Aus der Klinik für Gastroenterologie, Hepatologie und Infektiologie
der Heinrich-Heine-Universität Düsseldorf
Direktor: Univ.-Prof. Dr. Tom Lüdde

Evaluation of Clinical Scores for Early Detection and Prognostic Assessment of Severe Infections in a Resource-Limited Setting

Dissertation

zur Erlangung des Grades eines Doktors der Medizin
der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Birte Elisabeth Friederike Scholz
2025

Als Inauguraldissertation gedruckt mit Genehmigung der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf

gez.:
Dekan: Prof. Dr. med. Nikolaj Klöckner
Erstgutachter: Prof. Dr. med. Torsten Feldt
Zweitgutachter: Prof. Dr. med. Klaus Pfeffer

In Liebe und Dankbarkeit für meine Familie.

Planned publication:

Improvements in sepsis recognition and prediction of mortality for allocation of treatment resources in the resource-limited setting

Fuchs, A.*, Scholz, B.*, Tufa, T., Eggers, Y., Abdissa, S., Schneider, M., Hurissa, Z., Jarso, G., Orth, H., Pfeffer, K., Mackenzie, C., Luedde, T., Kaasch, A., Feldt, T.

*These authors contributed equally to this work.

Note: Part of the data presented in this thesis are part of a manuscript currently being prepared for publication. The title and author list are provisional.

Zusammenfassung

Sepsis zählt weltweit zu den häufigsten Todesursachen[1]. Eine zügige Diagnose ist entscheidend[2–4]. Zahlreiche klinische Scores zur Früherkennung wurden in *High Income*-Ländern entwickelt und kaum unter ressourcen-limitierten Bedingungen evaluiert[5–11]. Das Ziel dieser Arbeit ist daher die Evaluierung und Anpassung von Scores zur Risikoevaluation bei Sepsis in einem ressourcen-limitierten Umfeld in Äthiopien.

Patienten, die sich mit Anzeichen einer Infektion in den Notaufnahmen zweier äthiopischer Krankenhäuser vorstellten, wurden bei Einwilligung in die Studie eingeschlossen. Es erfolgte die Erhebung klinischer Parameter und Nachverfolgung bis Entlassung. Folgende Scores wurden evaluiert, mittels *Decision-Tree*-Analysen und Regressionsmodellen, nach Methoden von Moore et al.[12] an die Gegebenheiten vor Ort angepasst, sowie anschließend an einem Testdatensatz reevaluiert: *Systemic Inflammatory Response Syndrome* (SIRS)[5], *Modified Early Warning Score* (MEWS)[6], modifizierter *National Early Warning Score* (mNEWS)[7], *Quick Sequential Organ Failure Assessment* (qSOFA)[8] und *Universal Vital Assessment* (UVA)[12].

Bei den 851 Studienteilnehmenden lag die Krankenhaussterblichkeit bei 4,5% (n = 38). Die Flächen unter der *Receiver Operating Characteristic* Kurve (AUROC) (mit 95%-Konfidenz-Intervall) zur Vorhersage der Krankenhaus Mortalität betrugen: SIRS 0,63 (0,54–0,71), MEWS 0.70 (0.61–0.79), mNEWS 0,72 (0,63–0,81), qSOFA 0,67 (0,59–0,74) und UVA 0,87 (0.82–0.92). Im Testdatensatz zeigten die originalen und die adaptierten Scores folgende AUROCs: SIRS 0,64 (0,47–0,80), adaptierter SIRS 0,77 (0.67–0,86), qSOFA 0,73 (0,64–0,83), adaptierter qSOFA 0,82 (0,72-0,92); mNEWS 0,76 (0,59–0,92), adaptierter mNEWS 0,82 (0,72–0,92); UVA 0.86 (0,76–0,94) und adaptierter UVA 0,86 (0,76–0,94).

Die Analyse zeigte, dass der UVA Score den anderen getesteten Scores in der Mortalitätsvorhersage überlegen war. Dies dürfte darauf zurückzuführen sein, dass er an einer Kohorte aus Sub-Sahara Afrika entwickelt wurde[12]. Die Leistung von Scores kann in verschiedenen Populationen variieren und durch Anpassung an lokale Gegebenheiten verbessert werden[13]. Weitere Validierungen in größeren Kohorten sind erforderlich, um unsere angepassten Modelle zu validieren und verfeinern.

Summary

Sepsis is one of the most common causes of death worldwide[1]. Rapid diagnosis is crucial, but often a challenge[2–4]. Numerous clinical scores for early detection have been developed in high-income countries (HICs) and have hardly been evaluated under resource-limited conditions[5–11]. The aim of this work is to evaluate and adapt scores for sepsis risk evaluation in a resource-limited setting in Ethiopia.

Patients presenting with signs of infection to the emergency departments (EOPDs) of two Ethiopian hospitals were included in the study after providing informed consent. Clinical parameters were collected and followed up until discharge. The following established scores were evaluated, adapted to local conditions using decision tree analyses and regression models, as proposed by Moore et al.[12], and subsequently validated using a test dataset: Systemic Inflammatory Response Syndrome (SIRS)[5], Modified Early Warning Score (MEWS)[6], a modified National Early Warning Score (mNEWS)[7], Quick Sequential Organ Failure Assessment (qSOFA)[8], and Universal Vital Assessment (UVA)[12].

Among the 851 study participants, the hospital mortality rate was 4.5% (n = 38). The areas under the receiver operating characteristic curves (AUROC) (with 95% confidence intervals) for in-hospital mortality prediction were: SIRS 0.63 (0.54, 0.71), MEWS 0.70 (0.61, 0.79), mNEWS 0.72 (0.63, 0.81), qSOFA 0.67 (0.59, 0.74), and UVA 0.87 (0.82, 0.92). In the test dataset the original and the adapted scores showed the following AUROCs: SIRS 0.64 (0.47, 0.80), adapted SIRS 0.77 (0.67, 0.86), qSOFA 0.73 (0.64, 0.83), adapted qSOFA 0.82 (0.72, 0.92); mNEWS 0.76 (0.59, 0.92), adapted mNEWS 0.82 (0.72, 0.92); UVA 0.86 (0.76, 0.94), and adapted UVA 0.86 (0.76, 0.94).

In this study, UVA outperformed the other evaluated scores in predicting mortality among patients presenting at the EOPD with infections in a resource-limited setting. This may be due to the development of UVA in a cohort from Sub-Saharan Africa[12]. The performance of scores can vary in different populations[13]. Adapting scores to local conditions can improve the accuracy of mortality prediction. However, further validations in larger cohorts are necessary to refine and confirm our adapted models.

Table of contents

List of figures

List of tables

List of abbreviations

a.s.l.	Above sea level
AHMC	Adama Hospital Medical College
ATRH	Asella Teaching and Referral Hospital
AUROC	Area Under the Receiver Operating Characteristic Curve
AVPU	Alert Verbal Pain Unresponsive scale
CI	Confidence interval
EOPD	Emergency Outpatient Department
FiO_2	Fraction of inspired oxygen
GCS	Glasgow Coma Scale
Hgb	Hemoglobin
HIC	High-income country
HIV	Human immunodeficiency virus
ICU	Intensive care unit
IQR	interquartile range
KI	Konfidenz Intervall
LMIC	Low- and middle-income country
LODS	Logistic Organ Dysfunction System
MAP	Mean arterial pressure
MCV	Mean Corpuscular Volume
MEWS	Modified Early Warning Score
mNEWS	Modified National Early Warning Score
NEWS	National Early Warning Score
NPV	Negative Predictive Value
OR	Odds ratio
PaO_2	Partial pressure of oxygen in arterial blood
PIRO	Sepsis 2 Model: **P**redisposing conditions, detectability of **I**nfection, the host **R**esponse, and **O**rgan dysfunction
PLT	Platelets
PPV	Positive Predictive Value
qSOFA	Quick Sequential Organ Failure Assessment
RBC	Red Blood Cells
SIRS	Systemic Inflammatory Response Syndrome
SMOTENC	Synthetic Minority Over-sampling Technique for categorical and continuous variables
SOFA	Sequential Organ Failure Assessment
SSC	Surviving Sepsis Campaign
UVA	Universal Vital Assessment
WBC	White Blood Cells
WHO	World Health Organization

1 Introduction

1.1 Global and regional incidence and mortality of sepsis

Despite many years of sepsis research and implementations of new treatment strategies, sepsis remains a global challenge. Incidence estimates vary from 31.5 million to 48.9 million annual sepsis cases and 19.4 million severe sepsis cases worldwide[1,14]. These wide discrepancies in global incidence estimates emerge from differences in sepsis definitions and lack of data from low- and middle-income countries (LMICs), where 87% of the world's population lives and where the burden of sepsis is estimated to be greatest[1,14–16]. In line with this, in Sub-Saharan Africa alone, there are an estimated 16.7 million sepsis cases each year, including 1.1 million cases in Ethiopia[1]. A systematic review conducted by Fleischmann-Struzek et al.[17] found an increasing trend in number sepsis cases.

Moreover, not only is the incidence of sepsis high, but also its associated mortality. In-hospital mortality rate estimates range from 14.7% to 29.9% in the USA to as high as 49.0% and above in low-income countries[18,19]. 11 million people were estimated to die of sepsis globally per year and in Sub-Saharan Africa, the in-hospital mortality for sepsis is estimated to be particularly high, with 3.5 million deaths annually, including 270.000 deaths in Ethiopia[1].

Despite these high mortality figures, the burden of sepsis may still be underestimated in Sub-Saharan Africa due to limited data availability, particularly on sepsis patients' long-term follow-up data (longer than 28 days)[20]. Additionally, sepsis is often not diagnosed or registered as the reason for a patient's death[21].

1.2 The financial burden of sepsis

Although data on the financial burden of sepsis is available almost exclusively from HICs, estimates suggest that the financial burden of sepsis is enormous in any context[22]. According to the Agency for Healthcare Research and Quality, the United States spent more than 38 billion USD on the care of septic patients in 2017 and is thus ranked the most expensive condition treated in the United States[23].

