Sepsis is a life-threatening organ dysfunction caused by the dysregulation of the host’s response to infection, whose incidence rate and mortality have been high for a long time. The occurrence and development of sepsis is complex, which is the result of the interaction of many mechanisms and factors. Patients with sepsis often have a hidden onset and are easy to be ignored, resulting in the delay of treatment. Sepsis has long been a focus in the medical world because of its acute onset, high mortality, high consumption of medical resources, while early identification and prevention of sepsis can reduce its mortality. With the deepening of sepsis research, there are numerous scoring systems available for diagnosing and evaluating sepsis at the moment; however, no scoring system can completely and accurately evaluate the prognosis of patients with sepsis. Some scoring systems are too cumbersome, while others limit the time and occasion of use. This review briefly reviewed the composition, development, and clinical validation of those current international scoring systems for sepsis prognosis, to compare their shortcomings and advantages based on the clinical studies. The purpose of this review is to provide further guidance for the selection of prognostic evaluation methods of sepsis.

Keywords: Sepsis; score; prognosis; mortality

Received: 30 April 2021; Accepted: 17 June 2022; Published: 10 July 2022.
doi: 10.21037/biotarget-21-5
View this article at: https://dx.doi.org/10.21037/biotarget-21-5

Sequential Organ Failure Assessment (SOFA)

In 1994, scholars from the European Society of Intensive Care Medicine (ESICM) first proposed the SOFA score at a conference organized in Paris, to evaluate the degree of sepsis; however, because each scoring system has inherent advantages and disadvantages, it is necessary to compare and explore the scores further to develop a more accurate and timely prognosis evaluation system for septic patients. This review aims to systematically review the research progress of pertinent scoring standards for patients with sepsis.
organ failure as efficiently and objectively as possible (2). Those patients were scored from the respiratory system, coagulation system, liver function, cardiac function, nervous system, and renal function, while the worst score was recorded every day. SOFA score was first used to evaluate the condition of patients with sepsis, which has been widely used by more and more clinicians to evaluate the condition of critical patients caused by various reasons now. In clinical, the relative accuracy of SOFA score for septic patients’ severity is also widely recognized. In the 2016 revision of the “The Third International Consensus Definitions for Sepsis and Septic Shock” (1), not only the definition of sepsis was revised to emphasize the importance of organ failure in the occurrence and development of sepsis, but also the SOFA score was used to evaluate the degree of organ failure to facilitate the diagnosis of sepsis. At the same time, a quick sepsis-related organ failure assessment (qSOFA) score was proposed after the statistics and analysis of the database, aiming to screen and identify septic patients in the early stage by respiratory rate, the state of consciousness, and blood pressure. Compared with the SOFA score, the qSOFA score is more convenient and rapid, but its accuracy in the diagnosis and prognosis of sepsis has been also controversial. Raith et al. (3) evaluated the relationship between SOFA, systemic inflammatory response syndrome (SIRS), qSOFA, and prognosis accuracy of 184,875 patients in the intensive care unit (ICU), and the results showed that the SOFA score was more accurate for prognosis than SIRS and qSOFA score in ICU. At the same time, in a study on 28- and 90-day mortality of patients with sepsis (4), the accuracy of the SOFA score in evaluating the prognosis of patients with sepsis is much higher than qSOFA, despite that the qSOFA score being more concise and rapid. Due to the complexity of sepsis pathophysiology, many biomarkers have been confirmed to play a critical role in the diagnosis and early warning of sepsis, while biomarkers combined with SOFA score also showed relatively high accuracy in predicting the prognosis of sepsis. A multicenter study confirmed that the SOFA score combined with Apache II score and albumin has a higher predictive value for mortality in patients with sepsis than the single SOFA score. SOFA combined with the heparin-binding protein (HBP) (5), the acute gastrointestinal injury (AGI) grading system (6), the lactate (7), procalcitonin (PCT), and C-reactive protein (CRP) (8) all showed sound results in the diagnosis and prognosis evaluation of patients with sepsis. Although the usage of SOFA scores is increasingly frequent currently, there are still certain areas to be addressed. First of all, intestinal-related indicators are omitted due to their difficulty to obtain at the time the SOFA score was first established. However, the changes in intestinal flora (9) and the damage to intestinal mucosal barrier function (10) play a critical role in the onset and progression of sepsis. Second, a large number of research results have confirmed that (11-13) the SOFA score is still slightly inferior in the evaluation of the prognosis of septic patients compared with other scoring systems, although the SOFA score has a considerable effect on the diagnosis, and continuous SOFA score can help to observe the changes in septic patients’ condition.

