J Crit Care 90 (2025) 155174

Contents lists available at ScienceDirect

Journal of Critical Care

journal homepage: www.journals.elsevier.com/journal-of-critical-care

External validation of eight different models to predict sepsis mortality in intensive care units

Satyen Hargovan [a,b], Charlotte Simpson [c], Sayonne Sivalingam [d], Angus Carter [e], Ronny Gunnarsson [f,g,h,*]

[a] Department of Medicine, Cairns Hospital and Hinterland Health Service, Cairns, Queensland, Australia
[b] Adjunct Lecturer in Medicine, College of Medicine and Dentistry, James Cook University, Queensland, Australia
[c] Registrar in Intensive Care Medicine, Sunshine Coast University Hospital, Queensland, Australia
[d] Registrar, The Prince Charles Hospital, Brisbane, Queensland, Australia
[e] Department of Intensive Care Medicine, Bendigo Hospital, Victoria, Australia
[f] Research, Development, Education and Innovation, Primary Health Care, Region Västra Götaland, Sweden
[g] Centre for Antibiotic Resistance Research (CARe) at University of Gothenburg, Sweden
[h] General Practice/Family Medicine, School of Public Health and Community Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

ARTICLE INFO

Keywords:
Sepsis
Intensive care units
Mortality prediction
Area under curve
Validation study
Prognosis

ABSTRACT

Purpose: Sepsis is a complex, heterogenous syndrome defined as life-threatening organ dysfunction due to severe infection. Existing mortality prediction models may not adequately capture the complexities of sepsis. The objectives of this study were twofold; to clarify to what extent variables belonging to eight different mortality prediction models used in intensive care units (ICU) were collected in routine medical care, and to externally validate these models.

Material and methods: A retrospective cohort of 750 patients admitted to three ICU's with a final diagnosis of sepsis at ICU discharge were included. Mortality prediction models were evaluated by calculating the area under receiver operating curve (AUROC) for their ability to predict 30-day mortality.

Results: The CSM-4, when used 4 h after ICU admission, predicted ICU episode-of-care mortality best with an AUROC of 0.80. It used only a few variables which are frequently retrieved in routine medical care. ANZROD 24 was the best performing model to be applied 24 h after admission with AUROC of 0.83.

Conclusions: Time after admission may decide which prediction model is most useful. Early after ICU admission, the sepsis-specific CSM-4 mortality prediction model performed slightly better than other models. However, at 24 h after admission general models not specific for sepsis, like the ANZROD 24, performed well.

1. Introduction

Sepsis is a leading cause of morbidity and mortality globally, accounting for >50 million cases-per-year and causing 11 million deaths, or 19.7 % of deaths globally [1–3]. Sepsis is associated with individual patient morbidity, loss of quality of life, significantly increased resource use, financial cost, complications, increased length-of-stay and mortality. Despite this, recognition of septic adults, and those at high-risk of dying, remains a challenge [4]. This is crucial as early identification of at-risk septic patients, and their timely management with high-quality care, improves sepsis outcomes [5,6].

1.1. Being septic or having sepsis

Traditionally, sepsis has been considered as a severe illness with bacteria in the blood stream [7]. This definition has changed several times, with the current Sepsis-3 criteria considering sepsis a "life-threatening organ dysfunction caused by a dysregulated host response to infection" [1,8]. This current definition does not require bacteria in the blood stream. It can even be a viral infection. However, the transition from an older traditional definition to a new definition has not been completely implemented in clinical practice, where studies have demonstrated that only a third of patients in an intensive care unit (ICU) that fulfil the Sepsis-3 criteria are discharged with a final diagnosis of

* Corresponding author.
E-mail address: ronny.gunnarsson@gu.se (R. Gunnarsson).

https://doi.org/10.1016/j.jcrc.2025.155174
Received 7 February 2025; Accepted 3 July 2025
Available online 9 July 2025

sepsis [9,10]. Hence, not all patients fulfilling Sepsis-3 criteria are in clinical practice labelled as having had sepsis, and not all patients diagnosed with sepsis fulfil the Sepsis-3 criteria [10]. In reality, there is no gold standard diagnostic test for sepsis [9] and even expert panels may disagree on the final diagnosis [11]. Hence, there is always an uncertainty in previous retrospective research studies if included patients fulfilled the Sepsis-3 criteria or were only discharged with a clinical diagnosis of sepsis.

