



NAVIGATING MYSTERIOUS LANDS REGARDING THE PREDICTIVE VALUE OF THE SOFA SCORE: THE RACE DILEMMA

NAVEGANDO POR UN TERRITORIO MISTERIOSO EN CUANTO AL VALOR PREDICTIVO DE LA PUNTUACIÓN DEL SOFA: EL DILEMA DE LA RAZA

Ivan David Lozada Martínez^{1,2a}, Stehany Hernandez Muñoz^{3b}, Ailyn Zenith Angulo Lara^{4b}, Estefanía Ruiz Benítez^{5b}, Cindy Stephanie Buitrago Almanza^{6b}, Jhon Alejandro Hernández Gallego^{7b}, Laura Cristina Carreño Mancilla^{8b}, Nathalia Carolina Ordóñez Veloza^{9b}, Yohan Santiago Reyes Triviño^{10b}

ABSTRACT

Introduction: Critical syndromes are conditions that carry a high global burden of disease. Scoring systems are practical and reproducible aids that allow patients with more severe disease to be quickly identified and admitted to intensive care and to initiate structured and aggressive therapy. The Sequential Organ Failure Assessment (SOFA) score is one of the most widely used in the world, as there are several versions and it is simple. However, with the appearance of the COVID-19, several studies showed that there was a disparity in the estimation of mortality and associated outcomes, with respect to race, which culminated in an excess of preventable mortality in certain racial groups. Constant evaluation of the performance of these scoring systems must be performed due to definitional updates, which may vary the accuracy of the predictive value. There is a very large evidence gap in this regard, as the existing studies are from high-income countries where the predominant racial group is Caucasian, which should draw attention to the magnitude of the problem. Based on the above, the objective of this review is to discuss evidence on the performance of critical care scoring systems, particularly SOFA, and the impact that race has had on its predictive value.

Keywords: Organ Dysfunction Scores, Continental Population Groups, Predictive Value of Tests, Critical Care Outcomes. (Source: MESH-NLM)

RESUMEN

Introducción: Los síndromes críticos son condiciones que acarrear una elevada carga de enfermedad a nivel global. Los sistemas de puntaje, son ayudas prácticas y reproducibles que permiten identificar de manera rápida pacientes con enfermedad más grave e ingresarlos a cuidado intensivo e iniciar terapia estructurada y agresiva. El score Sequential Organ Failure Assessment (SOFA), es uno de los más utilizados en el mundo, al existir varias versiones y ser sencillo. No obstante, con la aparición de la COVID-19, diversos estudios demostraron que existía una disparidad en cuanto a la estimación de mortalidad y desenlaces asociados, respecto a la raza, lo que culminó en un exceso de mortalidad prevenible en ciertos grupos raciales. La evaluación constante del rendimiento de estos sistemas de puntaje, debe realizarse debido a actualizaciones de definiciones, las cuales pueden variar la precisión del valor predictivo. Existe una brecha muy grande en cuanto la evidencia al respecto, puesto que los estudios existentes provienen de países de altos ingresos, donde el grupo racial predominante son los caucásicos, lo que debe llamar la atención de la magnitud del problema. Sobre la base de lo anterior, el objetivo de esta revisión consiste en discutir evidencia al respecto sobre el rendimiento de sistemas de puntuación en cuidado crítico, particularmente del SOFA y el impacto que ha tenido la raza sobre su valor predictivo.

Palabras claves: Puntuaciones en la Disfunción de Órganos, Grupos de Población Continentales, Valor Predictivo de las Pruebas, Resultados de Cuidados Críticos. (Fuente: DeCS- BIREME)

¹ Medical-Surgical Research Center, Future Surgeons Chapter, Colombian Association of Surgery, Bogotá, Colombia.

² Prometheus Group and Biomedicine Applied to Clinical Sciences, Faculty of Medicine, University of Cartagena, Cartagena, Colombia.

³ Faculty of Medicine, Universidad Santiago de Cali, Cali, Colombia.

⁴ Faculty of Medicine, Universidad Libre, Barranquilla, Colombia.

⁵ Faculty of Medicine, Universidad del Rosario, Bogotá, Colombia.

⁶ Faculty of Medicine, National University, Colombia, Colombia.

⁷ Faculty of Medicine, Cooperative University of Colombia, Medellín, Colombia.

⁸ Faculty of Medicine, Autonomous University of Bucaramanga, Bucaramanga, Colombia.

⁹ Faculty of Medicine, University of Manizales, Manizales, Colombia.

¹⁰ Faculty of Medicine, Juan N. Corpas University Foundation, Bogotá, Colombia.

^a Medical Intern.

^b Physician.

Cite as: Lozada Martínez ID, Hernandez Muñoz S, Angulo Lara AZ, Ruiz Benítez E, Buitrago Almanza CS, Hernández Gallego JA, Carreño Mancilla LC, Ordóñez Veloza NC, Reyes Triviño YS. Navigating mysterious lands regarding the predictive value of the SOFA score: the race dilemma. Rev Fac Med Hum. 2022;22(4):833-840. doi 10.25176/RFMH.v22i4.5036

Journal home page: <http://revistas.urp.edu.pe/index.php/RFMH>

Article published by the Magazine of the Faculty of Human Medicine of the Ricardo Palma University. It is an open access article, distributed under the terms of the Creative Commons License: Creative Commons Attribution 4.0 International, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>), that allows non-commercial use, distribution and reproduction in any medium, provided that the original work is duly cited. For commercial use, please contact revista.medicina@urp.pe





INTRODUCTION

With the appearance of the COVID-19 pandemic, the need to validate and carry out massive research in critical care in regions that lack quality primary data became evident⁽¹⁾. Disease burden studies show that for the year 2004, there were 58,772 deaths globally, with critical syndromes such as sepsis, acute lung injury, and invasive mechanical ventilation being the most frequent and with the greatest burden of disease worldwide. (approximately 45,000 / 58,772 reported cases)⁽²⁾.