**Acute Physiology and Chronic Health Evaluation (APACHE II)**

In 1981, after three years of research, the medical team led by Dr. Knaus of the University of Washington put forward the APACHE I score, which can objectively assess the condition of patients in ICU by quantifying the degree of disorder in physiological changes, and classify patients according to the severity of the disease. APACHE I score consists of two parts: (I) acute physiology score (APS), 34 physiological parameters of patients (including blood lactic acid, blood pH, creatinine, etc.) were recorded and scored within 32 h after admission; (II) chronic health score (CHS), whose purpose is to know the health status of the patients 3 to 6 months before admission, such as whether the patients go to see the doctor every week, whether they can't go to work because of illness, whether they are limited in activities, etc. However, due to the complexity of the content of the APACHE I score and the results from clinical verification, although the APACHE I score has great value in predicting the prognosis of ICU patients, there is a certain degree of deviation in predicting the individual prognosis of non-ICU patients (14). Therefore, after nearly four years of statistical research on a large number of clinical samples, the team modified the content of the APACHE I score and put forward the APACHE II score standard in 1985, which was based on age, CHS, state of consciousness, and physiological index (including temperature, respiratory rate, white blood cells, etc.), and then the results were summarized. At the same time, they also put forward a formula to calculate the expected mortality (15). APACHE II score can be used not only for ICU patients but also for the evaluation and prediction of individual patients without ICU. In 1991, the team carried out a statistical analysis of the clinical data of 17,440 patients and found that there was a certain
deviation in the prediction of the mortality of patients with a low score by APACHE II. Therefore, they proposed the APACHE III score based on APACHE I and APACHE II scores to further improve the accuracy of the prediction of the hospital mortality of patients. The APACHE III score was based on whether the patients could open their eyes to pain or language stimulation and the degree of language and motor impairment, instead of the original Glasgow score. The age and CHSs were modified, and five physiological indexes, such as liver function and renal function, were added (16). Compared with the previous two generations of APACHE scores, APACHE III score not only improves the ability to predict the death risk of critically ill patients but also improves its simplicity and accessibility. However, with the continuous improvement of the medical level and policy in ICU, the length of ICU stay has decreased. Zimmerman et al. (17) made a statistical analysis of clinical data of 11,629 people in the United States from 2002 to 2003 to modify the APACHE III, then the APACHE IV was constructed, which is also the latest generation of APACHE scores in the world so far. It increased to 116 diseases based on the original 96 diseases and added five parameters such as mechanical ventilation and thrombolytic therapy for AMI. Besides, APACHE IV excludes other ICU patients and replaces the missing laboratory indicators with the indicators of the previous day. Moreover, an electronic scoring system is designed to facilitate the collection and calculation of data. The APACHE IV improves the accuracy of the prediction of ICU patients’ length of stay and mortality. However, due to the differences in the accuracy of the evaluation and prediction of APACHE IV score for patients with various diagnoses, and the statistical analysis based on the clinical data of the American population, there is a certain bias in the representativeness. Therefore, some Asian scholars have questioned whether the assessment of APACHE IV in different areas and physical groups is also applicable. Some scholars collected the clinical data of 82 elderly patients with sepsis in China and concluded that APACHE IV score underestimated the mortality of elderly patients with sepsis (18). Other scholars applied the APACHE IV score to Surgical Abdominal Sepsis (SABS) patients but found that the ability of the APACHE IV score to evaluate the prognosis of SABS patients is poor (19). Since the APACHE IV score has been used in clinical practice for a short time, there are few studies and validations on the APACHE IV score at present. Therefore, its clinical application is less common than that of APACHE II.