1.2. Mortality prediction models

Mortality prediction models estimate risk of death by combining multiple variables strongly correlated to mortality. Their ability to help risk stratify, prognosticate, assist patient-clinician discussion and shared decision-making make them a useful clinical decision support tool. Furthermore, they can help researchers to improve and evaluate safety and quality care via calculation of standardised mortality ratios and compare patient cohorts in randomised control trials [12,13].

Prediction models are often assessed at the first internal validation on calibration statistics, such as the Hosmer-Lemeshow goodness-of-fit test, discrimination statistics, such as the area-under-receiver-operator-curve (AUROC) and global fit statistic such as the Nagelkerke R2 statistic. Other desirable features include being globally applicable to the specific population of interest, clinically easy-to-use and calculated on cheap and widely available variables that best fit the relevant condition [14]. However, most models have limited clinical utility [15] and is rarely used in decision-making [16].

Although several mortality prediction models exist for use in critically ill patients most of them are not sepsis-specific, and their performance in septic patients is not well explored. Non-sepsis-specific models may not adequately capture the complexities of underlying sepsis pathophysiology. Furthermore, existing mortality prediction models vary in their time of use. Whilst some are meant to be used at time of ICU admission, within 4-h, others within 24-h and some up to 72-h.

This study had two equally important objectives; to clarify to what extent variables belonging to eight different mortality prediction models often used in ICU were collected in routine medical care and to externally validate these models.

2. Material and methods

A retrospective cohort study was performed across three hospital ICU's. The Townsville Hospital Human Research Ethics Committee provided over-arching ethical approval for the study (HREC/QTHS/91805), site-specific applications were also approved, and a Public Health Act waiver of consent allowing access to data was granted.

2.1. Eligibility criteria

Varied geography, demographics, patient case-mix and hospital size were considered to ensure generalisability of results. The three hospitals included were

1. Cairns and Hinterland Hospital and Health Service (CHHHS), a 531-bed tertiary referral centre in rural, tropical Far North Queensland, Australia serving a population of approximately 240,000; a large proportion of whom identify as Aboriginal and/or Torres Strait Islanders.
2. Sunshine Coast University Hospital (SCUH) – a 745-bed tertiary referral centre located in regional, subtropical Queensland serving a population of approximately 317,000.
3. The Prince Charles Hospital (TPCH) – a 673-bed, tertiary predominantly cardiothoracic referral hospital in metropolitan Brisbane, Queensland, Australia sharing service to a population of approximately 2 million.

The last 250 consecutive adults (aged ≥18) admitted to the intensive care unit of each of the hospitals from December 2022 and backwards with a final diagnosis at discharge of sepsis were included, totalling 750 patients. Identification of patients with sepsis were provided by the Australian and New Zealand Intensive Care Society Centre for Outcome and Resource Evaluation (ANZICS CORE) database. This is a bi-national peer review and quality assurance program which has provided audit and analysis of the performance of Australian and New Zealand ICU's since 1992.

A thorough chart review was done after identification of patients. Sepsis had to be suspected at admission to ICU and it also had to be the final diagnosis at discharge from the intensive care unit. Furthermore, discharge diagnosis had to be compatible with Acute Physiology and Chronic Health Evaluation (APACHE) III-j diagnostic codes 501–504: non-urinary sepsis, urinary sepsis, non-urinary sepsis with shock, and urinary sepsis with shock, respectively. These patients were then double-checked by the authors against the Sepsis-3 criteria [8], ensuring that patients had a change in acute SOFA score of ≥2 from baseline and that infection was suspected based on clinical documentation and ordering of relevant investigations to investigate infection (blood cultures, for example). Patients that did not meet the above criteria were not included.

