH HALF PRESENTED BACTEREMIA, WITH THE *P-VALUE* OF THE STATISTICAL ANALYSIS COMPARING THESE TWO GROUPS.

		Hemoculture negative (n = 10)	Hemoculture positive (n = 10)	*p-value*
Age (years), median (IQR)		60 (7)	58(22)	0.954*
ECMO (n/proportion)	No	8 (0.80)	7 (0.70)	1.000^{+}
	Yes	2 (0.20)	3 (0.30)	
With comorbidities (n/ proportion)	No	1 (0.10)	1 (0.10)	1.000^{+}
	Yes	9 (0.90)	9 (0.90)	
BMI, median (IQR)		26.27 (6.27)	27.17 (4.35)	0.744$^{\#}$
Nr. Days in ICU, median (IQR)		20(12)	18(14)	0.473$^{\#}$

*Students t-test, $^{+}$ Fishers exact test, $^{\#}$Mann-Whitney U

Fig. 1 represents the t-SNE plot of the 21 cytokines across the two groups of patients. The t-SNE algorithm converts similarities between data points into joint probabilities, minimizing the Kullback-Leibler (KL) divergence between these probabilities. The KL divergence measures the dissimilarity between two probability distributions. Interestingly, a clear pattern of data separation between the two groups emerged (Fig. 1), suggesting that the cytokine profiles can classify these patients.

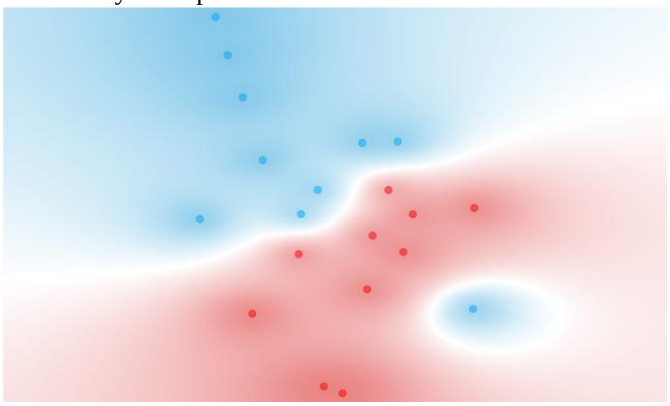

Fig. 1. t-SNE plot of patients with (blue) or without (red) bacteremia, based on 21 serum cytokines.

The Naïve Bayes supervised classification method was selected over other machine learning algorithms as it is a fast and efficient classifier, and typically performs well with modest-sized datasets, resulting in accurate classifications. This probabilistic model, based on the full set of serum cytokines, produced a reasonable output, with an Area Under the Receiver Operating Characteristic Curve (AUC-ROC) of 0.750 (Fig. 2), and sensitivity, specificity, precision, and accuracy all at 0.70.

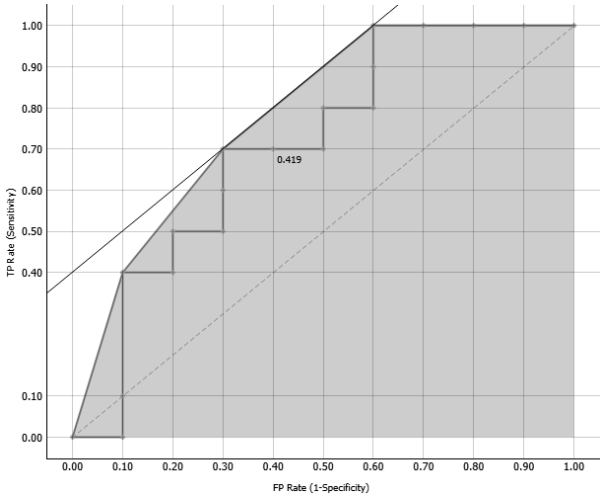

Fig. 2. AUC-ROC for the Naïve Bayes model, based on the full dataset of 21 serum cytokines, to predict bacteremia.

To improve the model's performance, the Information Gain algorithm was implemented for feature selection. This algorithm identifies each feature's contribution to reducing uncertainty (or entropy) in the dataset. The following three cytokines were highlighted: granulocyte-macrophage colony-stimulating factor (GM-CSF), interleukin 5 (IL-5), and interleukin 6 (IL-6). Interestingly, none of these three cytokines were statistically different between the two groups of patients (Fig. 3).