In addition to these direct costs that sepsis treatment cuts into a country's health care budget, indirect costs arise due to the economic loss of previously healthy working-age people[22]. These indirect costs are particularly grave in LMICs, where sepsis tends to affect younger people compared to HICs[13,24,25]. Despite its great financial and public health burden, sepsis has long been and possibly still is an underrated and underrecognized condition.

1.3 Addressing the burden of sepsis in Ethiopia

Combating the high burden of sepsis in the developing world requires a broad approach, which will only be achievable through the collaboration of national and international stakeholders. Several key aspects must be addressed to improve sepsis survival. Among others, these include training of health care workers, improving early recognition and timely treatment strategies, investment in intensive care units (ICUs) as well as laboratory resources, improving supply chains for diagnostic and treatment material, etc[26–29].

While in some contexts large investments in ICU infrastructure or laboratory equipment may not be feasible in the near future, addressing timely treatment and triage-based distribution of critical care resources is likely to be more feasible in many contexts and may improve sepsis outcomes[12].

Papali et al.[30] have developed a model to explain the delay in sepsis treatment in adults in resource-limited settings. According to the authors, there are three main issues to address when preventing delay in sepsis treatment:

1. Sepsis recognition and diagnosis at the time of triage
2. Initial focused resuscitation
3. Postresuscitation clinical monitoring and reassessment.

This thesis focuses on Papali's first reason for delay: sepsis recognition and diagnosis.

1.4 The benefit of early recognition and treatment

Mortality associated with sepsis can be reduced when the condition is recognized promptly and appropriate measures are taken[2–4], most importantly the early

administration of appropriate antibiotics[31,32]. Liu et al.[32] found that the absolute mortality rate for patients with severe sepsis increased by 0.3% for every hour delay in antibiotic treatment. For septic shock there was an increase of 1.8% per hour.

Furthermore, the identification of patients at risk of a poor outcome enables clinicians to allocate limited resources to those who would benefit the most[26,33]. A fast and precise recognition can also be more cost-effective by enabling early treatment, minimizing the progression of sepsis to severe cases, preventing complications, and often leading to shorter hospital stays[34].

1.5 The challenge of defining sepsis

Thus, a prompt, sensitive, and specific identification of patients at risk is essential. Unfortunately, diagnosing and defining sepsis has been a huge challenge so far. Many ways to describe sepsis have been evaluated, but none of them seem to grasp the whole concept. But what makes it so difficult to find a definition for sepsis?

Since no single highly sensitive biomarker to diagnose sepsis has been established, its diagnosis relies on a combination of clinical, laboratory, radiological, and physiological indicators[35]. Furthermore, the construct of sepsis describes a heterogenous condition. Depending on the causing pathogen, different complex inflammatory pathways are activated[16,21]. This diversity complicates the finding of an accurate definition for sepsis.

Generally, understanding sepsis has been a challenge due to several challenges in sepsis research. Results of sepsis studies are often hard to compare as they use different sepsis definitions, endpoints (i.e. ICU admission, in-hospital mortality, 28-day mortality, etc.), or are conducted in specific settings (ICU population, patients presenting at EOPD, preclinical setting)[36,37]. Retrospective sepsis studies can also be challenging as there is often no specific code for sepsis in hospitals' coding systems[35].

1.6 Sepsis definitions

Despite these challenges or maybe because of them, various sepsis definitions have been developed. In fact, sepsis has been studied throughout history. The first mention of the

word sepsis dates back over 2700 years as it has been found in Homer's poems. It derives from the Greek word "sepo", which is translated as "I rot"[38].

Stiefel et al.[39] describes that older reports of the concept of sepsis, being a result of infections, go back to a papyrus, found in Egypt, written in 1600 BC, which is believed to be a copy of an even older writing from 3000 BC. The manuscript describes the examination, symptoms, treatment, and prognosis of 48 surgical cases and mentions septic processes due to wound infections. Since, the term and concept emerged in literature and were used by famous authors like Aristotle, Matthaeus Silvaticus, and Hippocrates[40].

For a long time in history, there was no need for a precise definition of sepsis because of the lack of specific therapeutic options. With greater recognition of high case numbers of sepsis, development of more advanced therapy strategies, and a deeper understanding of physiological processes that lead to sepsis, the need for a more accurate definition for clinical, as well as research purposes grew[9].

1.6.1 Sepsis-1

This necessity was recognised in 1991 by members of the American College of Chest Physicians/ Society of Critical Care Medicine Consensus Conference. Their aim was to define sepsis and its sequelae in a precise way. Therefore, the authors agreed to differentiate between sepsis, severe sepsis, and septic shock and introduced the definitions that were later referred to as sepsis-1 definitions[5]. Like in earlier definitions, sepsis was defined as the host's "systemic inflammatory response to infection"[5,41]. The occurrence of sepsis in conjunction with organ dysfunction, hypotension or hypoperfusion was called severe sepsis, whereas remaining hypotension despite adequate volume management was referred to as septic shock, with hypotension being defined as systolic blood pressure <90 mmHg or reduction of baseline systolic blood pressure of ≥40 mmHg when no other causes for hypotension are identifiable[5].

To simplify the clinical diagnosis of sepsis the Systematic Inflammatory Response Syndrome (SIRS) criteria were introduced (see Table 1)[5].

SIRS is not limited to infectious diseases but can be associated with various clinical conditions, such as, e.g., pancreatitis, burns, or trauma. Only if a patient shows two or more SIRS criteria due to an infection, their condition can be termed sepsis according to the sepsis-1 definition[5,42].

Table 1: Systemic Inflammatory Response Syndrome (SIRS) Criteria by Bone et al.[5]

Parameter	Points
Temperature <36.0°C or >38.0°C	1
Heart rate ≥90 beats/minute	1
Respiratory Rate ≥20 breaths/minute	1
White blood cell count <4,000 or >12,000 or >10% bands	1

Table 1: Calculation of the Systemic Inflammatory Response Syndrome (SIRS) Criteria by Bone et al.[5] to identify patients with SIRS. The score is based on temperature, heart rate, respiratory rate, and white blood cell count. Score points are assigned to patients with suspected infection whose parameters deviate from the normal value range. Higher scores indicate a greater risk of SIRS.

According to Balk et al.[42] the SIRS criteria were easily established in clinical practice, because of their simplicity. Especially, in research, the introduction of the SIRS criteria was advantageous, as it led to similar inclusion criteria for clinical trials and thus to improved comparability of clinical study results. Balk et al.[42] further noted, that the standardization of language also led to clearer medical communication in daily life, at conferences, and in literature. Furthermore, SIRS proved to be practical, as it does not require additional information than that obtained regularly during clinical practice.

Nevertheless, opinions on the SIRS criteria are divided because some problems arise with the use of these criteria. First, as mentioned above, the SIRS criteria are not specific for sepsis but can occur in patients with various conditions[43,44]. In a study conducted by Sprung et al.[45], 93% of patients admitted to the ICU showed two or more SIRS criteria, demonstrating the low specificity of the SIRS criteria. If almost every ICU patient fulfils the SIRS criteria and sepsis is defined as the presence of at least two SIRS criteria and infection, then, due to this definition in the ICU, the terms sepsis and infection can almost be used interchangeably[44]. Furthermore, this definition does not distinguish between valuable facets of the host's response to infection and harmful ones, which lead to complications[44]. Even though the SIRS criteria are not a good tool to identify patients with sepsis, they remain suitable for infection recognition, as fever, increased white

blood cell count, and increased heart rate reflect an appropriate response of the body to infection[8].

1.6.2 Sepsis-2

The concerns mentioned above and an increase in clinical trials led to the need for a modification of the existing definitions. Therefore, in 2001, several European and American intensive care associations held a "Consensus Conference" to evaluate and improve the current definitions, as well as to find new tools for a more accurate diagnosis of sepsis as reported by Levy et al.[46]. Unfortunately, the delegates were not able to propose a more accurate definition of sepsis, due to the lack of data to support an adjustment. Thus, the 1991 criteria remained in use[5,44].

To the already existing definitions[5], the conference delegates added a long list of symptoms and laboratory parameters (see Suppl. Table 1) for the diagnosis of sepsis and proposed a sepsis staging model[46], inspired by the tumour-node-metastasis classification used in oncology[47]. The idea was to document disease extent, to identify patients at risk of a poor outcome, and divide patients into subgroups for whom the same therapeutic intervention is likely to be beneficial. The model was called the PIRO, as it focuses on **P**redisposing conditions, detectability of **I**nfection, the host **R**esponse, and **O**rgan dysfunction to stratify a patient's risk for a poor outcome. The conference participants envisioned the new model's advantage to be the discrimination between morbidity caused by infection and morbidity caused by the host's response to infection[46].

Until now, the PIRO system is not a fully developed staging system, but it is rather a template for further investigation[48]. Different aspects of sepsis can be investigated with the help of the model, but rapid recognition of patients likely to have a poor outcome is not possible.

1.6.3 Sepsis-3

Again, in 2016, a task force was appointed by the European Society of Intensive Care Medicine and the Society of Critical Care Medicine to review previous definitions in consideration of new pathobiological, immunological, and epidemiological understandings, based on the findings of at that time newly available large electronic

databases[8]. Thus, these are the first definitions that are not solely based on expert opinion but on the assessment of large electronic patient databases. Unlike other diseases, sepsis had so far been mostly defined by its clinical presentation (i.e., the SIRS criteria[5]).

At this point, the task force emphasised the need to differentiate between a definition that explains what the essence of the condition is and the description of clinical criteria to identify a patient who suffers from this condition. Furthermore, the view that sepsis is a disease that consists of various stages of the disease was abandoned[8].

New definition for sepsis

According to the Sepsis-3 definitions, sepsis is defined as a "life-threatening organ dysfunction caused by a dysregulated host response to infection", pointing out the urgency and possible deadliness of the condition[8,49]. In contrast to earlier definitions that focused on inflammation, this update emphasizes organ dysfunction[5,46]. If the body responds to an infection reasonably, the infection can be warded off. It is the inadequate host response that leads to organ dysfunction, which is now defined as sepsis. This dysregulated response of the body includes the activation of pro- and anti-inflammatory pathways[43,46].