2.2. Data collection

Hospital unit record number (URN) of patients from the ICU's of CHHHS, SCUH and TPCH that met the inclusion criteria were collected from the ANZICS (Australia and New Zealand Intensive Care Society) APD (Adult patient database). The ICU integrated electronic medical record (IEMR) database MetaVision (iMDSoft, Tel Aviv, Israel ®) was browsed for each URN by 3 clinicians (SH, SS and CS) and data were collected using a standardised data sheet adhering to definitions stated in the original publications of each prediction model of interest [16,17]. Updated versions of original publications were not used [17]. Once the data were collected a cleaning process was undertaken to minimise human error, manually reviewing inputs greater than 2 standard deviations away from the mean.

Variables that described the cohort were collected such as age, gender, most likely source of sepsis and 30-day mortality. Other variables obtained were those needed to calculate the mortality prediction models of interest; quick sequential organ failure assessment (qSOFA) [8], mortality probability model-admission (MPMo-III) [18], simplified acute physiology score (SAPS-III) [19,20], the 4-Hour Cairns Sepsis Model (CSM-4) [21], the Australian and New Zealand Risk of Death (ANZROD 24) [22], the acute physiology and chronic health evaluation (APACHE-II [23] and APACHE-IIIj [24]) and the sequential organ failure assessment (SOFA) [25]. See supplemental file 1 for more details about each prediction model. Imputation was not made for missing data. Hence, the prediction by each model was only calculated if all included variables were available [17].

2.3. Statistical analysis

We calculated the percentage patients where all variables required for calculating a prediction score was readily available in routine medical care. If all required variables were available, they were used to calculate scores for each mortality prediction model. For CSM-4 we obtained the raw beta coefficients for each included predictor and used this to calculate the predicted probability for death for each patient (Supplemental file 2).

The performance of these eight models to predict 30-day mortality was estimated by calculating the area under receiver operating curve (AUROC) with 95 % confidence interval. The AUROC analysis was performed using the software SPSS version 29.0.1.1 (IBM Corp, Armonk, NY ®). The AUROC for each prediction model was compared pairwise with the method described by Hanley & McNeil [26,27] using MedCalc

[28]. We strived for an overall level of significance of 0.05. We adjusted for multiple testing according to Bonferroni's method to decide the level of significance for each pairwise comparison. Since we performed 12 pairwise comparisons the level of significance was set to 0.0042. We also calculated the Brier score as an estimate of calibration for all models except qSOFA.

2.4. Sample size calculation

The CSM-4, the most recently published mortality prediction model, is intended for use early after ICU admission in patients where sepsis is suspected. This model was used for our sample size estimation. The internal validation study for CSM-4 had a 12 % ICU sepsis case fatality rate, with the model predicting mortality with an AUROC 0.90 (0.85–0.95, $p < 0.05$). Assumptions made were that the AUROC in this external validation will be weaker, with a point estimate of the AUROC being either 0.80 (or 0.85), and that the lower limit for the 95 % should not go below 0.70. With these scenarios in mind, at least 69 (or 29) fatalities and 552 (or 232) survivors were required to be included to reach statistical significance, a total of 621 (or 261) patients. This was rounded up to 750 total patients. Sample size estimation was made using Medcalc version 19.8 (MedCalc Software Ltd. ®).

3. Results

Retrospectively from December 2022, we identified 750 patients that met inclusion criteria, 250 from each participating hospital. On average, included patients were 63 years old (standard deviation 17 years, min-max 18–97 years). Most (n = 451, 60 %) were male (Table 1). The average ICU length of stay was 104 h (min-max 3,2–1300).