The availability of an intensive care unit (ICU) bed per 100,000 inhabitants continues to be a global health challenge. While high-income countries such as the United States, Germany, Canada, among others; have an average of between 15 and 20 beds per 100,000 inhabitants; low- and middle-income countries have on average ≤ 5 beds for the same proportion of the population⁽²⁾. At that time almost 20 years ago, promoting strategies that would improve disease burden indicators were mentioned as a global health objective, especially considering the political, economic, and health limitations of a large part of the world's population⁽³⁻⁶⁾.

Recently, a meta-analysis carried out by the Department of Integrated Health Services of the World Health Organization showed that the global incidence of sepsis with multi-organ dysfunction continues to be in ranges similar to those reported many years ago (9,300 cases per 100,000 inhabitants), and maintains a worrying incidence of the hospital- and ICU-acquired sepsis, which ranges between 25% - 50%⁽⁷⁾. Reported mortality currently exceeds 50%; it is necessary to highlight that this meta-analysis included 51 studies, and approximately half were from low- and middle-income countries, which is of great concern due to the previously described barriers⁽⁷⁾. Preventable deaths in critical care is a hot topic of discussion since different tools have been proposed to help prioritize or re-stratify in timely manner patients with higher support requirements for their survival^(6,8,9). The scales consisting of scoring systems are practical and reproducible aids that allow the rapid identification of patients with more severe disease, admission to the ICU, and initiation of

structured and aggressive therapy⁽¹⁰⁾. Many of these scales have been proposed and validated over the years, such as SOFA (Sequential Organ Failure Assessment)⁽¹¹⁾, APACHE (Acute Physiology and Chronic Health Evaluation)⁽¹²⁾, SAPS (Simplified Acute Physiologic Score)⁽¹²⁾, MPM0 (Admission Mortality Probability Model)⁽¹³⁾, among others^(8,9).

Various studies have evaluated the performance of these scales in different contexts, adapting them to new variables according to the population's behavior where they are used, with new versions emerging since their creation⁽¹⁰⁾. This is fundamental since, being a scale that defines the critically ill patient quickly, the prognostic value must be precise in those who apply it, regardless of the population group evaluated. Meta-investigations strongly discuss the heterogeneity of some studies and the implications that this would have in real practice⁽¹⁴⁾.

During the course of the COVID-19 pandemic, there were many differences in the performance of these scales, which unbalanced the validity of the evidence⁽¹⁵⁻¹⁷⁾. Particularly, during the year 2021, many studies suggested that there were disparities in the accuracy of the SOFA (one of the most used scales and with the best predictive value) in terms of the race of the population where it was used, having a substantial impact on the procedure doctor in the ICU and hospitalization, in patients with severe COVID-19 phenotype, who required invasive mechanical ventilation, scarce drugs, strict surveillance, among other interventions⁽¹⁵⁻¹⁷⁾; which were very limited back then.

Considering the current relevance of the global burden of disease caused by critical clinical syndromes, the need to apply meta-investigations that discuss the strengths and weaknesses of the evidence and to know possible biases that exist in prognostic tools in the management of critically ill patients, the objective of this review is to summarize evidence on the performance of scales used in critical care, and particularly the SOFA, which allows knowing the outcomes obtained over time, and to discuss the dilemma of the race when using this instrument.



METHODS

A bibliographic search was carried out using search terms such as "SOFA", "Score", "Critical Care" and "Race", as well as synonyms, which were combined with the Boolean operators "AND" and "OR", in the search engines. search and databases PubMed, ScienceDirect, Embase, EBSCO, and MEDLINE. As an inclusion criterion, it was defined that any article related to the evaluation of the SOFA scale in patients where performance was discriminated according to race or other relevant subgroups would be included, giving priority to original studies and systematic reviews, and meta-analyses. Also, articles related to other critical care scores with predictive values for mortality and associated outcomes were included.

In addition, they should be available in full text. As non-inclusion criteria, it was established that articles published in a language other than Spanish and English would not be included. Considering the breadth of the topic and the great variety of publications, articles published between 2000 and 2022 were included. A total of 242 potentially relevant articles were identified, with a review of the title and abstract of all of them, of which finally 57 articles were included after their discrimination according to the inclusion and non-inclusion criteria. Other useful references were included for the discussion of general concepts. The estimates and calculations found were expressed in their original measurements, whether frequencies, percentages, confidence intervals (CI), mean difference (MD), relative risk (RR), odds ratio (OR), and incidence rate (IRR). or hazard ratio (HR).

RESULTS

Predictive scoring systems for outcomes in severe disease.

APACHE

Starting in the 1980s, there was already talk about the APACHE score and its first versions(18). This system is made up of variables such as age, temperature, mean arterial pressure, heart and respiratory rate, creatinine, and Glasgow coma scale, among others; which can

range between 0 and 59, establishing a minimum cut-off score of 0-4 and a maximum of >34. The mortality of those in the minimum range fluctuates between 1% - 4%, for both postoperative and non-operative patients⁽¹⁹⁾. Those that reach the maximum cut, have mortality above 85%, an important value in decision-making. These data take into account the prognostic value of version number two of the score (APACHE II)⁽¹⁹⁾.