What was earlier defined as sepsis (i.e., infection and presence of more than one SIRS criteria) is now defined as infection. And the old term "severe sepsis" that used to refer to sepsis with organ dysfunction according to the old definition, becomes obsolete, because sepsis is defined likewise[8].

Clinical criteria for the identification of patients with sepsis

The task force also recognised the need for clinical criteria to identify patients with sepsis in the intensive care unit (ICU), in emergency departments, and non-clinical settings. They found that in patients, admitted to the ICU, the performance of the Sepsis-related Organ Failure Assessment (SOFA) in predicting in-hospital mortality was significantly greater than SIRS and not inferior to more complicated scores[50].

The SOFA is a score which was developed by Vincent et al.[51] in 1996 to describe organ dysfunction. It focuses on respiration, circulation, liver function, kidney function,

coagulation, and consciousness. Each parameter is scored from 0 to 4 points based on severity of dysfunction, resulting in a maximum score of 24. Table 2 shows how the SOFA score is calculated.

An acute increase in the SOFA score of ≥2 points in consequence of an infection indicates organ dysfunction and is associated with 10% excess in-hospital mortality[8].

Table 2: Sepsis-related Organ Failure Assessment Score by Vincent et al.[51]

Parameter	0	1	2	3	4
PaO_2/FiO_2, mmHg	≥400	<400	<300	<200 with respiratory support	<100 with respiratory support
Platelets $x10^3/mm^3$	≥150	<150	<100	<50	<20
Bilirubin, mg/dl (µmol/l)	<1.2 (20)	1.2-1.9 (20-32)	2.0-5.9 (33-101)	6.0-11.9 (102-204)	>12.0 (<204)
Hypotension	MAD ≥70	MAP <70	Dopamine ≤5 or dobutamine (any dose)*	Dopamine >5 or epinephrine ≤0.1 or norepinephrine ≤0.1	Dopamine >15 or epinephrine >0.1 or norepinephrine >0.1
Glasgow Coma Scale	15	13-14	10-12	6-9	<6
Creatinine, mg/dl (µmol/l) or urine output	<1.2 (110)	1.2-1.9 (110-170)	2.0-3.4 (171-299)	3.5-4.9 (300-440) or <500ml/day	>5.0 (>440) or >200 ml/day

Table 2: Calculation of the Sepsis-related Organ Failure Assessment Score (SOFA) by Vincent et al.[51] to assess organ dysfunction in critically ill patients. An acute increase of the SOFA score of ≥2 points as a result of an infection is a sign for organ dysfunction and is associated with increased mortality. *Adrenergic agents administered for at least 1h (doses given are in µg/kg/min). PaO_2= oxygen partial pressure; FiO_2= fraction of inspired oxygen.

According to Singer et al.[8], since laboratory parameters are used to calculate the SOFA score, it is not suitable as a quick screening tool. Therefore, for the use outside of the ICU, the quick Sequential (Sepsis-Related) Organ Failure Assessment (qSOFA) score was suggested by the task force as a bedside screening tool. The qSOFA is easy to use and only consists of three criteria: respiratory rate ≥22 breaths/minute, altered consciousness: Glasgow Coma Scale (GCS) <15, and systolic blood pressure ≤100 mmHg.

If a patient shows two or more of these criteria, the patient is more likely to have a poor outcome (see Table 3)[8].

Table 3: Quick Sequential (Sepsis-Related) Organ Failure Assessment by Singer et al.[8]

Parameter	Points
Respiratory rate ≥22 breaths/minute	1
Systolic blood pressure ≤100 mmHg	1
Altered mental status: Glasgow Coma Scale <15	1

Table 3: Calculation of the Quick Sequential (Sepsis-Related) Organ Failure Assessment by Singer et al.[8] to screen adult patients with suspected infection and identify those who are at risk of a poor outcome. The score is based on respiratory rate, systolic blood pressure, and mental status. Score points are assigned to patients, whose parameters deviate from the normal value range. A score ≥2 is considered positive and is associated with an increased mortality[8].

New definition for septic shock

Septic shock was defined as "a subset of sepsis in which underlying circulatory and cellular metabolism abnormalities are profound enough to substantially increase mortality", adding the relevance of cellular alterations and deformities, as well as a clearly higher probability of a poor outcome compared to sepsis according to the old definitions[8,52].

Clinical criteria for the identification of patients with septic shock

To identify criteria used to diagnose septic shock, a meta-analysis and systematic review were conducted[52]. The systematic review revealed the broad spectrum of septic shock definitions in use. Moreover, the meta-analysis showed the resulting variability in mortality rates, pointing out the need for established definitions for clinical practice, as well as research.

Based on the conducted studies, the task force agreed upon the following clinical criteria to identify septic shock among patients with sepsis: the need for vasopressors to maintain a mean arterial pressure (MAP) above 65 mmHg and a serum lactate greater than 2 mmol/L despite adequate resuscitation[8,52]. The new criteria were evaluated in large electronic health records provided by the University of Pittsburgh Medical Centre and the organization Kaiser Permanente Northern California[52]. Patients identified by the new septic shock criteria had a mortality rate over 40% higher than patients with sepsis alone[8].

Advantages and limitations of the newest definition

An advantage of the sepsis-3 definitions is the clear distinction between a definition of sepsis, which explains its pathophysiological nature, and clinical criteria to recognise patients with sepsis and septic shock, as well as the proposed screening tool, qSOFA, which facilitates the prompt recognition of at-risk patients[53].

On the other hand, the sepsis-3 definitions have been criticized because they are still very broad, including patients with various pathogens and comorbidities[9]. There is no division of patients with similar characteristics into subgroups. The division of sepsis into sepsis subtypes, which are characterized by cellular changes leading to the expression of specific biomarkers, may enable specific treatment for each subgroup, which may be beneficial in improving sepsis outcome[9,54,55].

1.7 The potential positive impact of scoring systems

In addition to the summarized sepsis definitions and clinical criteria, further scores have been developed to either identify patients with sepsis or to detect those at risk of a poor outcome. Among others, the following scores have been introduced: The National Early Warning Score (NEWS)[7], the Modified Early Warning Scores (MEWS)[6], and the Universal Vital Assessment (UVA)[12].

In HIC, the use of scores to identify patients at risk of a poor outcome has led to earlier responses, which are often crucial for survival[56,57]. However, there is little evidence regarding the impact of the implementation of these scores in Africa. One study conducted in Rwanda reported that after the implementation of a Paediatric Early Warning for Resource-Limited Settings score, there were more calls from nurses to treating physicians to start early interventions[57]. In Egypt, Badr et al.[58] were able to show that the implementation of the NEWS was associated with earlier interventions for patients with a critical illness and significant patient outcome improvement.

1.8 Limitations of clinical scoring systems in resource-limited settings

A lack of laboratory equipment and therapeutic measures may cause difficulties to use certain clinical scoring systems and sepsis definitions, in resource-limited environments[9,10]. For instance, calculating the Logistic Organ Dysfunction Score, a weighted score to assess the extent of organ dysfunction, requires laboratory measurements such as bilirubin, blood-urea-nitrogen, and prothrombin[59]. In certain settings, attaining these parameters may not be feasible. Similar difficulties may arise when using the SOFA score, which requires the availability of blood gas analysis, bilirubin concentration, platelet count, and creatinine concentration[51]. Furthermore, the clinical criteria that define septic shock according to the sepsis-3 definition require the use of vasopressors and lactate measurements, which also may not be available[8–10,52].

However, for the calculation of the tools mentioned above, SIRS, MEWS, NEWS, qSOFA, and UVA, no complex laboratory measures are required. Out of these five scores, only SIRS and UVA require laboratory results (i.e., white blood cell count and human immunodeficiency virus (HIV) status), which are likely to be assessable in most settings.

Even though the highest sepsis incidence is found in low-income countries, most diagnostic tools have been derived solely from cohorts in HICs and have rarely been validated in low-resource settings[1,9–11,15]. This raises concerns about their applicability in LMICs, where patient populations, healthcare infrastructure, and treatment options may differ substantially[9]. To ensure the clinical practicability, accuracy, and reliability of scores, context-specific validation or the development of context-specific tools is an essential step towards improving sepsis recognition on a global level[10,60,61].

In a previous study conducted by Fuchs et al.[60], at one of the hospital where the current study was conducted at, patients presenting at the hospital with signs of infection and positive SIRS criteria were included. In this cohort of 170 patients, the performance of qSOFA in identifying patients with sepsis, defined as SOFA score ≥2, was evaluated. Sensitivity of qSOFA for the diagnosis of sepsis in this SIRS positive cohort was 0.54, which was higher than those reported in literature. Specificity, however, was relatively

low at 0.66. The Area Under the Receiver Operating Characteristic Curve (AUROC) was 0.60. The poor performance of qSOFA to recognise sepsis in this study emphasizes the need for further evaluation of existing clinical scores for sepsis recognition and the detection of patients at risk of a poor outcome[60].

1.9 Aim of the study

1.9.1 Main objective

The aim of this study is to evaluate the ability of risk assessment scores to predict in-hospital mortality in patients presenting with signs of infection to the Emergency Outpatient Department (EOPD) in a resource-limited setting.

1.9.2 Secondary objectives

The secondary objectives of this study are to adapt the evaluated scores to enhance their performance in a resource-limited setting, to raise awareness of sepsis, and to raise awareness of global health care inequities.

1.9.3 Hypothesis

I. We hypothesize that risk prediction scores developed in similar contexts demonstrate higher performance compared to those derived from different settings.

II. We hypothesize that the adaptation of scores to regional settings is feasible and can improve score performance.

2 Methods

2.1 Research sites

This is a prospective cohort study conducted in a resource-limited setting in Ethiopia. Prior to the data collection, ethical approval (Institutional Ethical Review Board of Arsi University in Asella, Ethiopia: AU/H/516/120/4445/4/18; Ethics Committee of the Faculty of Medicine, Heinrich-Heine-University, in Düsseldorf, Germany: study number 5729) was obtained. Study participants were included at Asella Teaching and Referral Hospital (ATRH) and Adama Hospital Medical College (AHMC).