Most variables for seven of the models were found in the medical electronic charts. However, for SAPS-III required variables were only found in 47 % of cases. The mortality among patients who had a full record for SAPS-III were 9.6 % compared to 13 % if SAPS-III could not be retrieved (p = 0.20, two-sided chi-square). In most variables the patient's condition was similar between hospitals indicating that the three hospitals cared for patients that were similar (Table 2).

Table 1
– Cohort of ICU sepsis patients used for external validation.

	CHHHS[a] N = 250	SCUH[a] N = 250	TPCH[a] N = 250	All hospitals N = 750
Demographic data				
Female/Male	103/147	101/149	95/155	299/451
Age; years (standard deviation)	61 (16)	65 (16)	63 (17)	63 (17)
Admitted from other hospital ward	62	60	97	219
History of renal disease	25	24	13	62
Most likely source of sepsis % (n)				
Abdominal	13 % (32)	16 % (39)	14 % (34)	14 % (105)
Bones/joints	1.6 % (4)	0.0 % (0)	4.8 % (12)	21 % (16)
Genitourinary	18 % (46)	29 % (73)	23 % (57)	23 % (176)
Lungs	37 % (93)	10 % (25)	31 % (77)	26 % (195)
Lungs+Genitourinary	0.0 % (0)	0.0 % (0)	0.40 % (1)	0.13 % (1)
Brain, brainstem or other nerves	0.0 % (0)	1.2 % (3)	0.80 % (2)	0,6 % (5)
Skin/soft tissue	18 % (45)	16 % (40)	18 % (44)	17 % (129)
Unknown	12 % (30)	28 % (70)	9.2 % (23)	16 % (123)
Average length of stay and mortality				
Average length of stay; hours	116	82	115	104
Died within 30 days: n (%)	25 (10 %)	24 (9.6 %)	13 (5.2 %)	62 (8.3 %)

[a] CHHHS = Cairns and Hinterland Hospital and Health Service, SCUH = Sunshine Coast University Hospital, TPCH = The Prince Charles Hospital.

3.1. External validation of mortality prediction models

Of the four models to be applied early after admission, CSM-4 had the highest AUROC (Table 3) and was statistically better than qSOFA (Table 4). ANZROD 24 was performing best of the four models to be applied 24 h after admission (Table 3) where ANZROD 24 was statistically better than SOFA (Table 4).

4. Discussion

This external validation of eight different mortality prediction models identified the CSM-4 as a slightly better model to use early after admission (4 h after admission), while ANZROD 24 was the best performing models to be applied 24 h after admission.

Table 2
Patients condition during the first 24 h in ICU.

	CHHHS[a] N = 250	SCUH[a] N = 250	TPCH[a] N = 250	All hospitals N = 750
History of renal disease at admission	25 (10 %)	24 (9.6 %)	13 (5.2 %)	62 (8.3 %)
Vasopressors at admission[b] 0	65 (26 %)	81 (32 %)	74 (30 %)	220 (29 %)
1	136 (54 %)	144 (58 %)	110 (44 %)	390 (52 %)
2	31 (12 %)	20 (8.0 %)	44 (18 %)	95 (13 %)
3	17 (6.8 %)	5 (2.0 %)	21 (8.4 %)	43 (5.7 %)
4	1 (0.40 %)	0 (0.0 %)	1 (0.40 %)	2 (0.27 %)
Glasgow Coma Scale at admission				
Mild (13–15 scores)	212 (85 %)	214 (86 %)	200 (80 %)	626 (83 %)
Moderate (9–12 scores)	5 (2.0 %)	6 (2.4 %)	12 (4.8 %)	23 3.1 %)
Severe (3–8 scores)	33 (13 %)	30 (12 %)	38 (15 %)	101 (13 %)
Median values at admission (n)				
Lactate (mmol/L)	1.8 (250)	1.4 (245)	1.6 (241)	1.6 (736)
Bicarbonate (mmol/L)	20 (244)	20 (247)	20 (239)	20 (730)
AST (U/L)	50 (244)	39 (214)	43 (161)	44 (619)
Lactate dehydrogenase ((U/L)	280 (244)	310 (209)	340 (160)	300 (613)
S-Albumine (g/L)	24 (244)	27 (214)	27 (155)	26 (613)
S-Magnesium (mmol/L)	0.74 (244)	0.72 (213)	0.81 (161)	0.76 (618)
Median worst values (n) during the first 24 h after admission				
Worst prothrombine time (s)	15 (220)	16 (169)	16 (137)	16 (526)
Worst bilirubine (umol/L)	12 (250)	20 (248)	17 (239)	17 (737)
Worst mean arterial pressure (mmHg)	62 (250)	62 (250)	60 (250)	61 (750)