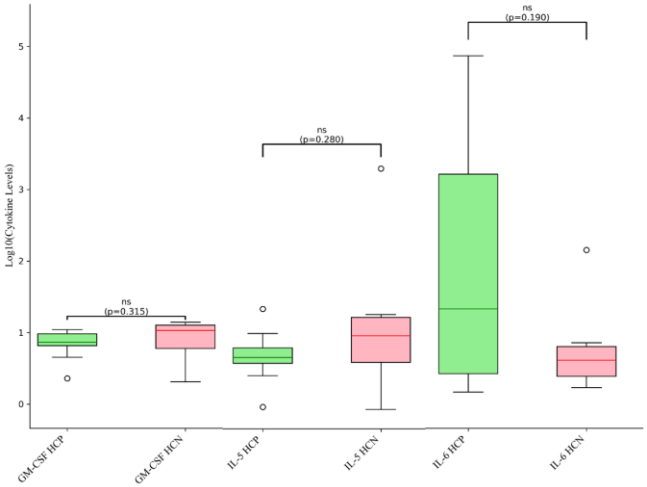

Fig. 3. GM-CSF, IL-5, and IL-6 levels in patients with bacteremia (hemoculture positive, HCP) and without infection (hemoculture negative, HCN). Cytokine levels were log-transformed for visualization. Statistical significance between groups was assessed using either the Student's t-test or Mann-Whitney U test, depending on normality. The corresponding p-values are displayed for each comparison.

The t-SNE based on these three cytokines presented a data cluster that separated almost all samples from the two patient groups (Fig. 4), performing nearly as well as the t-SNE based on the full set of 21 cytokines (Fig. 1). This suggests that this set of three cytokines, when considered together, can classify patients. To further explore this, a Naïve Bayes model was built using these three cytokines, resulting in an excellent AUC-ROC of 0.950 (Fig. 5), with sensitivity, specificity, precision, and accuracy of 0.900, 0.800, 0.818, and 0.850, respectively.

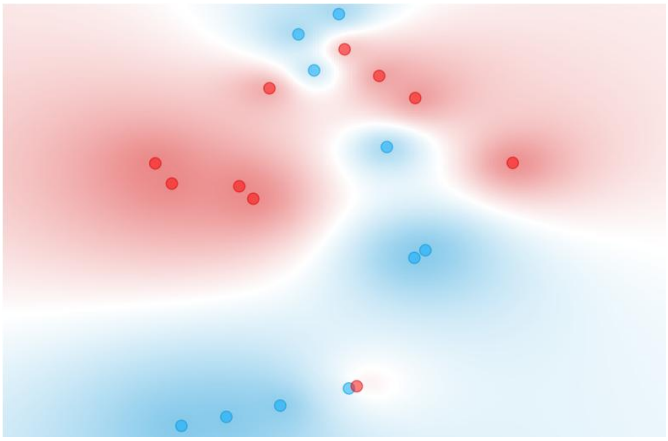

Fig. 4. t-SNE of patients with (blue) or without (red) bacteremia, based on the serum GM-CSF, IL-5, and IL-6 cytokines.

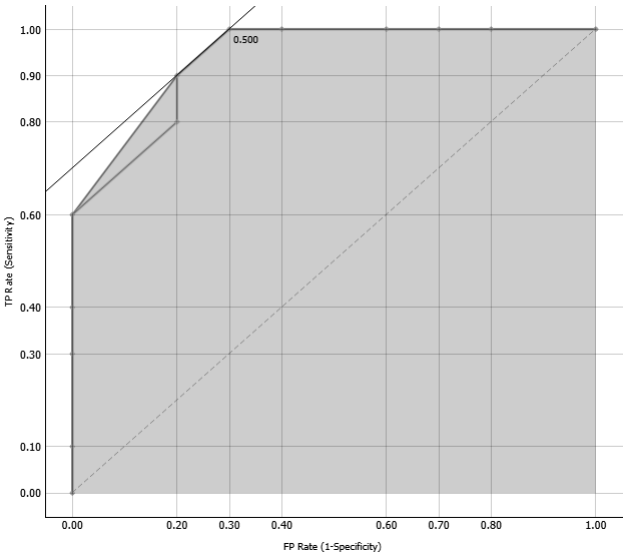

Fig. 5. AUC-ROC for the Naïve Bayes models based on the serum GM-CSF, IL-5, and IL-6, to predict bacteremia.