In 1991, Knaus et al.⁽¹⁸⁾ conducted a multicenter prospective study with more than 17,000 patients in the United States to increase the prognostic accuracy of the APACHE (resulting in a new version). The authors showed that a 5-point increase in the new score (APACHE III) was independently associated with an increased mortality risk (OR 1.10) in each of the 78 categories. Similarly, it was observed that the equation monopolized the variability in the presented mortality rate ($r^2 = 0.90$; $p < 0.0001$)(18). In the 2000s, numerous studies were carried out that tried to reproduce and validate the performance of these criteria, but including more and more variables, according to the subgroups that were presented, taking into account the cause of the critical clinical syndrome⁽¹⁹⁻²¹⁾.

Ho et al.⁽¹⁹⁾ carried out a retrospective cohort study on that date, where they evaluated more than 11,000 admissions to the ICU after non-cardiovascular surgery, performing a reassessment of the patient's status at 24 hours, showing that the average score obtained was 12 and 15, respectively; and that the predicted mortality was 15% and 19%, correlating with the real mortality, which was 16%. The area under the curve (AUC) calculated was 83.8% and 84.6% for both cuts, with no significant difference ($p=1.0$)⁽¹⁹⁾. This allowed us to conclude that the score maintained adequate performance from its first calculation, regardless of the worsening of the patient 24 hours after admission.



In 2006, two very large cohort studies were published, both by Zimmerman et al.^(20,21), who included more than 130,000 ICU admissions from 45 hospital centers in the United States, giving rise to the fourth version of the score (APACHE IV). This included new variables, mainly regarding chronic health and ICU admission diagnosis. The authors evidenced adequate validation (AUC: 0.88), without finding significant differences in terms of mortality variation in 90% of cases⁽²⁰⁾. Regarding ICU stay, the same authors in the second study found that the predicted value correlated adequately with the real value (in this case, 3.78 vs. 3.86; $p < 0.001$), with no significant differences among 93% of the diagnoses made⁽²¹⁾.

It was concluded that the APACHE IV was a tool that provided clinical utility to critically ill patients regarding their ICU stay and mortality risk^(20,21). However, it was highlighted that due to a large number of categories, the score was dynamic and should be interpreted rationally, since it was individualized. Recent studies continue to evaluate the performance of this last score, emphasizing some factors that can significantly influence the performance of the predictive value, according to the subgroup evaluated.

For example, Xu et al.⁽²²⁾ evaluated the score in post-transplant patients, finding that age, the use of hormones, and the presence of respiratory failure must be considered since they modify the association between the score and the actual mortality⁽²²⁾. Xiao et al.⁽²³⁾, performed a secondary analysis of the eICU collaborative database, showing that baseline platelet count is negatively associated with all causes of mortality, both in the hospital (RR: 0.87; 95% CI, 0.84-0.91) and in the ICU (RR: 0.87; 95% CI, 0.83-0.92)⁽²³⁾. Other authors have shown that the clinical course and establishment of treatment influence the predictive value of scores in critical care (5% in predicting mortality and up to 4 hours of hospital stay)⁽²⁴⁾. Procalcitonin has been described as a biomarker that also influences, although not significantly, the prediction of survival at different times, and it is necessary to consider it in depth in those with organ

involvement where this peptide is produced and released⁽²⁵⁾. Today, the APACHE is still used in critical care and has shown to have adequate performance in general, although it has the limitation that it is individualized and that variables continue to arise that can affect the predictive value since it has many categories to evaluate.

SAPS II

In 1993⁽²⁶⁾, the study that developed and validated the SAPS II score was published, consisting of 17 variables (12 physiological, age, type of admission, and presence of 3 underlying diseases [cancer and immunodeficiency syndrome acquired]). This study included more than 13,000 patients and obtained an AUC of 0.88 for its development and 0.86 for its validation⁽²⁶⁾. Like the APACHE, it has been replicated throughout the world, and to date, studies continue to appear that report factors that modify the performance of this score.

Recently, it was shown that colonization by *Clostridioides difficile* increases the SAPS II score, although it does not modify mortality or the frequency and/or severity of diarrhea⁽²⁷⁾. Another study found that in patients with trauma, homocysteine, D-dimer, and procalcitonin are parameters that are associated with greater disease severity and, therefore, mortality risk in those patients stratified by SAPS II ($p < 0.05$), considering independent factors of unfavorable prognosis⁽²⁸⁾. Likewise, it has been observed that extracorporeal membrane oxygenation (ECMO) is a complementary risk factor that increases the risk of death in those patients with a moderate or high SAPS II score⁽²⁹⁾. Also, SAPS II has been found to be a good predictor of outcomes in difficult-to-access tracheostomized patients admitted to specialized rehabilitation units⁽³⁰⁾.

In the third version of this score (SAPS III), which was used in a timely manner in the management of critically ill patients with COVID-19, it showed that in diabetic and non-diabetic Austrian patients, the performance was inadequate due to inaccuracy in predicting mortality⁽³¹⁾.





Although some authors tried to calibrate it, it was not possible to achieve it. Therefore, it is suggested to use it with caution or not in this population⁽³¹⁾. This could be the least used score and the one most influenced by different factors, despite being made up of a few categories.