[a] CHHHS = Cairns and Hinterland Hospital and Health Service, SCUH = Sunshine Coast University Hospital, TPCH = The Prince Charles Hospital.
[b] Number of vasopressors at ICU admission. Vasopressors are any of Dopamine, Epinephrine, Norepinephrine, Phenylephrine or Vasopressin.

Table 3
External validation of models to predict sepsis mortality in intensive care units.

	Prediction models				Internal validation (in original study)			External validation (in this study)		
Time[a]	Label[b]	Sepsis specific	Number of predictors	Published (Year)	AUROC (95 % CI), p-value[c]	Nagelkirke R-square[d]	*P*-value for HLGOF[e]	AUROC (95 % CI), p-value[c]	Brier score[f]	Data available[g]
0	**qSOFA**	x	3	2016	0.61 (CI not stated), p =?	–	–	0.67 (0.60–0.73); $p < 0.0001$	–	97 %
0	**MPMo-III**		16	2007	0.82 (CI not stated), p =?	–	0.31	0.74 (0.69–0.79); p = 0.00014	0.18	100 %
1	**SAPS-III**		20	2005	0.85 (CI not stated), p =?	–	0.39	0.73 (0.64–0.82); p < 0.0001	0.16	47 %
4	**CSM-4**	x	10	2020	0.90 (0.84–0.95), p < 0.0001	0.51	0.081	0.80 (0.73–0.86); p < 0.0001	0.084	80 %
24	**ANZROD 24**		13	2013	0.91 CI not stated), p =?	–	–	0.83 (0.79–0.88); p < 0.0001	0.096	99 %
24	**APACHE-II**		17	1985	0.86 (CI not stated), p =?	0.319	–	0.76 (0.71–0.81); p < 0.0001	0.13	97 %
24	**APACHE-IIIj**		18	1991	0.89 (CI not stated), p =?	–	–	0.79 (0.74–0.84); p < 0.0001	0.12	100 %
24	**SOFA**	x	8	1996	0.88 (CI not stated), p =?	–	0.80	0.70 (0.63–0.76); p < 0.0001	0.15	96 %

[a] Time in hours from admission to the intensive care unit (ICU) to when the prediction model is recommended to be applied as stated in the original publication. 0 means the model is applicable at admission to ICU.
[b] qSOFA = quick Sequential Organ Failure Assessment; MPMo-III = Mortality Probability Model at admission; SAPS-III = Simplified Acute Physiology Score; CSM-4 = 4-Hour Cairns Sepsis Model; ANZROD 24 = Australian and New Zealand Risk of Death, APACHE-II and APACHE-IIIj = Different versions of Acute Physiology and Chronic Health Evaluation; SOFA = sequential organ failure assessment.
[c] Area under curve, with 95 % confidence interval, is an estimation of discrimination. The *p*-value for this AUC explores if it differs from random prediction (? denotes that the *p*-value for AUC was not provided in the original study).
[d] Nagelkirke R^2. Is an estimate obtained during logistic regression as a measure of overall model fit and varies between 0.0 and 1.0. The higher the better.
[e] Hosmer–Lemeshow goodness of fit (HLGOF) is an estimation of calibration. HLGOF investigates if the model describes the observations better than pure chance. A low p-value indicates that it is not better than pure chance while a high p-value says your model has good calibration. This test is too sensitive for large data sets.
[f] Brier score is an estimation of calibration. The lower the better. qSOFA was not suitable to be transformed to a probability for death.
[g] This column states the proportion of patients where all predictors were available. This study did not order extra tests or investigations for research purposes. It was an important aim of this study to investigate to what extent required variables for each prognostic model was retrieved in routine care.