ATRH is a hospital located in Asella, in the Oromia Region of Ethiopia, serving about 3.5 million people in the area[62]. In 2016, over 15.000 people presented to the emergency department[62]. The AHMC is also located in Oromia, about 75 km north of Asella. It serves a population of around 6 million people[63]. Patients travel from Oromia, Amhara, Afar, Somali, and Dire-Dawa to be treated at AHMC. It has 232 beds and serves on average 1000 patients a day[63].

2.2 Inclusion criteria and primary endpoint

Between December 2018 and September 2020, patients presenting at the Emergency Department at ATRH and AHMC during the working hours of our study team were screened for prospective study inclusion.

All patients who were included in this study showed clinical signs of infections, were aged 15 years or above, were able and willing to give informed consent, as well as to adhere to the procedures of the study (in particular, allow blood samples to be taken if required).

Furthermore, patients who had antibiotics administered in the Emergency Outpatient Department (EOPD) of the respective study centre before the clinical assessment for study inclusion were not included. Antibiotics administered in a referral institution were not an exclusion criterion. A complete list of inclusion and exclusion criteria is shown in Table 4. The clinical signs of infection included the following: fever/chills, skin alterations (redness, swelling, pain), neurological (headache, stiffness of neck,

altered consciousness etc.), respiratory (e.g. shortness of breath, cough), cardiac (chest pain, palpitations), abdominal symptoms (e.g. and. pain, nausea/vomiting, diarrhoea).

Table 4: Inclusion and Exclusion criteria

Inclusion criteria:
• Able and willing to give informed consent
• Age: ≥15 years
• Clinical signs of infection as determined by health officer, resident, or study nurse in charge
• Able and willing to adhere to the study procedures (in particular, allow two complete pairs of blood cultures to be taken if required)
Exclusion criteria:
• Not able and willing to give informed consent or to comply to the study protocoll
• Administration of antibiotics during the current admission at the Emergency Outpatient Department (EOPD) of the respective study centre, prior study participation invitation. Antibiotics administered before the current admission were not an exclusion criterion.

The primary endpoint of the study was in-hospital mortality, which was assessed by our study team, who regularly monitored the patients, communicated with the treating clinicians, and reviewed hospital charts.

2.3 Data collection

The data collectors were trained for data retrieval from medical charts, measurement of vital signs, blood sampling, as well as the use of the Glasgow Coma Scale (GCS), and the Alert Verbal Pain Unresponsive scale (AVPU).

After informed consent was obtained, baseline data were collected by a trained study nurse. All participants' data were pseudonymized by assigning a unique study number to each patient. Each patient was interviewed by a data collector concerning her or his chief complaints, regular and current medication, previous contacts to the health care system, comorbidities, as well as alcohol, cigarettes, and khat usage (i.e. a commonly chewed plant for its stimulant effects, mainly due to its cathinone content, a compound similar to amphetamine). Moreover, the patient's sex, age, address, telephone number,

date and time of inclusion, as well as time of admission to the EOPD, constitution, and the occurrence of open wounds were recorded.

Also assessed on admission were vital parameters including pulse rate, respiratory rate, systolic and diastolic blood pressure, temperature, oxygen saturation, GCS, and AVPU.

Blood samples were collected from the participants to measure various parameters. These included red blood cells (RBC), haemoglobin (Hgb), mean corpuscular volume (MCV), platelets (PLT), white blood cells (WBC), and the subtypes of WBC, including lymphocytes, granulocytes, and mid-cells. In addition, metabolic and organ function were assessed by measuring blood glucose, lactate, and creatinine levels. Finally, a blood smear and thick film to test for malaria and spirochetes, as well as an HIV-rapid-test, were also performed. If the HIV test was positive, a test for cryptococcal-antigen was applied.

Furthermore, the study team extracted data from the medical charts including the clinical diagnosis of admission and the diagnosis of sepsis according to the treating clinician. If the medical charts did not provide this information, the clinician was asked by the data collector.

At the end of each patient's treatment, information about the patient's treatment and hospital discharge were gathered. The discharge information included in case of fatality the reason for death and the date of decease. For all other cases, the reason for and date of discharge were documented. To assess 28-day mortality, each patient or their contact person was called 28 days after hospital discharge. If they could not be reached, the study nurse team tried to contact them or their contact person at least two more times.

The consent forms and worksheets can be found in the appendix (Suppl. Fig. 1-Suppl. Fig. 8). All data were entered independently by two different study team members and crosschecked to minimise data entry errors.

2.4 Statistical analysis

Statistical analyses were performed using SPSS version 29.0 and Python version 3.9.7 within a Jupyter notebook environment. The following libraries were used: pandas, numpy, sklearn, imblearn, statsmodels, scipy, matplotlib, seaborn, math, functools,

copy, and re. For all statistical analyses, we considered a p-value of <0.05 as significant, except for Bonferroni adjusted analyses.

First, we summarized the demographic and clinical characteristics of the study participants.

Second, we assessed the performance of previously established scoring systems in predicting mortality within the resource-limited environment described above. This analysis included the application of the existing scores and the evaluation of their predictive accuracy and discriminative power.

Third, we modified the existing scoring systems by adjusting the cutoff values of each parameter within each score to optimize the performance of mortality prediction in this specific context.

2.4.1 Demographic and clinical characteristics of the study population

To assess the distribution of the data, the Shapiro-Wilk test, as well as visual methods, were used. No normally distributed continuous variables were included in the statistical evaluation. Categorical data are presented as frequencies (proportions), and continuous data as median (interquartile range (IQR)). Differences between survivors and non-survivors were assessed using Fisher's exact test, χ^2, or Mann-Whitney-U-test as appropriate.

2.4.2 Evaluation of Score Performance

Five scores were selected for evaluation. These scores are easy to use and need little to no laboratory equipment.

Systemic Inflammatory Response Syndrome (SIRS) Criteria

As described in the introduction, the SIRS criteria are part of the sepsis-1 definition[5]. Because of their long history of usage and their popularity, we decided to evaluate the SIRS criteria in this study. The SIRS criteria are fulfilled when two or more of the following criteria are present: temperature <36.0°C or >38.0°C, heart rate ≥90 beats/minute, respiratory rate ≥20 breaths/minute, or white blood cell count <4,000, >12,000, or the presence of more than 10% immature neutrophils (see Table 1)[5].

Modified Early Warning Score (MEWS)

The Modified Early Warning Score was developed to identify patients at risk of deterioration, who might benefit from a higher level of care, e.g., a transition to an intensive care unit[6]. MEWS is calculated using the following variables: respiratory rate, temperature, systolic blood pressure, heart rate, and AVPU as described in Table 5[6].

Table 5: Modified Early Warning Score by Subbe et al.[6]

Parameter	3	2	1	0	1	2	3
Respiratory Rate, breaths/minute		<9		9-14	15-20	21-29	≥30
Temperature, °C		<35.0		35.0-38.4		≥38.5	
Systolic blood pressure, mmHg	<70	71-80	81-100	101-199		≥200	
Heart rate, beats/minute		<40	41-50	51-100	101-110	111-129	≥130
Altered mental status: AVPU				A	V	P	U

Table 5: Calculation of the Modified Early Warning Score according to Subbe et al.[6] to identify patients at risk of deterioration. The score is based on respiratory rate, temperature, systolic blood pressure, heart rate, and mental status. Score points are assigned to patients with suspected infection whose parameters deviate from the normal value range. Higher scores indicate a greater risk of deterioration. A = Alert; V = Response to verbal stimuli, P = Response to painful stimuli, U = Unresponsive.

National Early Warning Score (NEWS)

In 2012, the Royal College of Physicians created the National Early Warning Score to standardise the detection of people with acute illness within the National Health System in the United Kingdom[7]. The NEWS is calculated using the following parameters: respiratory rate, oxygen saturation, supplemental oxygen, temperature, systolic blood pressure, heart rate, and altered mental status (see Table 6)[7].

There was a high number of missing values for oxygen supplementation in our dataset. Therefore, we decided to compute a modified NEWS (in the following called mNEWS) by eliminating oxygen supplementation as a score feature.

Table 6: National Early Warning Score by the Royal College of Physicians[7]

Parameter	3	2	1	0	1	2	3
Respiratory Rate, breaths/minute	≤8		9-11	12-20		21-24	≥25
Oxygen saturation, %	≤91	92-93	94-95	≥96			
Any Supplemental Oxygen*		yes		no			
Temperature, °C	≤35.0		35.1-36.0	36.1-38.0	38.1-39.0	≥39.1	
Systolic blood pressure, mmHg	≤90	91-100	101-110	111-219			≥220
Heart rate, beats/minute	≤40		41-50	51-90	91-110	111-130	≥131
Altered mental status: AVPU				A			V, P, U

Table 6: Calculation of the National Early Warning Score by the Royal College of Physicians[7] to identify patients with acute illness. Higher scores indicate a greater risk of acute illness. In this study, a modified version of the NEWS (mNEWS) was used, not including the parameter 'any supplemental oxygen', due to high number of missing data for this variable. *Not included in mNEWS. A = Alert; V = Response to verbal stimuli, P = Response to painful stimuli, U = Unresponsive.

Quick Sequential Organ Failure Assessment (qSOFA)

As described in the introduction, the qSOFA is calculated using only the following three criteria: respiratory rate ≥22 breaths/minute, altered mentation: GCS <15, and systolic blood pressure ≤100 mmHg (see Table 3)[8]. Two or more positive criteria in a patient are associated with a poorer outcome[8].

Universal Vital Assessment (UVA)

The UVA was developed to identify patients with the highest risk of death among hospitalized adults in Sub-Saharan Africa[12]. Additionally, to six clinical parameters, the UVA also takes the HIV serostatus into account. The score is calculated as reported in Table 7[12].

Table 7: Universal Vital Assessment by Moore et al.[12]

Parameter	Points
Respiratory rate ≥30 breaths/minute	1
Heart rate ≥120 beats/minute	1
HIV anamnesis positive	2
Systolic blood pressure <90 mmHg	1
Oxygen saturation <92%	2
Termperature <36°C	2
Altered mental status: Glasgow Coma Scale <15	4

Table 7: Calculation of the Universal Vital Assessment by Moore et al.[12] to identify patients with the highest risk of death. The score is based on respiratory rate, heart rate, HIV anamnesis, systolic blood pressure, systolic blood pressure, oxygen saturation, temperature, and mental status. Score points are assigned to patients with suspected infection whose parameters deviate from the normal value range. Higher scores indicate a greater risk of death.