MPMO

This score is made up of 15 variables, which evaluate the patient's evaluation time according to admission, morbid and paraclinical history, clinical status, and treatment performed. Like the other previously described scores, this one was created in the 1980s and initially validated in a multicenter manner in the 1990s⁽³¹⁾. To date, the third version of this score (MPMO-III) is used. Studies such as the one by Higgins et al.⁽³²⁾ have prospectively evaluated the performance of this version in more than 55,000 patients, finding that the predicted mortality was 7,456, which correlated satisfactorily with the actual mortality (7,331; 13.2. %), obtaining a mortality ratio of 0.983 (95% CI, 0.963 – 1.001)⁽³²⁾. Until that moment, it was concluded that the model was robust and had optimal external validity in the American population, which was where the primary data were obtained.

Less than 10 years ago, new parameters appeared that influenced the performance of this score, as well as others, which must be taken into account when establishing the prognosis in critical care. Ho et al.⁽³³⁾ conducted a cohort study in Australia, where they evaluated the impact of the anion gap and other similar paraclinical on the mortality of 6,878 individuals⁽³³⁾.

The authors showed that 13.4% died (n=924), highlighting the differences between the acid-base markers between the groups of survivors vs. not survivors. The anion gap added to the lactate presented an AUC of 0.631; while in isolation, it presented an AUC of 0.521. When adjusting the mortality predicted by MPMO-III according to the markers, it was shown that arterial lactate correlated with the mortality variability, compared to the anion gap, which lost its predictive value⁽³³⁾. More precisely, in low- and middle-income countries, there have been reports criticizing the accuracy of this score, making it necessary to adapt it

under certain limitations, as is the case of the R-MPM score in Rwanda⁽³⁴⁾.

However, studies that have compared the APACHE, SAPS, and MPMO scores have found that although all three have good calibration, the APACHE IV is superior in terms of its ability to predict mortality (AUC 0.745); while SAPS and MPMO present an AUC of 0.700 and 0.670, respectively⁽³⁵⁾. Chen et al.⁽³⁶⁾ compared eight different scores to evaluate the prediction of general mortality at 28 days, finding that APACHE III was superior to the rest in both outcomes, with an AUC of 0.817⁽³⁶⁾. Another study compared APACHE II, SAPS III and MPMO in 9549 patients, of whom 1276 died⁽³⁷⁾. It was found that although APACHE II had better AUC (0.845) than SAPS III (0.836) and MPMO-III (0.807), SAPS III had better calibration and overall performance (slope of curve 1.03, R² 0.297) compared to the other two⁽³⁷⁾. Up to this point, it is understood that of the three scores studied, APACHE has better performance, regardless of subgroups, having some variations in the calibration according to certain parameters but maintaining an adequate AUC. It is observed that MPMO is not so reliable and is dependent on many variable adjustments.

SOFA

This score was designed after an event in 1994, made up of clinical and paraclinical criteria, which allow evaluation of the presence of multi-organ damage. It has been closely associated with infections and sepsis, the diagnoses from which the most robust studies have been made to-date⁽³⁸⁻⁴³⁾. Retrospective analyzes of more than 180,000 patients have found discrimination of hospital mortality given by an AUC 0.753 (99% CI, 0.750 - 0.757); being above summary versions and other criteria of the systemic inflammatory response⁽³⁸⁾. Favorably, when performing multiple sensitivity analyses for the outcome of in-hospital mortality, the estimated prognosis was sustained, which may reflect adequate external validity. Currently, the breadth of applicability of this score has been discussed, supported by internationally recognized critical care





care scientific societies. It has been postulated as a reference marker of efficacy in clinical trials since the variations between the different studies and compared to other parameters are minimal⁽³⁹⁾. Liu et al.⁽⁴⁰⁾ evaluated 1865 patients in China who had sepsis and were admitted to the ICU, in whom serum lactate, SOFA and quick SOFA (qSOFA) were measured to compare the predictive value regarding mortality. It was found that SOFA presented higher AUC (0.686; 95% CI, 0.661-0.710), compared to the other two parameters (serum lactate, AUC 0.664; 95% CI, 0.639-0.689 vs. qSOFA, AUC 0.547; 95% CI). %, 0.521 – 0.574)⁽⁴⁰⁾. In acute decompensated heart failure, SOFA is significantly associated with overall and 30-day mortality, with an AUC of 0.765 (95% CI, 0.733-0.798) and 0.706 (95% CI, 0.676-0.736), respectively⁽⁴¹⁾. In patients with suspected infection who are evaluated in the emergency department, a SOFA score ≥ 2 is independently associated with mortality up to 2 years later (HR 1.90; 1.83 - 1.98)⁽⁴²⁾. This precision has been obtained in other studies that have also studied representative samples in various parts of the world⁽⁴³⁾.

These results show that SOFA is probably the most accurate and widely used score globally today. Although it requires paraclinical for its calculation, it would be difficult to establish the prediction in a timely manner in low- and middle-income countries, where there are not many specialized centers and the availability of individual teams by departments or treatment units.

SOFA performance over time and the dilemma of race

During the COVID-19 pandemic, due to the variety of unknown information on the management of these patients, various analyzes allowed us to observe discrepancies and correlations not previously studied in depth, which have the potential to modify the performance of scores in critical care, such as the SOFA. Over time, this scoring system has been studied in depth, being sustained since the beginning of the year 2000, as a simple but effective score in predicting mortality⁽⁴⁴⁾. As new definitions appear, it is necessary to revalidate the the performance of the scores. In 2017,

with the appearance of the third definition of sepsis, Matics & Sanchez-Pinto⁽⁴⁵⁾ published the results of the adaptation and validation of this system in the pediatric population through the evaluation of 8,711 events, obtaining a mortality of 2.6% and an AUC of 0.94 (95% CI, 0.92-0.95)⁽⁴⁵⁾. Other studies, such as those by Pawar et al.⁽⁴⁶⁾ and Arakawa et al.⁽⁴⁷⁾ discussed the variability of SOFA against different infectious states and disseminated intravascular coagulation (DIC), respectively. The former showed that the AUC ranged between 0.59 (95% CI, 0.49 - 0.70) and 0.79 (95% CI, 0.69 - 0.90), for cases with endocarditis and bacteremia. isolated⁽⁴⁶⁾.