Table 4
Pairwise comparison of different sepsis prediction models [a].

	p-value[b]
Models to be applied early after ICU admission (0–4 h)	
qSOFA < MPMo-III	0.068
qSOFA < SAPS-III	0.23
qSOFA < CSM-4	**0.0031**
MPMo-III > SAPS-III	0.83
MPMo-III < CSM-4	0.14
SAPS-III < CSM-4	0.21
Models to be applied later after ICU admission (24 h)	
ANZROD 24 > APACHE-II	0.042
ANZROD 24 > APACHE-IIIj	0.21
ANZROD 24 > SOFA	**0.0007**
APACHE-II < APACHE-IIIj	0.40
APACHE-II > SOFA	0.13
APACHE-IIIj > SOFA	0.021

[a] Pairwise comparison of AUC with the method described by Hanley & McNeil [26,27].
[b] We strived for an overall level of significance of 0.05. Adjusting for multiple testing according to Bonferronis method shows that the level of significance for each pairwise comparison should be set to 0.0042 and any p-value below this is highlighted in bold.

4.1. Variables relevant to include in a model

Sepsis is heterogenous, making risk stratification for prognostication difficult. Sepsis can be staged by looking at predisposition, infection, response and organ dysfunction variables, each of which are independently associated with hospital mortality [29]. Mortality prediction models that incorporated variables from each of the sepsis PIRO staging categories [29] performed better than models that took a narrow view on organ dysfunction only, the SOFA score, for example. Furthermore, some models are more difficult to use with the APACHE IIIj requiring 18 variables to be obtained. Others such as the SOFA score may have fewer variables, but they require invasive measures to be obtained. For example, an arterial line for a mean arterial pressure, or arterial blood gas for the partial pressure of oxygen. These are barriers to mortality prediction model utilisation as they may impede clinical workflow [30].

4.2. Sepsis specific models

Some mortality prediction models are sepsis specific, developed specifically for septic patients, whilst other are general mortality prediction models for use with any condition (Table 3). Among models to be applied early after admission, the sepsis-specific CSM-4 model outperformed other more general mortality prediction models as well as the sepsis-specific qSOFA. The sepsis-specific SOFA model was inferior to other more general mortality prediction models when applied 24 h after admission. The use of a sepsis specific model is, logically, more likely to be advantageous in the early phase. General mortality prediction models also include variables that are unlikely to contribute to sepsis mortality prediction, potassium levels for example in the SAPS-3, whereas lactate and number of vasopressors required in the CSM-4 are much more sepsis-specific.

4.3. Strengths and weaknesses

The major strength of this study is that multiple sites were used with varied geography and population demographics to improve generalisability. Cairns and Hinterland Hospital and Health Service was used for constructing the CSM-4 model and now also being one of three hospitals participating in this external validation. It should be clarified that the cohort at Cairns and Hinterland Hospital and Health Service used to

initially construct the CSM-4 model was from 2014 to 2018 and not the same cohort as the one now used from 2021 to 2022 in this study for the external validation. Hence, we deem it appropriate to include all three hospitals. However, we did a sensitivity analysis excluding all cases from the Cairns and Hinterland Hospital and Health Service showing minor changes to AUROC (Supplemental file 3, Table S1 and S2).