Because of its simplicity, applicability, and relative accuracy, the APACHE II score is currently recognized by most clinicians and has also been studied and verified by a large number of experts, which has been widely used in clinical practice (20). APACHE score also showed a good ability to evaluate the prognosis of patients with sepsis. Sadaka et al. (21) compared the ability of Apache II and Apache III to evaluate the prognosis of patients with sepsis, both of which showed a high capacity for evaluation, and there was no significant difference (P>0.005). However, compared with the APACHE II score, the collection of parameters involved in the APACHE III score is a more time-consuming and complicated process. By comparing the APACHE II score with the SOFA score, some scholars have concluded that although the APACHE II score on the day of admission of septic patients correlates with their prognosis, the SOFA score on the first week of admission, especially on the third day, can more accurately evaluate the prognosis of septic patients (22). In recent years, with the wide application of the APACHE II score, to further improve the evaluation level of APACHE II score for the prognosis of patients with sepsis, many scholars have actively explored the indicators (such as microperfusion level, coagulation, and infection) that may be missing in APACHE II score which has a certain correlation with the occurrence and development of sepsis and construct a joint prediction model with APACHE II Model. Among them, the APACHE II score with the red cell distribution width (23), thyroid hormone profile (24), uric acid concentration (25) all showed a strong predictive ability for the prognosis of patients with sepsis, whose accuracy was higher than that of single APACHE II score.

**Mortality in emergency department sepsis score (MEDS)**

MEDS was first proposed by Shapiro in 2003, which is a fast, concise, and targeted scoring system based on the objective conditions of the emergency department. Shapiro et al. (26) carried out a prospective study, in which early fluid resuscitation was included in the observation indicators, while nine indicators related to mortality were identified from a large number of variables. Compared with other scoring systems, MEDS was more concise with the fixed score, whose validation results suggested a good ability to predict the mortality risk of emergency patients with sepsis.

Pong et al. (12) analyzed the 30-day in-hospital mortality
and 72-hour in-hospital mortality of 249 patients with suspected sepsis and SIRS in Singapore from 2014 to 2016 and compared the scores of MEDS, APACHE II, and SOFA. The results showed that the area under the curve (AUC) of MEDS was 0.87 and 0.88, respectively, which were higher than the APACHE II score and SOFA score, while MEDS showed good specificity. Another study confirmed that MEDS is more accurate than the qSOFA score in evaluating the 30-day mortality of emergency septic patients (13). Some scholars also concluded that the MEDS score had a strong positive correlation with the 28- and 90-day mortality of septic patients (27). MEDS is better than the NEWS score, Charlson Comorbidity Index (CCI), and SOFA score in the risk stratification and prognosis evaluation of emergency sepsis patients. It is critical in identifying high-risk individuals early, early warning, diagnosing, and intervening in sepsis (28,29). Some scholars believe that the “terminal illness, alternate mental status” in MEDS are not only affected by race and the establishment of norms but also lack objective evaluation criteria, so the overall evaluation results are interfered with by human factors to a certain extent (30). Therefore, further clinical data are needed to verify the MEDS and improve its rigor and objectivity.