The latter developed the SOFACOMB, finding that in the DIC, the AUC at 2, 4, and 7 days was much higher compared to the original version ($p < 0.002$)⁽⁴⁷⁾. In COVID-19 patients, SOFA was used to predict the development of a severe phenotype (AUC 0.908; 95% CI, 0.857-0.960) and overall and 60-day mortality (AUC 0.995; 95% CI, 0.985-1000)⁽⁴⁸⁾. In octogenarians, accuracy varies depending on variables such as the day of score evaluation, neurological failure, and polypharmacy on admission, which can generate an AUC between 0.71 and 0.91⁽⁴⁹⁾. In this order of ideas, particularly as new definitions appear and new parameters are incorporated into their use, it is necessary to reassess the variability of the scores according to the stratification of patients by subgroups^(50,51).

A very interesting situation that was evidenced during the management of COVID-19 patients was the impact that race had on the general and specific outcomes and the predictive value of SOFA⁽⁵²⁻⁵⁴⁾. In early 2021, Rodriguez et al.⁽⁵²⁾ analyzed the American Heart Association cardiovascular disease registry of those hospitalized for COVID-19, including 7,868 patients. The study group was heterogeneous, composed of Hispanics (33%), non-Hispanic Afros (25.5%), Asians (6.3%), and non-Hispanic Caucasians (35.2%). It was observed that the general mortality was 18%, and 53% was concentrated in Hispanic and Afro patients. When adjusting for morbidity and mortality, by race, this has the greatest burden, although Asians more frequently presented cardiopulmonary severity due to COVID-19⁽⁵²⁾.



Gershengorn et al. ⁽⁵³⁾ conducted a bicenter retrospective cohort study with 1127 in the United States for the same date, who also had some ethnic heterogeneity (63.1% Caucasian, 28.7% Afro-descendant; 54.2% of the total were Hispanic patients). The authors found no association between ethnicity and variability in the prediction of mortality when taking the Caucasian group as a reference (Afro-descendants [IRR 1.00; 95% CI, 0.89 - 1.12]; Asians [IRR 0, 95, 95% CI, 0.62 - 1.45], multiracial [IRR 0.93, 95% CI, 0.72 - 1.19] ⁽⁵³⁾ Ashana et al. ⁽⁵⁴⁾ also evaluated this phenomenon through from the analysis of more than 113,000 patients, 24.4% of whom were Afro-descendants.

When comparing the SOFA with the Laboratory-based Acute Physiology Score (LAPS2), it was shown that the latter discriminates more accurately the outcomes in Afro-descendant patients (AUC 0.76; 95% CI, 0.76-0.77 vs. AUC 0.68; 95% CI, 0.68-0.69). The LAPS2 was better calibrated in predicting mortality in both racial groups. the creatinine item of the SOFA, the miscalibration was reduced ⁽⁵⁴⁾. The authors concluded that it is necessary to develop more equitable scores, and clearly, it can be evidenced that when evaluating more robust samples, it is o more easily observing the existing gap. It can be inferred then that in countries where this race predominates, there should be greater disparities with this score.

Authorship contributions: All authors participated in the research through the development of the project, collection, and analysis of information, as well as in the preparation of the manuscript of this research.

Funding sources: Self financed.

Correspondence: : Ivan David Lozada Martinez.

Address: Grupo Prometeus y Biomedicina Aplicada a las Ciencias Clinicas, Facultad de Medicina, Universidad de Cartagena, Cartagena, Colombia.

E-mail: ilozadam@unicartagena.edu.co

REFERENCES

1. Immovilli P, Morelli N, Antonucci E, Radaelli G, Barbera M, Guidetti D. COVID-19 mortality and ICU admission: the Italian experience. *Crit Care*. 2020; 24(1):228. DOI: 10.1186/s13054-020-02957-9
2. Adhikari NK, Fowler RA, Bhagwanjee S, Rubenfeld GD. Critical care and the global burden of critical illness in adults. *Lancet*. 2010; 376(9749):1339-46. DOI: 10.1016/S0140-6736(10)60446-1
3. Baker T, Schultz MJ, Dünser MW; Global Intensive Care Working Group of the European Society of Intensive Care Medicine. Critical illness in developing countries: dying in the dark. *Lancet*. 2011; 377(9775):1405. DOI: 10.1016/S0140-6736(11)60572-2.
4. Fowler RA, Adhikari NK, Bhagwanjee S. Clinical review: critical care in the global context--disparities in burden of illness, access, and economics. *Crit Care*. 2008; 12(5):225. DOI: 10.1186/cc6984
5. Vincent JL, Marshall JC, Namendys-Silva SA, François B, Martin-Loeches I, Lipman J, et al. Assessment of the worldwide burden of critical illness: the intensive care over nations (ICON) audit. *Lancet Respir Med*. 2014; 2(5):380-6. DOI: 10.1016/S2213-2600(14)70061-X
6. Dijkema LM, Dieperink W, van Meurs M, Zijlstra JG. Preventable mortality evaluation in the ICU. *Crit Care*. 2012; 16(2):309. DOI: 10.1186/cc11212
7. Markwart R, Saito H, Harder T, Tomczyk S, Cassini A, Fleischmann-Struzek C, et al. Epidemiology and burden of sepsis acquired in hospitals and intensive care units: a systematic review and meta-analysis. *Intensive Care Med*. 2020; 46(8):1536-1551. DOI: 10.1007/s00134-020-06106-2
8. Serviá L, Llupart-Pou JA, Chico-Fernández M, Montserrat N, Badia M, Barea-Mendoza JA, et al. Development of a new score for early mortality prediction in trauma ICU patients: RETRASCORE. *Crit Care*. 2021; 25(1):420. DOI: 10.1186/s13054-021-03845-6
9. Deliberato RO, Escudero GG, Bulgarelli L, Neto AS, Ko SQ, Campos NS, et al. EVERITAS: An externally validated mortality prediction for critically ill patients in low and middle-income countries. *Int J Med Inform*. 2019; 131:103959. DOI: 10.1016/j.ijmedinf.2019.103959