The identification of these three cytokines as predictors of bacteremia aligns with findings from other researchers. For example, GM-CSF plays a critical role in sepsis, as it affects various immune cells, including monocytes, macrophages, neutrophils, eosinophils, and basophils, by promoting increased cell survival, proliferation, differentiation, and activation (as reviewed in [8]). IL-5 is part of a group of cytokines, including IL-13, that are associated with the type 2 hyperinflammatory response in COVID-19 patients [9]. For example, IL-5 has been detected in critically ill COVID-19 patients at the time of intubation, with higher levels observed in the survivor group. [9]. Similarly, IL-6 is well known to be associated with cytokine storm syndrome and has been detected at higher levels in cases of bacteremia [10], as observed in the current patients.

In conclusion, excellent predictive models of bacteremia in critically ill patients under invasive mechanical ventilation were developed based on a small set of cytokines: GM-CSF, IL-5, and IL-6. Based on these outputs, it is recommended that these findings be further validated in a larger cohort.

Acknowledgment

This work was supported by the R-DICIP project (reference IPL/IDI&CA2024/R-DICIP_ISEL) granted by Instituto Politécnico de Lisboa (Portugal) and PREMO (reference DSAIPA/DS/0117/2020) granted by the Fundação para a Ciência e a Tecnologia (Portugal) (FCT). C. Rekowski, T. Fonseca, and R. Araújo, acknowledges the PhD grants from FCT, 2023.01951/BD, 2024.02043.BD and 2021.05553.BD, respectively.

References

[1] R. Araújo, L.F.N. Bento, T.A.H. Fonseca, C.P. Von Rekowski, B. Ribeiro da Cunha, C.R.C Calado, "Infection Biomarkers Based on Metabolomics", Metabolites, 12(2), pp. 92, 2022. https://doi.org/10.3390/metabo12020092

[2] J.R. Prowle, J.E. Echeverri, E.V. Ligabo, N. Sherry, G.C. Taori, T.M. Crozier, *et al.*, "Acquired bloodstream infection in the intensive care unit: incidence and attributable mortality", Crit Care, 15, p. R100, 2011. https://doi.org/10.1186/cc10114

[3] C. Adrie, C. Alberti, C. Chaix-Couturier, E. Azoulay, A. De Lassence, Y. Cohen, *et al.*, "Epidemiology and economic evaluation of severe sepsis in France: age, severity, infection site, and place of acquisition (community, hospital, or intensive care unit) as determinants of workload and cost", J Crit Care, 20, pp. 46-58, 2005. https://doi.org/10.1016/j.jcrc.2004.10.005

[4] C. Adrie, M. Garrouste-Orgeas, W. Ibn Essaied, C. Schwebel, M. Darmon, B. Mourvillier, et al., "Attributable mortality of ICU-acquired bloodstream infections: Impact of the source, causative micro-organism, resistance profile and antimicrobial therapy", J Infect, *74*(2), pp. 131–141, 2017. https://doi.org/10.1016/j.jinf.2016.11.001

[5] W. Schulte, J. Bernhagen, R. Bucala, "Cytokines in Sepsis: Potent Immunoregulators and Potential Therapeutic Targets—An Updated View", Mediators Inflamm., 1, pp. 16, 2013. https://doi.org/10.1155/2013/165974

[6] U. Trahtemberg, N. Bazak, S. Sviri, M. Beil, S. Paschke….PV. Heerden, "Cytokine patterns in critically ill patients undergoing percutaneous tracheostomy", Clin. Exp. Immunology, 198(1), pp. 121–129, 2019. https://doi.org/10.1111/cei.13333

[7] J. Demšar, A. Erjavec, T. Hoˇcevar, M. Milutinoviˇc, M. Možina, M. Toplak, L. Umek, J. Zbontar, B. Zupan, "Orange: Data Mining Toolbox in Python", J. Mach. Learn. Res. 14, pp. 2349-53, 2013.