**Simplified acute physiology score II (SAPS II)**

SAPS II is a scoring system proposed by Le Gall in 1993, which is helpful to assess the risk of death in ICU patients. It is based on the statistical analysis of the data of 13,152 clinical patients (31). In 1984, Le Gall adjusted the deficiency of APACHE I score, deleted the chronic physiology score in APACHE I, and added the age reference index, to improve the accuracy of the APACHE I score in the assessment of in-hospital mortality of ICU patients, naming the improved scoring system as SAPS I (32). Then, to further improve the accuracy of SAPS I prediction and accurately calculate the death probability of patients, Le Gall used logistic regression analysis to assign values to evaluation parameters based on SAPS I and constructed SAPS II, which includes 17 reference indexes, such as white blood cells, and serum bilirubin and also puts forward Predicted Hospital Mortality (PHM). Some scholars collected the clinical data of 2,470 patients with sepsis from the medical information mart for intensive care III (MIMIC-III) database and concluded that the accuracy of the SAPS II score for the prognosis of patients with sepsis was higher than the SOFA score through calculation and analysis (33). However, SAPS II did not contain any coagulation-related parameters, since Sepsis is a process that involves inflammation and coagulation. In addition to the SAPS II and SOFA scores, antithrombin and prothrombin ratio were indicated to have an evaluation effect on the mortality of patients with sepsis, whose evaluation accuracy was better than that of SAPS II and SOFA scores (34). Studies have displayed that when the SAPS II score 60, the intravascular tissue factor increased significantly, which was positively correlated with the mortality of patients with sepsis, and greatly increased the risk of death of patients with sepsis (35), and it may further increase the accuracy of the SAPS II score in the prognosis evaluation of patients with sepsis by actively looking for coagulation-related parameters affecting the development of sepsis and exploring their correlation with SAPS II or combined application with SAPS II. SAPS II is based on clinical data from European and North American populations. However, there are few studies on the clinical research and validation of SAPS II from Asian cases with sepsis, while its applicability needs to be verified by further clinical data.

**Rapid Emergency Medicine Score (REMS)**

REMS was derived from the rapid acute physiology score (RAPS), which was proposed by Olsson in 2004 (36). RAPS is a reduced version of the APACHE II score proposed by Rhee in 1987, where all parameters can be obtained in all patients in case of emergency, including pulse, blood pressure, respiratory rate, and Glasgow Coma Score (GCS), to help doctors and nurses to evaluate the condition of critically ill patients before transportation, being fast, simple, and convenient (37). Olsson included peripheral oxygen saturation and age into the raps score to construct the REMS and conducted a prospective study on 12,006 emergency non-surgical patients, whose results suggest that the REMS not only retains the rapidity and convenience of the RAPS but also shows a good predictive ability for the prognosis of emergency non-surgical patients (36). At present, REMS has been validated as a suitable tool for evaluating the condition of critically ill patients in the emergency department (38), but few studies on its evaluation ability, sensitivity, and specificity of prognosis of patients with sepsis, and its applicability to the Asian population level. In comparison to other scores, the REMS is more succinct, speedier, and better suited to the emergency department environment, which serves as the first line of defense against sepsis. Early risk
stratification and prognosis evaluation of sepsis patients can aid in intervention measures in the process of early sepsis occurrence and development and improve the survival rate. Increasing the REMS's clinical utility for sepsis patients and further exploring its use in evaluating septic patients' situations should be predicted to contribute to improving the prognosis of sepsis patients.