Other authors have recently shown that disparities continue to exist regarding the priority of granting ventilators and prioritizing care through standard care in crisis events, which culminated in excess mortality (up to 43.9% in Afro-descendants) preventable ⁽⁵⁵⁻⁵⁷⁾. This, added to the underestimation that can occur in certain clinical contexts, can increase morbidity, mortality or disability among racial groups, mainly Afro-descendants. Therefore, it is necessary to reevaluate or design scores that have greater precision, considering the considerations discussed regarding race and the different clinical-pathological contexts.

CONCLUSION

Currently, numerous scoring systems have a useful predictive value for mortality and associated outcomes in critical care, which should be used rationally depending on the various clinical contexts. The SOFA score has stood out for having a satisfactory performance over the years, even being adapted and validated in pediatrics and in its fast version. No obstante, parece verse afectado por la raza, subestimando la mortalidad principalmente en pacientes afrodescendientes, lo que puede culminar en un exceso de mortalidad prevenible.

Conflicts of interest: None of the authors has a conflict of interest in accordance with their declaration

Received: June 28, 2022

Approved: August 16, 2022

10. Magee F, Wilson A, Bailey M, Pilcher D, Gabbe B, Bellomo R. Comparison of Intensive Care and Trauma-specific Scoring Systems in Critically Ill Patients. *Injury*. 2021; 52(9):2543-2550. DOI: 10.1016/j.injury.2021.03.049
11. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA*. 2001; 286(14):1754-8. DOI: 10.1001/jama.286.14.1754
12. Capuzzo M, Valpondi V, Sgarbi A, Bortolazzi S, Pavoni V, Gilli G, et al. Validation of severity scoring systems SAPS II and APACHE II in a single-center population. *Intensive Care Med*. 2000; 26(12):1779-85. DOI: 10.1007/s001340000715
13. Higgins TL, Teres D, Copes WS, Nathanson BH, Stark M, Kramer AA. Assessing contemporary intensive care unit outcome: an updated Mortality Probability Admission Model (MPM0-III). *Crit Care Med*. 2007; 35(3):827-35. DOI: 10.1097/01.CCM.0000257337.63529.9F
14. Fuchs PA, Czech JJ, Krzych LJ. Mortality Prediction Using SOFA Score in Critically Ill Surgical and Non-Surgical Patients: Which Parameter Is the Most Valuable? *Medicina (Kaunas)*. 2020; 56(6):273. DOI: 10.3390/medicina56060273
15. Miller WD, Han X, Peek ME, Charan Ashana D, Parker WF. Accuracy of the Sequential Organ Failure Assessment Score for In-Hospital Mortality by Race and Relevance to Crisis Standards of Care. *JAMA Netw Open*. 2021; 4(6):e2113891. DOI: 10.1001/jamanetworkopen.2021.13891
16. Tolchin B, Oladele C, Galusha D, Kashyap N, Showstark M, Bonito J, et al. Racial disparities in the SOFA score among patients hospitalized with COVID-19. *PLoS One*. 2021; 16(9):e0257608. DOI: 10.1371/journal.pone.0257608
17. Roy S, Showstark M, Tolchin B, Kashyap N, Bonito J, Salazar MC, et al. The potential impact of triage protocols on racial disparities in clinical outcomes among COVID-positive patients in a large academic healthcare system. *PLoS One*. 2021; 16(9):e0256763. DOI: 10.1371/journal.pone.0256763
18. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest*. 1991; 100:1619. DOI: 10.1378/chest.100.6.1619
19. Ho KM, Dobb GJ, Knuiam M, Finn J, Lee KY, Webb SA. A comparison of admission and worst 24-hour Acute Physiology and Chronic Health Evaluation II scores in predicting hospital mortality: a retrospective cohort study. *Crit Care* 2006; 10:R4. DOI: 10.1186/cc3913
20. Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med*. 2006; 34(5):1297-310. DOI: 10.1097/01.CCM.0000215112.84523.F0
21. Zimmerman JE, Kramer AA, McNair DS, Malila FM, Shaffer VL. Intensive care unit length of stay: Benchmarking based on Acute Physiology and Chronic Health Evaluation (APACHE) IV. *Crit Care Med*. 2006; 34(10):2517-29. DOI: 10.1097/01.CCM.0000240233.01711.D9
22. Xu Y, Chao S, Niu Y. Association between the Predicted Value of APACHE IV Scores and Intensive Care Unit Mortality: A Secondary Analysis Based on EICU Dataset. *Comput Math Methods Med*. 2022; 2022:9151925. DOI: 10.1155/2022/9151925
23. Xiao C, Qin Z, Xiao J, Li Q, He T, Li S, et al. Association between basal platelet count and all-cause mortality in critically ill patients with acute respiratory failure: a secondary analysis from the eICU collaborative research database. *Am J Transl Res*. 2022; 14(3):1685-1694. DOI: 10.1186/s13054-022-03986-2