[8] B. Mathias, B.E. Szpila, F.A. Moore, P.A. Efron, L.L. Moldawer, "A Review of GM-CSF Therapy in Sepsis", Medicine (Baltimore), 94(50):e2044. doi: 10.1097, 2015. https://doi.org/10.1097/MD.0000000000002044

[9] X. Ju, K. Son, R. Jamil, *et al.*, "Eosinophil-independent IL-5 levels are increased in critically ill COVID-19 patients who survive", Allergy Asthma Clin Immunol, 19, 58, 2023. https://doi.org/10.1186/s13223-023-00810-6

[19] L.Y.C. Chen, R.L. Hoiland, S. Stukas, C.L. Wellington, M.S. Sekhon, "Confronting the controversy: interleukin-6 and the COVID-19 cytokine storm syndrome", Eur Respir J., 1;56(4):2003006, 2020. https://doi.org/10.1183/13993003.03006-2020

[20] P.R. Reddy, M. Cherukuri, V.K. Eshwara, C.U. Kudru, R.K. Prabhu, "Diagnostic Potential of Serum Interleukin-6 in Predicting Bacteremia in Adult Patients with Sepsis: A Prospective Observational Study", Indian J Crit Care Med, 28(7):637-644, 2024. https://doi.org/10.5005/jp-journals-10071-24754

R Araújo holds several engineering degrees (Mechanical, Biomedical). A researcher at the Laboratory of Engineering and Health (ISEL) since 2017, focusing on biomarkers through metabolomics and proteomics for applications in toxicology, pharmacology, and more. His current PhD research, part of an FCT-approved project, studies COVID-19 biomarkers in critically ill patients using machine learning. https://orcid.org/0000-0002-9369-6486

L Ramalhete works in Biomedical Sciences and Health, focusing on Medical Biotechnology and Laboratory Technology, particularly in Histocompatibility and Immunogenetics. As the Technical and Scientific Supervisor of the CSTL-T Histocompatibility and Immunogenetics Lab, he oversees daily operations, ensures quality compliance, and provides clinical guidance for transplantation programs. His expertise includes flow cytometry, PCR-based assays, and immunological testing. He is also involved in research and development in immunobiology and oncogenomics and collaborates on national and international projects. https://orcid.org/0000-0002-8911-3380

T Fonseca holds a bachelor's degree in health sciences (2019) and a master's in biomedical engineering (2022) from ISEL. He is currently pursuing a Ph.D. in Biomedicine at Nova Medical School. Since 2020, he has been a research fellow at ISEL, working on the FCT-approved PREMO project. His research focuses on metabolomics and FTIR analysis, using blood serum samples from ICU patients. His primary goal is to identify biomarkers that can predict patient outcomes in critical care units, hence improving predictive models for morbidity and mortality. https://orcid.org/0000-0003-0741-2211

C Von Rekowski holds a degree in Clinical Physiology (2019) from Escola Superior de Tecnologia da Saúde de Lisboa and a master's in biomedical engineering (2022) from ISEL. She has been a research fellow at ISEL since 2020, working on an FCT-approved project focusing on biomarker identification, using metabolomics, proteomics, and machine learning. She is currently pursuing a Ph.D. in Biomedicine at Nova Medical School, while also supporting the work of various master's students. Her work focuses on building databases, performing statistical analysis of demographic, clinical, and laboratory data, and using this data to develop predictive models for morbidity and mortality in critically ill patients. https://orcid.org/0009-0009-6843-1935

L Bento completed his PhD in Clinical Research in 2018, at *Universidade Nova de Lisboa | Faculdade de Ciências Médicas*, and his Medical degree in 1994, at the same institution. He is the Coordinator of the Medical Emergency Unit at *Centro Hospitalar Universitário Lisboa Central EPE*, Head of the Intensive Care Medicine Department, and a Consultant in Internal Medicine. He also serves as an Assistant Professor at *Universidade Nova de Lisboa* and is a specialist in Internal Medicine and Intensive Care Medicine. He has worked in areas such as acute-on-chronic liver failure, acute kidney injury, multi-organ failure, hepatic transplant, and hemodynamic instability, with a focus on intensive care and emergency medicine. https://orcid.org/0000-0002-0260-003X

CRC Calado is Coordinating professor with Aggregation at ISEL. She has a BSc in Biochemistry, an MSc and a PhD in Biotechnology and an Aggregation in Biochemistry. She is the coordinator of the BSC in Biomedical Engineering and the R&D Laboratory in Health & Engineering. Her research focuses on Bioprocess Engineering and Development of Platforms for Drugs and Biomarkers Discovery.
https://orcid.org/0000-0002-5264-9755