The qSOFA and SOFA models were chosen since they are recommended for use in the latest sepsis-3 guidelines. The SAPS-III and MPMo-III are globally well recognised and frequently used in research. The APACHE-IIIj, the ANZROD and CSM-4 are used in Australia where this study was done, hence, they were included.

All patients included fulfilled the Sepsis-3 criteria and were discharged with a final diagnosis of sepsis. We cannot exclude that there were patients initially fulfilling the Sepsis-3 criteria but for reasons unknown to the authors were not diagnosed with a final diagnosis of sepsis since these patients were never presented to us by the ANZICS CORE database.

Not all variables were available for all models. However, this reflects current standard practice where many variables included in the SAPS-III model were often not routinely obtained. This is important information and may explain why such models have little clinical impact. This also introduces uncertainty about our validation of the SAPS-III model.

4.4. Conclusions and future directions

Time after admission may decide which prediction model is most useful. Early after admission to the ICU, the sepsis-specific CSM-4 mortality prediction model were slightly better than other models recommended to be used early after admission. However, at 24 h, the general ANZROD 24 model was the better one. As characteristics, treatment and outcomes of sepsis evolve, so too should mortality prediction models.

CRediT authorship contribution statement

Satyen Hargovan: Writing – original draft, Methodology, Conceptualization, Writing – review & editing, Project administration, Data curation. **Charlotte Simpson:** Writing – review & editing, Conceptualization, Data curation. **Sayonne Sivalingam:** Writing – review & editing, Conceptualization, Data curation. **Angus Carter:** Methodology, Writing – review & editing, Conceptualization. **Ronny Gunnarsson:** Writing – review & editing, Methodology, Conceptualization, Writing – original draft, Formal analysis.

Funding

No specific funding for this study.

Declaration of competing interest

None.

Acknowledgements

No acknowledgement.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jcrc.2025.155174.

Data availability statement

Raw data cannot be shared due to restrictions imposed by the participating hospitals.