Missing data and score calculation

For the evaluation of the score performance, missing values were assumed to be in the normal range, and no additional score points were added to calculate the scores. SIRS, MEWS, mNEWS, qSOFA, and UVA were calculated at admission for all study participants as previously described (see Table 1, Table 3, and Table 5 - Table 7).

Statistical analysis

Univariate logistic regression models were used to calculate the odds ratio (OR) with 95% CIs for in-hospital mortality with each score point increase.

Model discrimination was assessed by calculating sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and AUROC. For this analysis, we focused on score cutoff values previously used in the literature[5–8,12,64]. We calculated the confidence intervals (CI) for the AUROCs using bootstrapping (10,000 iterations). To evaluate whether the differences between the AUROCs are statistically significant, we used DeLong's test[65].

Additionally, to account for baseline risk, multivariable logistic regression models were calculated including age, sex, and laboratory-confirmed HIV serostatus. There were 10 patients without laboratory HIV test results. For these 10 patients, we used the anamnestic HIV status. Each score, SIRS, MEWS, mNEWS, qSOFA, and UVA, was

added separately as a continuous variable to the baseline risk model. Adjusted ORs with 95% CIs were calculated to estimate the association of increase in one score point and in-hospital mortality.

Model discrimination of the adjusted models was assessed using AUROCs with 95% CIs. DeLong's method[65] was used to determine significant differences between the AUROCs.

2.4.3 Score adaptation and evaluation of the adapted scores

Finally, we adjusted the scores mentioned above to our specific resource-limited context, in order to improve mortality prediction and identification of patients who are more likely to have a poor outcome.

Score adaptation

First, the dataset was split into training and testing datasets, using a testing data set size of 30% of the cases in the complete dataset. For the cutoff value analysis, missing nominal and ordinal data were imputed using the most frequent value. Missing metric values were iteratively imputed, modelling each feature that has missing values based on the other features.

The dataset was balanced using the Synthetic Minority Over-sampling Technique for categorical and continuous variables (SMOTENC)[66], an upsampling approach that generates synthetic samples by interpolation, rather than duplicating existing ones.

The identification of the ideal feature cutoff for each score and the assignment of the optimal score point values were performed as suggested by Moore et al.[12]. First, decision trees for each of the scores' features on in-hospital mortality were used to determine optimal cutoff values. The threshold that resulted in the smallest Gini impurity within the node was selected as the cutoff value.

For each variable, we determined whether each case had a value above or below/equal to the optimal split point that we identified in the decision tree analysis. We then created a Boolean variable indicating whether each case had a higher probability of in-hospital mortality based on that specific variable. These Boolean variables were used as independent variables in a linear regression model, while in-hospital mortality served as the dependent variable. The relative weights determined by the linear regression

model were then translated into a point-based scoring system. This was achieved by dividing each beta coefficient by the smallest absolute value beta coefficient. The normalized beta coefficients were rounded and used as score points assigned to the particular feature. For each study participant, we then summed up all assigned points to calculate the new adapted scores.

Evaluation of the adapted scores

To evaluate the new cutoff values determined using the training dataset, the modified scores were calculated for the testing dataset. AUROCs, as well as their associated CIs, of all scores (original and adapted) were calculated in both the training and the testing datasets. To assess whether the differences between the AUROCs of all evaluated scores are statistically significant, we again used DeLong's method[65]. Logistic regression models were used to calculate the OR and 95% CIs for in-hospital mortality with each score point increase for the original and the adapted scores in the two datasets.

For the original scores, we calculated sensitivity, specificity, PPV, and NPV using the standard cutoff values as described above. For the adapted scores, we used various cutoff values and determined which ones may be helpful for clinical settings to decide who needs an escalation of care. Our goal was to achieve a high sensitivity, to identify as many patients at risk of death as possible, while maintaining an acceptable specificity and PPV, to avoid unnecessary use of limited resources.

3 Results

876 patients, who presented at the emergency departments during the working hours of our staff, fulfilled the inclusion criteria and gave their consent to participate in the study. 25 patients were retrospectively excluded due to missing data on in-hospital mortality, leaving a total of 851 patients for analysis.

3.1 Missing Data

The overall percentage of missing data in the whole dataset was 2.73%. Of the values needed to calculate the scores, 5.68% were missing. The highest percentage of missing values among the score parameters was found for the NEWS parameter 'Supplementary Oxygen administered'. As we only started collecting data on oxygen supplementation after the inclusion of the first 271 patients, data for 44.1% (n = 375) was missing. We therefore calculated a modified NEWS without data on supplementary oxygen. For the other score parameters, the percentage of missing data ranged from 0% (n = 0) to 7.3% (n = 62). For more information on missing values, see Suppl. Table 2.

3.2 Demographic and clinical characteristics

Of the 851 patients included in the analysis, 569 (66.9%) patients were included in Asella, while 282 (33.1%) were included at the study site in Adama. 456 (53.6%) study participants were male. The median (IQR) age was 28 (22.0–40.0) years. The in-hospital mortality rate was 4.5% (n = 38), while 7.1% (n = 56) participants died within 28 days after presentation at the emergency department.

A detailed summary of the descriptive statistics of the study population can be found in Suppl. Table 3. A selection of the most important data is presented in Table 8.

3.2.1 Symptoms and diagnosis at presentation

The most common symptoms the participants showed were fever, abdominal symptoms, and neurological symptoms, which were present in 67.3% (n = 573), 44.7% (n = 380), and 43.0% (n = 366) of the participants, respectively. Similarly, abdominal infection was the most frequent main diagnosis at admission, accounting for 30.2% (n = 256)

diagnoses, followed by 25.4% (n = 216) respiratory tract infections, 12.4% (n = 105) meningitis/encephalitis, 8.1% (n = 69) urinary tract infection, 3.7% (n = 31) skin infections/ abscesses, 0.6% (n = 5) surgical site infections, and 0.1% (n = 1) osteomyelitis. In 9 (1.1%) cases, the diagnosis was unclear at admission. For 142 (16.7%) participants, the treating clinician did not find one of the aforementioned main diagnoses appropriate and added an alternative diagnosis. Among these other diagnoses, malaria, infective endocarditis, and puerperal sepsis were common. The treating physician diagnosed sepsis in 42.0% (n = 342) of the cases.

3.2.2 Medical history and physical constitution

Regarding the medical history of the patients, 11.5% (n = 98) of the participants were taking medication on a regular basis. 10.7% (n = 91) had a non-infectious chronic disease, while 8.1% (n = 69) of the patients had a chronic infectious disease. 46 (5.4%) patients or their companions reported during the medical history taking that the participant had been diagnosed with HIV. We collected information on HIV treatment for 44 patients, of whom 35 (79.5%) were receiving antiretroviral therapy. Among the HIV-positive patients, 2 (0.2%) also reported being diagnosed with tuberculosis. An additional 18 (2.1%) participants reported to be positive for tuberculosis (open or latent) only.

91.6% (n = 759) of the participants were estimated by the study nurses to have a regular body weight, while 7.7% (n = 64) were estimated to be massively underweight, and 0.7% (n = 6) to be massively overweight.

3.2.3 Vital signs

The median body temperature, pulse rate, respiratory rate, and oxygen saturation levels at admission were 37.9°C (IQR: 37–38.6), 102.5 beats per minute (IQR: 88–118), 26 breaths per minute (IQR:24–31.5), and 95% (IQR: 92–97). The median systolic blood pressure was 100 mmHg (90–110), while the median of diastolic blood pressure was 60 mmHg (60–70).

The majority of the patients did not have an altered consciousness. The median GCS was 15 (IQR: 15–15). A total of 93.2% (n = 779) were assessed to be alert, while 3.2% (n = 27)

responded to verbal stimuli, 2.5% (n = 21) responded to painful stimuli, and 1.1% (n = 9) were non-responsive.

3.2.4 Lab results

13 (1.6%) patients were tested positive for malaria in the blood smear test. Of those, 6 (0.7%) were infected with *Plasmodium falciparum*, while 4 (0.5%) were tested positive for *Plasmodium vivax/ovale.* We tested 10.2% (n = 86) of the participants positive for HIV.

3.2.5 Differences between survivors and non-survivors

Survivors and non-survivors were also analysed separately. Differences between the two groups were examined. Descriptive statistics of the two groups are reported in detail in Suppl. Table 3.

Significant differences, after Bonferroni adjustment for multiple comparisons, were observed in the following variables: sepsis diagnosis according to the treating clinician, presence of an open wound, respiratory rate, administration of supplementary oxygen, altered consciousness (GCS), as well as laboratory results (platelet count, lactate levels, blood glucose, and blood culture results). Further details on the significant variables as well as the parameters used in the aforementioned scores can be found in Table 8.