24. Candel BGJ, Raven W, Lameijer H, Thijssen WAMH, Temorshuizen F, Boerma C, et al. The effect of treatment and clinical course during Emergency Department stay on severity scoring and predicted mortality risk in Intensive Care patients. *Crit Care*. 2022; 26(1):112. DOI: 10.1186/s13054-022-03986-2
25. Makkar N, Soneja M, Arora U, Sood R, Biswas S, Jadon RS, et al. Prognostic utility of biomarker levels and clinical severity scoring in sepsis: a comparative study. *J Investing Med*. 2022; jim-2021-002276. DOI: 10.1136/jim-2021-002276
26. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA* 1993; 270:2957. DOI: 10.1001/jama.270.24.2957
27. Spagnó MF, Signori D, Comerlato PH, Tonietto TA, Caierão J, Pasqualotto AC, et al. High rate of Clostridioides difficile colonization in patients admitted to intensive care: A prospective cohort study. *Anaerobe*. 2022; 102538. DOI: 10.1016/j.anaerobe.2022.102538
28. Tian W, Meng W. Associations of Homocysteine, Procalcitonin, and D-Dimer Levels with Severity and Prognosis of Patients with Multiple Trauma. *Clin Lab*. 2022; 68(3): DOI: 10.7754/Clin.Lab.2021.210525
29. Tremli B, Breitkopf R, Bukumiric Z, Bachler M, Boesch J, Rajcic S. ECMO Predictors of Mortality: A 10-Year Referral Centre Experience. *J Clin Med*. 2022; 11(5):1224. DOI: 10.3390/jcm111051224
30. Costi S, Brogneri A, Bagni C, Pennacchi G, Beneventi C, Tabbi L, et al. Rehabilitation of Difficult-to-Wean, Tracheostomized Patients Admitted to Specialized Unit: Retrospective Analyses over 10-Years. *Int J Environ Res Public Health*. 2022; 19(5):2982. DOI: 10.3390/ijerph19052982
31. Aziz F, Reisinger AC, Aberer F, Sourij C, Tripolt N, Siller-Matula JM, et al. Simplified Acute Physiology Score 3 Performance in Austrian COVID-19 Patients Admitted to Intensive Care Units with and without Diabetes. *Viruses*. 2022; 14(4):777. DOI: 10.3390/v14040777
32. Higgins TL, Kramer AA, Nathanson BH, Copes W, Stark M, Teres D. Prospective validation of the intensive care unit admission Mortality Probability Model (MPM0-III). *Crit Care Med*. 2009; 37(5):1619-23. DOI: 10.1097/CCM.0b013e31819ded31
33. Ho KM, Lan NS, Williams TA, Harahsheh Y, Chapman AR, Dobb GJ, et al. A comparison of prognostic significance of strong ion gap (SIG) with other acid-base markers in the critically ill: a cohort study. *J Intensive Care*. 2016; 4:43. DOI: 10.1186/s40560-016-0166-z
34. Rivielo ED, Kiviri W, Fowler RA, Mueller A, Novack V, Banner-Goodspeed VM, et al. Predicting Mortality in Low-Income Country ICUs: The Rwanda Mortality Probability Model (R-MPM). *PLoS One*. 2016; 11(5):e0155858. DOI: 10.1371/journal.pone.0155858
35. Ko M, Shim M, Lee SM, Kim Y, Yoon S. Performance of APACHE IV in Medical Intensive Care Unit Patients: Comparisons with APACHE II, SAPS 3, and MPM0 III. *Acute Crit Care*. 2018; 33(4):216-221. DOI: 10.4266/acc.2018.00178
36. Chen BH, Tseng HJ, Chen WT, Chen PC, Ho YP, Huang CH, et al. Comparing Eight Prognostic Scores in Predicting Mortality of Patients with Acute-On-Chronic Liver Failure Who Were Admitted to an ICU: A Single-Center Experience. *J Clin Med*. 2020; 9(5):1540. DOI: 10.3390/jcm9051540
37. Ho KM, Williams TA, Harahsheh Y, Higgins TL. Using patient admission characteristics alone to predict mortality of critically ill patients: A comparison of 3 prognostic scores. *J Crit Care*. 2016; 31(1):21-5. DOI: 10.1016/j.jccr.2015.10.019
38. Raith EP, Udy AA, Bailey M, McLaughlin S, MacIsaac C, Bellomo R, et al. Prognostic Accuracy of the SOFA Score, SIRS Criteria, and qSOFA Score for In-Hospital Mortality Among Adults With Suspected Infection Admitted to the Intensive Care Unit. *JAMA*. 2017; 317(3):290-300. DOI: 10.1001/jama.2016.20328
39. Lambden S, Laterre PF, Levy MM, Francois B. The SOFA score-development, utility and challenges of accurate assessment in clinical trials. *Crit Care*. 2019; 23(1):374. DOI: 10.1186/s13054-019-2663-7
40. Liu Z, Meng Z, Li Y, Zhao J, Wu S, Gou S, et al. Prognostic accuracy of the serum lactate level, the SOFA score and the qSOFA score for mortality among adults with Sepsis. *Scand J Trauma Resusc Emerg Med*. 2019; 27(1):51. DOI: 10.1186/s13049-019-0609-3
41. Elias A, Agbarieh R, Saliba W, Khoury J, Bahouth F, Nashashibi J, et al. SOFA score and short-term mortality in acute decompensated heart failure. *Sci Rep*. 2020; 10(1):20802. DOI: 10.1038/s41598-020-77967-2