References

[1] Shankar-Hari M, Phillips GS, Levy ML, Seymour CW, Liu VX, Deutschman CS, et al. Developing a new definition and assessing new clinical criteria for septic shock: for the third international consensus definitions for Sepsis and septic shock (Sepsis-3). JAMA 2016;315(8):775–87.
[2] Fleischmann C, Scherag A, Adhikari NK, Hartog CS, Tsaganos T, Schlattmann P, et al. Assessment of global incidence and mortality of hospital-treated Sepsis. Current estimates and limitations. Am J Respir Crit Care Med 2016;193(3):259–72.
[3] Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kievlan DR, et al. Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the global burden of disease study. Lancet 2020;395(10219):200–11.
[4] Vincent JL. The clinical challenge of Sepsis identification and monitoring. PLoS Med 2016;13(5):e1002022.
[5] Evans L, Rhodes A, Alhazzani W, Antonelli M, Coopersmith CM, French C, et al. Surviving Sepsis campaign: international guidelines for Management of Sepsis and Septic Shock 2021. Crit Care Med 2021;49(11). e1063-e143.
[6] Hargovan S, Groch T, Brooks J, Sivalingam S, Bond T, Carter A. Indigenous Australians critically ill with sepsis: characteristics, outcomes, and areas for improvement. Aust Crit Care 2024;37(4):548–57.
[7] Petersen E, Zumla A. To have sepsis or to be septic-is the difference between these clinical conditions important? Int J Infect Dis 2016;48:118–9.
[8] Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The third international consensus definitions for Sepsis and septic shock (Sepsis-3). JAMA 2016;315(8):801–10.
[9] Lengquist M, Lundberg OHM, Spangfors M, Annborn M, Levin H, Friberg H, et al. Sepsis is underreported in Swedish intensive care units: a retrospective observational multicentre study. Acta Anaesthesiol Scand 2020;64(8):1167–76.
[10] Lengquist M, Friberg H, Frigyesi A. Sepsis is underreported in ICUs: a conclusion that still holds. Acta Anaesthesiol Scand 2021;65(2):277.
[11] Loots FJ, Smits M, Hopstaken RM, Jenniskens K, Schroeten FH, van den Bruel A, et al. New clinical prediction model for early recognition of sepsis in adult primary care patients: a prospective diagnostic cohort study of development and external validation. Br J Gen Pract 2022;72(719). e437-e45.
[12] Lee J, Dubin JA, Maslove DM. Mortality prediction in the ICU. Secondary analysis of electronic health records. Cham (CH) 2016:315–24.
[13] Sadaka F, EthmaneAbouElMaali C, Cytron MA, Fowler K, Javaux VM, O'Brien J. Predicting mortality of patients with Sepsis: a comparison of APACHE II and APACHE III scoring systems. J Clin Med Res 2017;9(11):907–10.
[14] Ramspek CL, Jager KJ, Dekker FW, Zoccali C, van Diepen M. External validation of prognostic models: what, why, how, when and where? Clin Kidney J 2021;14(1):49–58.
[15] Siontis GC, Tzoulaki I, Ioannidis JP. Predicting death: an empirical evaluation of predictive tools for mortality. Arch Intern Med 2011;171(19):1721–6.
[16] Cox EGM, Meijs DAM, Wynants L, Sels JEM, Koeze J, Keus F, et al. The definition of predictor and outcome variables in mortality prediction models: a scoping review and quality of reporting study. J Clin Epidemiol 2025;178:111605.
[17] Cox EGM, Wiersema R, Eck RJ, Kaufmann T, Granholm A, Vaara ST, et al. External validation of mortality prediction models for critical illness reveals preserved discrimination but poor calibration. Crit Care Med 2023;51(1):80–90.
[18] 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.
[19] Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, et al. SAPS 3–From evaluation of the patient to evaluation of the intensive care unit. Part 2: development of a prognostic model for hospital mortality at ICU admission. Intensive Care Med 2005;31(10):1345–55.
[20] Metnitz PG, Moreno RP, Almeida E, Jordan B, Bauer P, Campos RA, et al. SAPS 3–From evaluation of the patient to evaluation of the intensive care unit. Part 1: objectives, methods and cohort description. Intensive Care Med 2005;31(10):1336–44.
[21] Hargovan S, Gunnarsson R, Carter A, De Costa A, Brooks J, Groch T, et al. The 4-hour Cairns Sepsis model: a novel approach to predicting sepsis mortality at intensive care unit admission. Aust Crit Care 2021;34(6):552–60.
[22] Paul E, Bailey M, Pilcher D. Risk prediction of hospital mortality for adult patients admitted to Australian and New Zealand intensive care units: development and validation of the Australian and New Zealand risk of death model. J Crit Care 2013;28(6):935–41.
[23] Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985;13(10):818–29.
[24] 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(6):1619–36.
[25] Vincent JL, Moreno R, Takala J, Willatts S, De Mendonca A, Bruining H, et al. The SOFA (Sepsis-related organ failure assessment) score to describe organ dysfunction/failure. On behalf of the working group on Sepsis-related problems of the European Society of Intensive Care Medicine. Intensive Care Med 1996;22(7):707–10.
[26] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143(1):29–36.
[27] Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148(3):839–43.

[28] Comparison of AUC of independent ROC curves. https://www.medcalc.org/calc/comparison_of_independentROCtest.php;; 2024 [accessed 23rd of August.2024].
[29] Rathour S, Kumar S, Hadda V, Bhalla A, Sharma N, Varma S. PIRO concept: staging of sepsis. J Postgrad Med 2015;61(4):235–42.
[30] Gupta T, Puskarich MA, DeVos E, Javed A, Smotherman C, Sterling SA, et al. Sequential organ failure assessment component score prediction of in-hospital mortality from Sepsis. J Intensive Care Med 2020;35(8):810–7.