Table 8: Comparison of Survivors and Non-Survivors

Parameter	Total (n = 851)	Survivors (n = 813)	Non-Survivors (n = 38)	p-value
Sepsis Diagnosis				
Sepsis diagnosis according to treating clinician, n (%)	342 (42)	310 (39.9)	32 (86.5)	**<0.001**
Medical history				
Known non-infectious chronic disease, n (%)	91 (10.7)	83 (10.2)	8 (21.1)	0.035
Infectious chronic diseases				0.022
No known infectious chronic disease, n (%)	779 (91.9)	747 (92.2)	32 (84.2)	
Known infectious chronic disease, n (%)	69 (8.1)	63 (7.8)	6 (15.8)	
HIV, n (%)	46 (5.4)	41 (5.1)	5 (13.2)	
Tuberculosis, n (%)	18 (2.1)	18 (2.2)	0 (0.0)	

Tuberculosis and HIV, n (%)	2 (0.2)	2 (0.2)	0 (0.0)	
Viral Hepatitis, n (%)	3 (0.4)	2 (0.2)	1 (2.6)	
Physical examination				
Open wound, n (%)	36 (4.3)	30 (3.8)	6 (15.8)	**<0.001**
Temperature (°C), median (IQR)	37.9 (37.0–38.6)	37.9 (37.0–38.6)	38 (36.12–38.58)	0.218
Pulse rate (beats/min), median (IQR)	102.5 (88–118)	102 (88–117.5)	110 (92–125)	0.081
Respiratory rate (breaths/min), median (IQR)	26 (24–31.5)	26 (24–30)	32 (25–36)	**<0.001**
Systolic blood pressure (mmHg), median (IQR)	100 (90–110)	100 (90–110)	90 (80–110)	0.099
Oxygen saturation (%), median (IQR)	95 (92–97)	95 (92–97)	92.5 (89–96)	0.009
Supplementary oxygen application, n(%)	33 (6.9)	25 (5.5%)	8 (42.1)	**<0.001**
Altered consciousness (Glasgow Coma Scale), median (IQR)	15 (15–15)	15 (15–15)	14 (13–15)	**<0.001**
Altered consciousness (AVPU)				0.001
Alert, n (%)	779 (93.2)	757 (94.5)	22 (62.9)	
Responds to verbal stimuli, n (%)	27 (3.2)	24 (3)	6 (17.1)	
Responds to painful stimuli, n (%)	21 (2.5)	15 (1.9)	4 (11.4)	
Unresponsive, n (%)	9 (1.1)	5 (0.6)	3 (8.6)	
Laboratory results				
WBC x10^3/µl (IQR)	8.6 (5.55–12.5)	8.5 (5.5–12.4)	11 (7.2–15.4)	0.037
PLT x10/µl (IQR)	239 (170–318)	240 (174–320)	162 (88–222)	**<0.001**
Lactate (IQR)	2.3 (1.5–4.07)	2.3 (1.5–3.7)	4.6 (2.8–7)	**<0.001**
Blood glucose (IQR)	110 (96–128)	110 (96–127)	130 (112–165)	**0.001**
Blood culture results				**<0.001**
No blood culture taken, n (%)	243 (28.6)	242 (29.8)	1 (2.6)	
Sterile, n (%)	524 (61.6)	497 (61.1)	27 (71.1)	
Positive, n (%)	84 (9.9)	74 (9.1)	10 (26.3)	
HIV, n (%)	86 (10.2)	77 (9.6)	9 (24.3)	0.004

Table 8: Comparison of Survivors and Non-Survivors. P-values written in bold indicate significance after Bonferroni adjustment (p <0.00104). A detailed version of this table including all 48 comparisons, is reported in Suppl. Table 3. WBC = white blood count; PLT = platelets; IQR = interquartile range; AVPU = Alert, response to verbal stimuli, response to painful stimuli, unresponsive score.

3.3 The predictive validity of clinical gestalt

We assessed the clinical diagnosis of sepsis according to the treating clinician. The AUROC for clinician's diagnosis of sepsis as a mortality predictor was 0.73 (95% CI: 0.67, 0.79), with a sensitivity of 0.86 (95% CI: 0.74, 0.97) and a specificity: 0.60 (95% CI: 0.56, 0.64). The PPV and the NPV were 0.09 (95% CI: 0.06; 0.12) and 0.99 (95% CI: 0.98; 1.00), respectively.

3.4 The predictive validity of clinical scores

3.4.1 Distribution of clinical scores and mortality

The scores chosen for evaluation were calculated as described in the section: Evaluation of Score Performance. Fig. 1 to Fig. 5 show the distribution of patients across score values and the associated in-hospital mortality rates.

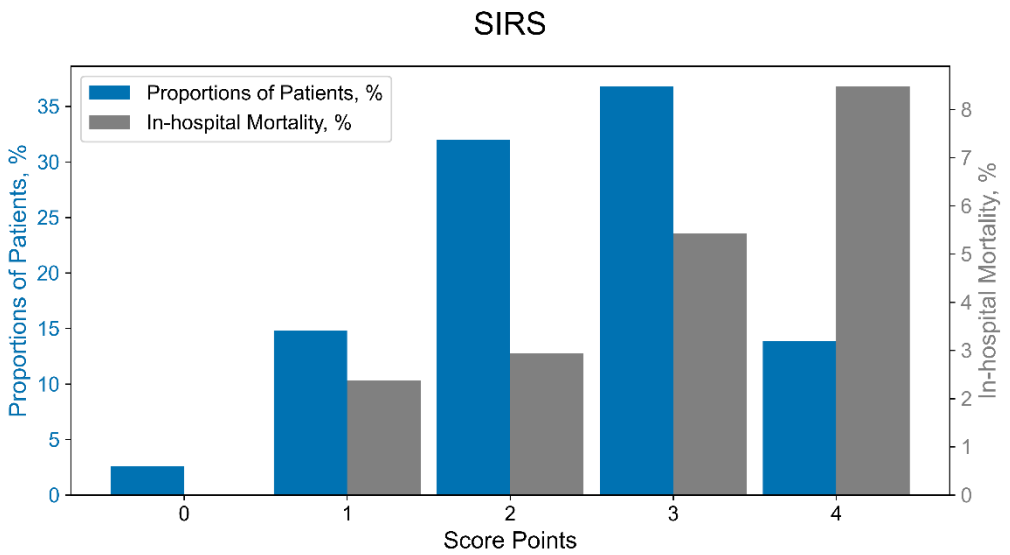


Fig. 1: SIRS distribution and mortality rates across score values
The distribution of patients across SIRS score values and associated in-hospital mortality rates. The blue bars represent the proportion of patients in percent at each SIRS score included in the study. The grey bars show the in-hospital mortality in percent at each SIRS score.

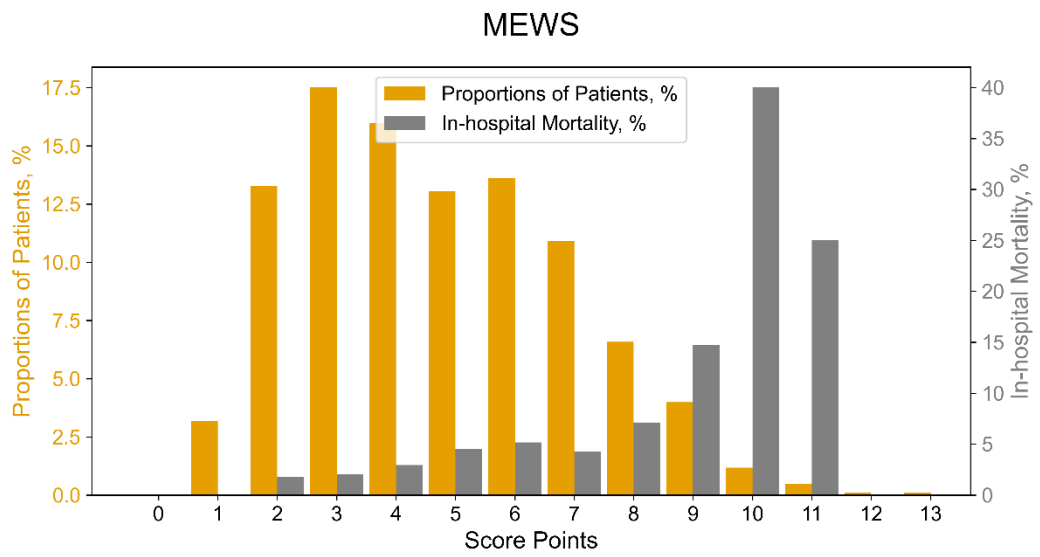


Fig. 2: MEWS distribution and mortality rates across score values
The distribution of patients across MEWS values and associated in-hospital mortality rates. The yellow bars represent the proportion of patients in percent at each MEWS included in the study. The grey bars show the in-hospital mortality in percent at each MEWS.

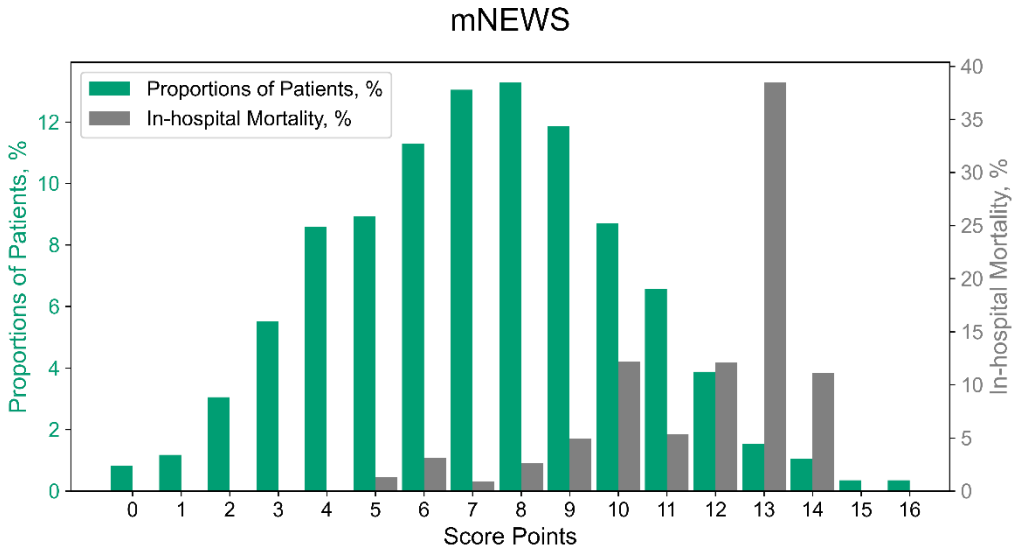


Fig. 3: mNEWS distribution and mortality rates across score values
The distribution of patients across mNEWS values and associated in-hospital mortality rates. The green bars represent the proportion of patients in percent at each mNEWS included in the study. The grey bars show the in-hospital mortality in percent at each mNEWS.