42. Chafranska L, Stenholt OO, Sørensen RH, Abdullah SMOB, Nielsen FE. Predictors for mortality in patients admitted with suspected bacterial infections - A prospective long-term follow-up study. *Am J Emerg Med*. 2022; 56:236-243. DOI: 10.1016/j.ajem.2022.04.002
43. Ko RE, Kwon O, Cho KJ, Lee YJ, Kwon JM, Park J, et al. Quick Sequential Organ Failure Assessment Score and the Modified Early Warning Score for Predicting Clinical Deterioration in General Ward Patients Regardless of Suspected Infection. *J Korean Med Sci*. 2022; 37(16):e122. DOI: 10.3346/jkms.2022.37.e122
44. Vincent JL, de Mendonça A, Cantraine F, Moreno R, Takala J, Suter PM, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Working group on "sepsis-related problems" of the European Society of Intensive Care Medicine. *Crit Care Med*. 1998; 26(11):1793-800. DOI: 10.1097/00003246-199811000-00016
45. Matics TJ, Sanchez-Pinto LN. Adaptation and Validation of a Pediatric Sequential Organ Failure Assessment Score and Evaluation of the Sepsis-3 Definitions in Critically Ill Children. *JAMA Pediatr*. 2017; 171(10):e172352. DOI: 10.1001/jamapediatrics.2017.2352
46. Pawar RD, Shih JA, Balaji L, Grossestreuer AV, Patel PV, Hansen CK, et al. Variation in SOFA (Sequential Organ Failure Assessment) Score Performance in Different Infectious States. *J Intensive Care Med*. 2021; 36(10):1217-1222. DOI: 10.1177/0885066620944879
47. Arakawa M, Levy JH, Fujimori K, Kondo K, Iba T. A new SOFA score calculation to improve the predictive performance for mortality in sepsis-associated disseminated intravascular coagulopathy patients. *J Crit Care*. 2021; 64:108-113. DOI: 10.1016/j.jccr.2021.03.008
48. Yang Z, Hu Q, Huang F, Xiong S, Sun Y. The prognostic value of the SOFA score in patients with COVID-19: A retrospective, observational study. *Medicine (Baltimore)*. 2021; 100(32):e26900. DOI: 10.1097/MD.00000000000026900
49. Loyrion E, Agier L, Trouve-Buisson T, Gavazzi G, Schwebel C, Bosson JL, et al. Dynamic SOFA score assessments to predict outcomes after acute admission of octogenarians to the intensive care unit. *PLoS One*. 2021; 16(8):e0253077. DOI: 10.1371/journal.pone.0253077
50. Li Y, Yan C, Gan Z, Xi X, Tan Z, Li J, et al. Prognostic values of SOFA score, qSOFA score, and LODS score for patients with sepsis. *Ann Palliat Med*. 2020; 9(3):1037-1044. DOI: 10.21037/apm-20-984
51. Kilinc Toker A, Kose S, Turken M. Comparison of SOFA Score, SIRS, qSOFA, and qSOFA + L Criteria in the Diagnosis and Prognosis of Sepsis. *Eurasian J Med*. 2021; 53(1):40-47. DOI: 10.5152/eurasianjmed.2021.20081
52. Rodriguez F, Solomon N, de Lemos JA, Das SR, Morrow DA, Bradley SM, et al. Racial and Ethnic Differences in Presentation and Outcomes for Patients Hospitalized With COVID-19: Findings From the American Heart Association's COVID-19 Cardiovascular Disease Registry. *Circulation*. 2021; 143(24):2332-2342. DOI: 10.1161/CIRCULATIONAHA.120.052278
53. Gershengorn HB, Holt GE, Rezk A, Delgado S, Shah N, Arora A, et al. Assessment of Disparities Associated With a Crisis Standards of Care Resource Allocation Algorithm for Patients in 2 US Hospitals During the COVID-19 Pandemic. *JAMA Netw Open*. 2021; 4(3):e214149. DOI: 10.1001/jamanetworkopen.2021.4149
54. Ashana DC, Anesi GL, Liu VX, Escobar GJ, Chesley C, Eneanya ND, et al. Equitably Allocating Resources during Crises: Racial Differences in Mortality Prediction Models. *Am J Respir Crit Care Med*. 2021; 204(2):178-186. DOI: 10.1164/rccm.202012-4383OC
55. Wunsch H, Hill AD, Bosch N, Adhikari NKJ, Rubenfeld G, Walkey A, et al. Comparison of 2 Triage Scoring Guidelines for Allocation of Mechanical Ventilators. *JAMA Netw Open*. 2020; 3(12):e2029250. DOI: 10.1001/jamanetworkopen.2020.29250
56. Rivielo ED, Dechen T, O'Donoghue AL, Cocchi MN, Hayes MM, Molina RL, et al. Assessment of a Crisis Standards of Care Scoring System for Resource Prioritization and Estimated Excess Mortality by Race, Ethnicity, and Socially Vulnerable Area During a Regional Surge in COVID-19. *JAMA Netw Open*. 2022; 5(3):e221744. DOI: 10.1001/jamanetworkopen.2022.1744
57. Butler CR, Webster LB, Sakata VL, Tonelli MR, Diekema DS, Gray MM. Functionality of Scarce Healthcare Resource Triage Teams During the COVID-19 Pandemic: A Multi-Institutional Simulation Study. *Crit Care Explor*. 2022; 4(1):e0627. DOI: 10.1097/CCE.0000000000000627