**National Early Warning Score (NEWS)**

MEWS is an improved version of the early warning score (EWS), which was formally proposed by Subbed in 2001, the purpose of which is to identify patients with acute diseases and potentially critical diseases at an early stage. MEWS is composed of five indexes: temperature, heart rate, blood pressure, respiratory rate, and altered mental status. The higher the score is, the more serious the condition is and the worse the prognosis is (39). In 2012, the Royal College of Physicians (RCP) added oxygen saturation and any supplementary oxygen to form NEWS based on the MEWS, aiming to play an early warning role in the condition of adult critically ill patients. Now some scholars gradually use it as one of the tools for risk stratification and prognosis evaluation of emergency septic patients (40). Ruangsomboon et al. (41) statistically analyzed the clinical data of 1,622 emergency patients with suspected sepsis in Thailand, of which results demonstrated that NEWS and REMS had a better early warning effect on sepsis than the qSOFA score and SIRS, as well as a similar predictive ability, without significant difference (P>0.005). However, the REMS is more accurate in predicting the 7-day mortality of septic patients. Some studies have proved that NEWS can be used as an effective tool to evaluate the prognosis of patients with sepsis, but its specificity is poor. It is necessary to further explore the clinical risk factors related to sepsis and improve the specificity of the NEWS (42). A previous retrospective study by Goulden et al. (43) also found that the prediction of NEWS relative to the mortality of sepsis patients was similar to the qSOFA score, and there was no significant statistical difference between them (P>0.005), which was more accurate than the SIRS score. However, the sensitivity of NEWS was higher than the qSOFA score with poor specificity. Goulden believes that any scoring system for identifying sepsis should be more sensitive than specific, as the consequences of false negatives (delayed or missed treatment) are arguably much greater than those of false positives (unnecessary antibiotics). However, some scholars have proposed that the sensitivity of NEWS for patients with structural brain injury is low, so patients with intracranial lesions should be excluded from the use process (44). NEWS is widely used in the Commonwealth of Nations, but the clinical applications in other regions and countries are rarely reported. The ability of the NEWS to evaluate the prognosis of patients with sepsis and whether it applies to sepsis caused by various causes still needs to be verified by a large number of clinical data.

**Oxford Acute Severity of Illness Score (OASIS)**

OASIS is a relatively simple scoring system for disease severity based on the APACHE IV score constructed by Johnson in 2013 (45), who used the clinical data of 9,786 patients in 86 ICUs of 49 hospitals in the United States as the model, and the clinical data of 23,618 patients as the external validation of the prediction model. Johnson selected 10 variables at any point during day 1 and admission following elective surgery as being important for developing the new severity score by the genetic algorithm (GA). Finally, OASIS includes age, pre-ICU admission length of stay, GCS, heart rate, mean arterial pressure (MAP), temperature, respiratory rate, urine output, ventilator, and elective surgery. He also verified that the OASIS not only simplifies the APACHE IV scoring system but also has better efficacy in evaluating the prognosis of critically ill patients than the APACHE IV score. In 2019, a prospective study was conducted on the evaluation ability of OASIS for the prognosis of septic patients, of which results prove that OASIS does have a certain evaluation ability for sepsis patients. Although it is slightly lower than the SOFA score, no significant difference is found when the SAPS II is less than 50 (P>0.005), but for patients with SAPS II greater than 50, the OASIS should be used cautiously (46). Although the OASIS is slightly inferior to the SOFA score in the prognosis evaluation of patients with sepsis, its simplicity and easy access to reference indicators have been also very attractive. On the whole, the OASIS has great potential and exploitability for the prognosis evaluation of patients with sepsis.

**Charlson’s weighted index of comorbidities (WIC)**

WIC, a quantitative scoring system that quantifies the previous underlying diseases of patients, was first proposed by Charlson et al. in 1987, to assess the risk of death in patients, while all parameters can be obtained from the patient's medical records without any laboratory indicators.
and analyzed the inpatient medical records and 1-year follow-up information of 559 patients. A total of 19 comorbidities significantly related to the prognosis of patients were included by Cox regression models, including diabetes, leukemia, liver disease, etc. Then, they combined the severity of the disease to give weight and finally constructed the WIC. Gradually, as WIC increased, patient mortality climbed. Subsequently, Charlson included 685 breast cancer patients to further verify the evaluation performance of WIC, which showed a high correlation with mortality and had a good performance in evaluating the prognosis of patients (47). Some clinical studies displayed that when the WIC is less than or equal to 2, the mortality rate is 18.18%, and when the WIC is more than 2, the mortality rate is 29.41%. WIC is related to the risk of death, but its ability to predict the prognosis of patients with sepsis is less than the SOFA score (48). Zhang et al. (49) carried out a prospective study on the 28-day mortality of 118 patients with sepsis. They used PCT, SOFA score, WIC, and the prediction model for the 28-day prognosis of patients with sepsis. The PCT, SOFA score, WIC, and their combined prediction model were used to compare the 28-day prognosis of patients with sepsis, while those results showed that the AUC of the prediction model for the 28-day prognosis of patients with sepsis was the largest, which was 0.943, indicating that the combined prediction of the three factors was relatively the most accurate. When evaluating the prognosis of patients with sepsis, some scoring systems will exclude patients with previous diseases to minimize the impact of previous diseases on therapy and to enhance overall prediction accuracy, although this will undoubtedly result in some degree of selective bias. Compared with other scoring systems, WIC has the advantages of convenience and low cost, since it does not contain any time-consuming reference indicators. At the same time, it closely links the previously combined diseases with the prognosis of sepsis. Although WIC is not good at evaluating the prognosis of sepsis patients, it can be considered as a supplementary part of other scores to better improve the accuracy of the prognosis of patients with sepsis.