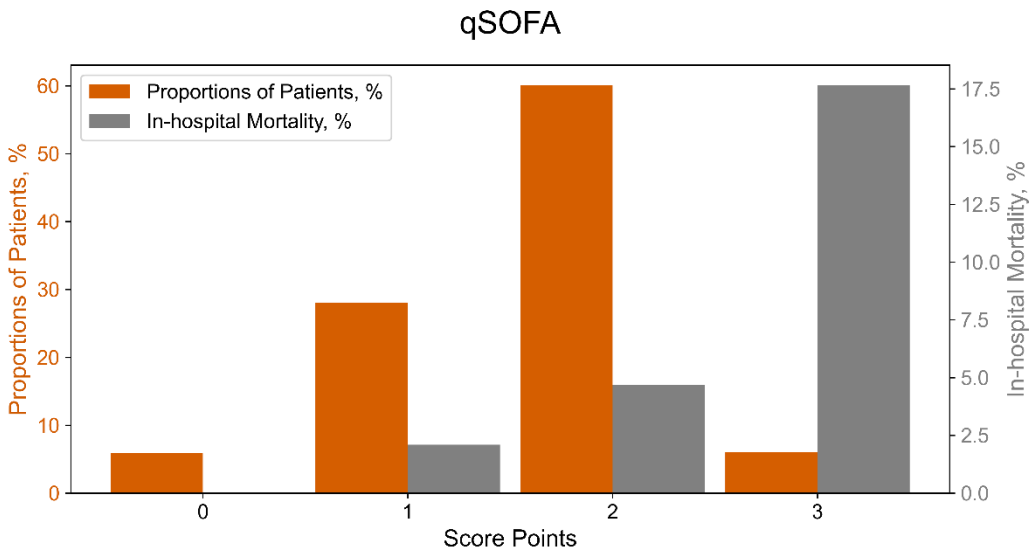


Fig. 4: qSOFA distribution and mortality rates across score values
The distribution of patients across qSOFA score values and associated in-hospital mortality rates. The orange bars represent the proportion of patients in percent at each qSOFA score included in the study. The grey bars show the in-hospital mortality in percent at each qSOFA score.

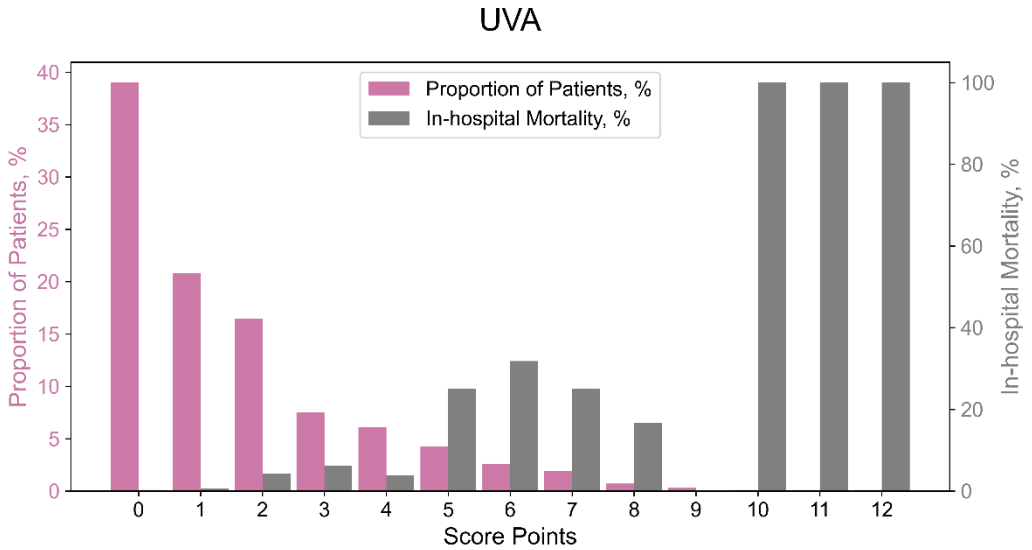


Fig. 5: UVA distribution and mortality rates across score values
The distribution of patients across UVA score values and associated in-hospital mortality rates. The pink bars represent the proportion of patients in percent at each UVA score included in the study. The grey bars show the in-hospital mortality in percent at each UVA score.

The scores ranged as follows: SIRS 0–4 (mode: 3), MEWS 1–13 (median 5, IQR 1–6), mNEWS 0–16 (median: 7, IQR 5–9), qSOFA 0–3 (mode: 2), and UVA 0–12 (median: 1, IQR 0–2).

Mortality was higher at higher score values, although not every increase in score points value resulted in an increase in mortality. Particularly, patients who had a very high score for MEWS or mNEWS survived their hospital stay.

3.4.2 qSOFA calculation

Out of the 851 enrolled patients, 511 (60%) and 51 (6%) patients had a qSOFA score of 2 and 3, respectively. That is a positivity rate of 66% (n = 562). The most common positive parameter combination for a qSOFA score of 2 was a respiratory rate ≥22 and a systolic blood pressure ≤100 (482 out of 511 patients; 94.3%). Less common were the combination of GCS <15 and increased respiratory rate (n = 23; 4.5% of the qSOFA positive participants), as well as the combination of decreased systolic blood pressure and decreased GCS (n = 6; 1.2% of the qSOFA positive participants).

3.4.3 Score performance

As described in section 2.4.2 the scores were evaluated unadjusted (Fig. 6, Fig. 7) and adjusted for baseline risk (Fig. 8, Fig. 9).

Unadjusted score performance

The univariate logistic regression models showed a significant association between one point score increase and in-hospital mortality. The odds of death increased by 70% per one point score increase for SIRS (OR 1.70; 95% CI: 1.18, 2.45), 40% for MEWS (OR 1.40; 95% CI: 1.21, 1.61), 32% for mNEWS (OR 1.32; 95% CI: 1.17, 1.48). The qSOFA showed the strongest association (OR 3.49; 95% CI: 1.93, 6.30), while one point score increase of the UVA was associated with 77% increase in odds of death (OR 1.77; 95% CI: 1.54, 2.04). For more information, see Suppl. Table 4.

The AUROC for SIRS without taking the baseline risk into account was 0.63 (95% CI: 0.55, 0.71), for MEWS 0.70 (95% CI: 0.61, 0.79), for mNEWS 0.72 (95% CI: 0.63, 0.81), for qSOFA 0.67 (95% CI: 0.59, 0.74), and for UVA 0.87 (95% CI: 0.82, 0.92).

UVA demonstrated a significantly higher AUROC compared to all evaluated scores ($p \leq 0.001$). Additionally, mNEWS showed a significantly higher AUROC than SIRS ($p = 0.033$). No significant differences were observed among the AUROCs of the remaining scores (see Suppl. Table 5).

Furthermore, the UVA showed a significant higher AUROC than the clinicians' diagnosis of sepsis in predicting mortality. This was not the case for the other scores. The AUROC of the clinicians' diagnosis of sepsis as a predictor for mortality was significantly larger than that of the SIRS.

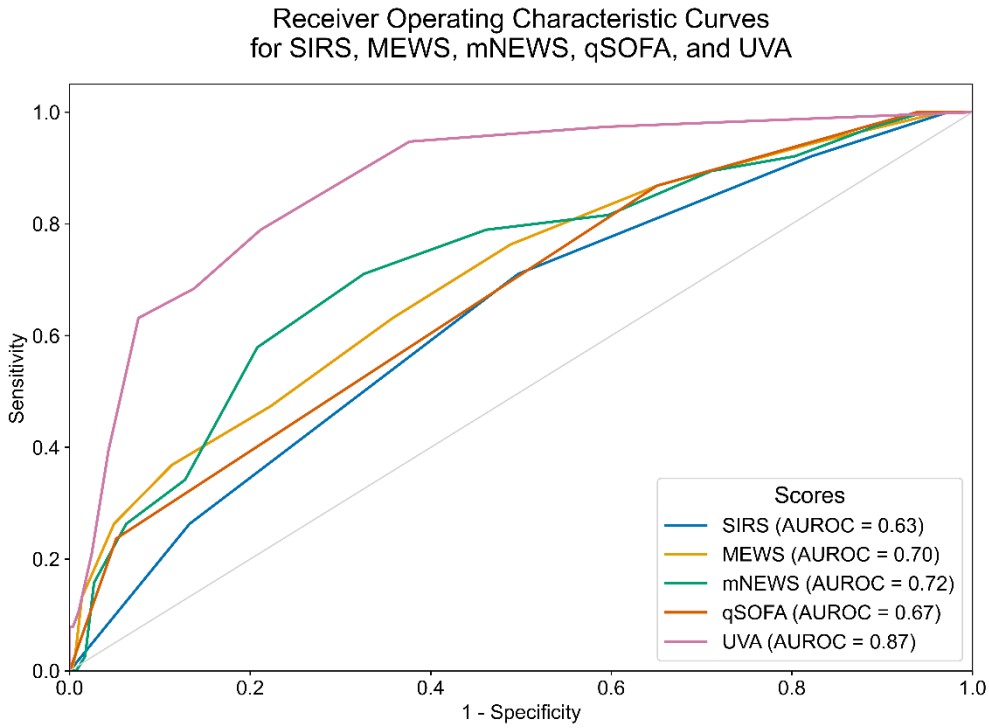


Fig. 6: Receiver operating characteristic curves for mortality discrimination
AUROC = Area under the receiver operating characteristic curve.
Receiver operating characteristic curves for original SIRS, MEWS, mNEWS, qSOFA, and UVA for in-hospital mortality among patients with suspected infection calculated using the complete dataset (n = 851). The curves are not adjusted for baseline risk. The grey line represents a model with no discrimination (random classification).

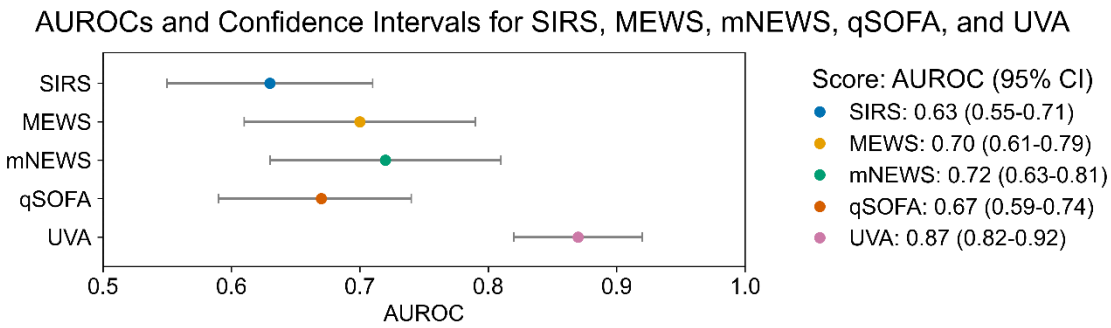


Fig. 7: Area under the receiver operating characteristic curves with 95% confidence intervals (CI) for mortality discrimination

Forest plot for better score comparability. Area under the receiver operating characteristic curves with 95% confidence intervals for original SIRS, MEWS, mNEWS, qSOFA, and UVA for in-hospital mortality among patients with suspected infection calculated using the complete dataset (n = 851). The area under the receiver operating characteristic curves are not adjusted for baseline risk.