**Sepsis Severity Score (SSS)**

SSS is a relatively new clinical score for patients with sepsis, which was first proposed in 2014, the formation of which came from the clinical data of 23,428 patients in the surviving sepsis campaign database. All patients were diagnosed with severe sepsis or septic shock, while the variables related to the mortality rate of sepsis were selected. Finally, a logistic regression model was constructed including 34 clinical variables, and they evaluated and verified the model, while the results showed that the calibration and fitting of the model were good, and the AUC area in the model group and the validation group were 0.736 and 0.7348, respectively, which was relatively accurate in evaluating the prognosis of patients with sepsis (50). Khwannimit et al. (51) compared the SSS with other related scores. In terms of the ability to assess the in-hospital mortality of patients with sepsis, it was concluded that although SSS had a good ability to assess the mortality of patients with sepsis, there was no significant difference between SSS and APACHE II score, SAPS II (P>0.005), and the calibration and fitting degree were poor. The establishment of SSS is based on the clinical data of patients with severe sepsis or septic shock, which has selective bias. Therefore, Khwannimit et al. (51) suggested that the SSS should be applied to more comprehensive septic patients, including different severity, and different races, to further analyze its applicability and evaluation ability. Some scholars have come to the same conclusion that SSS does have a certain predictive ability on the mortality evaluation of septic patients, but the calibration degree is poor, while the effectiveness of the SSS should be evaluated again to improve its performance of the SSS (5,11,52). Therefore, the current SSS is not mature, and its external verification is still less. It should be revised and reformulated in conjunction with additional research and verification to improve the deficit and prediction ability, as well as to facilitate future clinical use.

**Conclusions**

In recent years, as public understanding of sepsis has deepened, the pathophysiological mechanism of sepsis has been explored continuously, the treatment plan has been improved constantly, and the importance of early diagnosis, early treatment, and early warning of high-risk septic patients has been discovered continually, various scoring systems related to sepsis prognosis have gradually emerged, while the classic scoring system has also been better and better. But so far, no scoring system can be used to evaluate the prognosis of sepsis very accurately, while now an increasing number of scholars and specialists have recognized the predictive model's structure. The combined application of various parameters and scoring systems can
make up for those deficiencies, which is of great help to improve the accuracy of prognosis evaluation of sepsis. At present, the majority of the classic scoring systems widely used in the world are based on the statistical analysis of clinical data of patients in Europe and the United States. Due to different countries or regions that have varying levels of understanding and treatment of sepsis, the applicability of these scoring systems to diverse regions and races also requires a significant number of further clinical investigations.

Acknowledgments

Funding: This study was supported by the National Natural Science Foundation of China Youth (No. 81801893).

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://biotarget.amegroups.com/article/view/10.21037/biotarget-21-5/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

15. Knaus WA, Draper EA, Wagner DP, et al. APACHE II:


39. Gole AR, Srivastava SL, Neeraj. Prognostic accuracy of
SOFA score, SIRS criteria, NEWS and MEWS scores for in-hospital mortality among adults admitted to ICU with suspected sepsis. J Assoc Physicians India 2020;68:87.


44. Teasdale GM. National early warning score (NEWS) is not suitable for all patients. BMJ 2012;345:e5875.