Table 9 shows the performance metrics (sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV)) at various score cutoffs for each score.

At a SIRS cutoff value of ≥2, the sensitivity was 0.92, while the specificity was 0.18. For qSOFA with the same cutoff of ≥2, the sensitivity was 0.87, and the specificity was 0.35. For MEWS, the sensitivity at a cutoff value of ≥3 was 0.95, and the specificity was 0.17. At a cutoff value of ≥4 and ≥5, the sensitivity decreased to 0.87 and 0.76, respectively, while the specificity increased to 0.35 and 0.51, respectively. In the case of mNEWS, sensitivity and specificity were 0.79 and 0.54, respectively, for a cutoff of ≥7, and for a cutoff of ≥8, sensitivity decreased to 0.79, while specificity increased to 0.54. For UVA with a cutoff of ≥2, sensitivity was 0.95 and specificity was 0.62. An increase of the cutoff value to ≥4 led to a decreased sensitivity of 0.69 and an increased specificity of 0.86.

Table 9: Predictive capacity of SIRS, MEWS, mNEWS, qSOFA, and UVA at different cutoff values

Score	Sensitivity	Specificity	PPV	NPV
SIRS				
SIRS ≥1	1.00	0.03	0.05	1.00
SIRS ≥2	0.92	0.18	0.05	0.98
SIRS ≥3	0.71	0.50	0.06	0.97
MEWS				
MEWS ≥3	0.95	0.17	0.05	0.99
MEWS ≥4	0.87	0.35	0.06	0.98
MEWS ≥5	0.76	0.51	0.07	0.98
MEWS ≥6	0.63	0.64	0.08	0.97

MEWS ≥7	0.47	0.78	0.09	0.97
MEWS ≥8	0.37	0.89	0.13	0.97
MEWS ≥9	0.26	0.95	0.2	0.97
MEWS ≥10	0.13	0.99	0.31	0.96
MEWS ≥11	0.03	0.99	0.17	0.96
mNEWS				
mNEWS ≥4	0.97	0.11	0.05	0.99
mNEWS ≥5	0.92	0.20	0.05	0.98
mNEWS ≥6	0.89	0.29	0.06	0.98
mNEWS ≥7	0.82	0.40	0.06	0.98
mNEWS ≥8	0.79	0.54	0.07	0.98
mNEWS ≥9	0.71	0.67	0.09	0.98
mNEWS ≥10	0.58	0.79	0.12	0.98
mNEWS ≥11	0.34	0.87	0.11	0.97
qSOFA				
qSOFA ≥1	1.00	0.06	0.05	1.00
qSOFA ≥2	0.87	0.35	0.06	0.98
qSOFA ≥3	0.24	0.95	0.18	0.96
UVA				
UVA ≥2	0.95	0.62	0.11	1.00
UVA ≥3	0.79	0.79	0.15	0.99
UVA ≥4	0.68	0.86	0.19	0.98
UVA ≥5	0.63	0.92	0.28	0.98
UVA ≥6	0.39	0.96	0.30	0.97

Table 9: Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the original SIRS, MEWS, mNEWS, qSOFA, and UVA at different cutoff values. Commonly used cutoff values are written in bold.

Score performance adjusted for baseline risk

The ORs for each increase in score points were adjusted for age, sex, and laboratory-confirmed HIV status. The adjusted ORs were as follows: SIRS 1.71 (95% CI: 1.19, 2.47), MEWS 1.41 (95% CI: 1.22, 1.63), mNEWS 1.30 (95% CI: 1.15, 1.46), qSOFA 3.24 (95% CI: 1.79, 5.85), UVA 1.80 (95% CI: 1.56, 2.09). Full logistic regression model estimates are reported in Suppl. Table 6.

Fig. 8 and Fig. 9 show the AUROC of the baseline risk model, which showed an AUROC of 0.64 (0.55-0.72). SIRS, MEWS, mNEWS, qSOFA, and UVA were added as a continuous variable to the baseline risk model, changing the AUROCs to 0.68, 0.77, 0.75, 0.71, and 0.87, respectively. SIRS and qSOFA added to the baseline risk did not show a significantly higher AUROC than the baseline model. MEWS, mNEWS, and UVA in

combination with the baseline risk showed significantly higher AUROCs (see Suppl. Table 7).

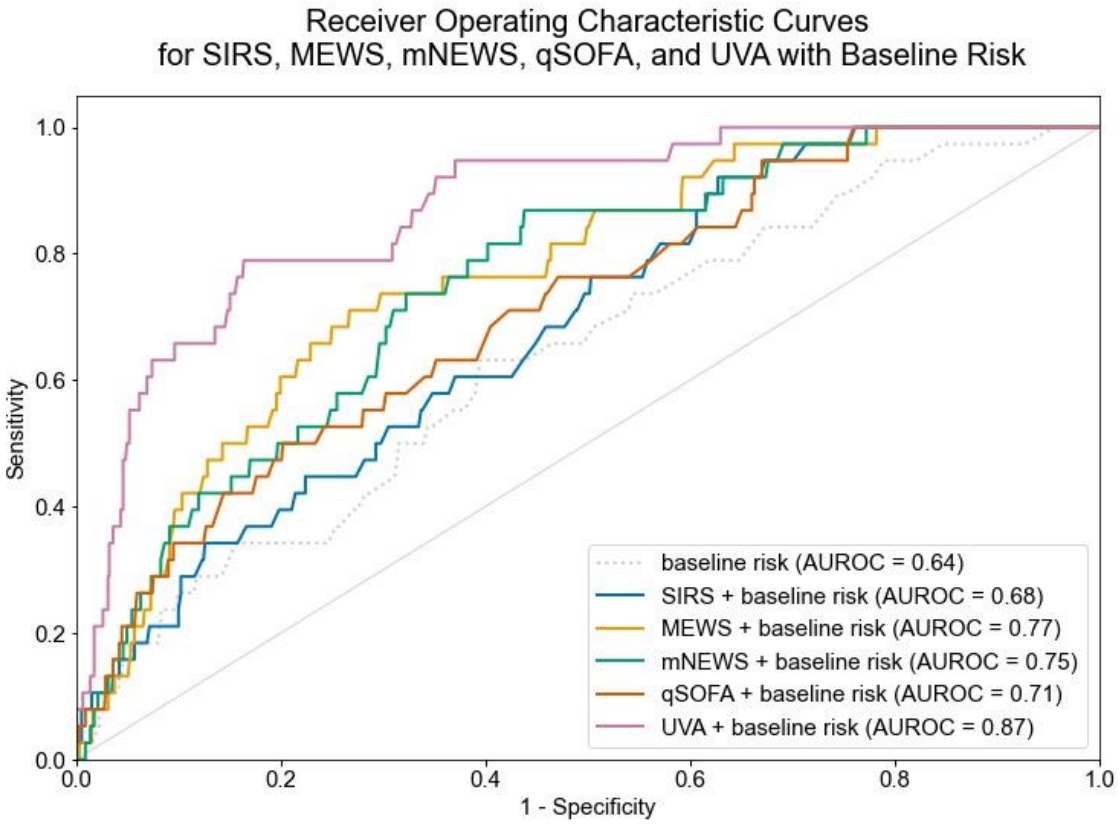


Fig. 8: Receiver operating characteristic curves for mortality discrimination adjusted for baseline risk

AUROC = Area under the receiver operating characteristic curve.

Receiver operating characteristic curves for original SIRS, MEWS, mNEWS, qSOFA, and UVA for in-hospital mortality among patients with suspected infection calculated using the complete dataset (n = 851). The curves are adjusted for baseline risk. The baseline risk model was calculated using age, sex, and HIV status. The grey line represents a model with no discrimination (random classification).

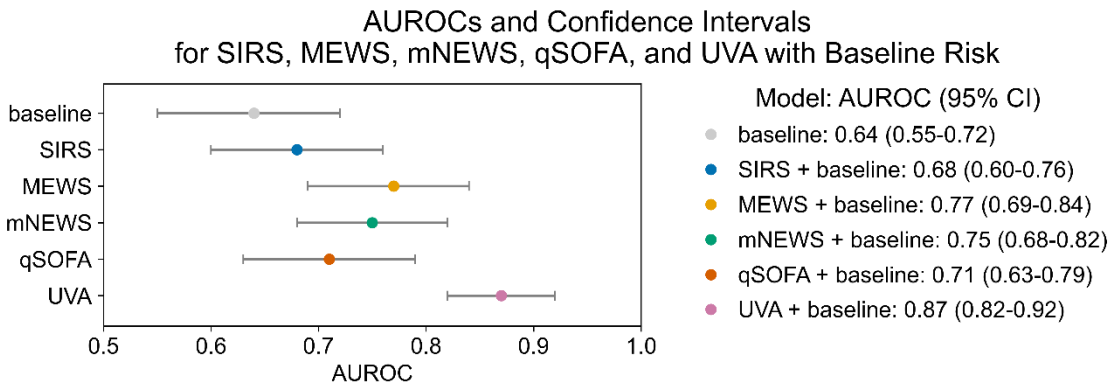


Fig. 9: Area under the receiver operating characteristic curves with 95% confidence intervals (CI) for mortality discrimination adjusted for baseline risk

Forest plot for better score comparability. Area under the receiver operating characteristic curves with 95% confidence intervals for original SIRS, MEWS, mNEWS, qSOFA, and UVA for in-hospital mortality among patients with suspected infection calculated using the complete dataset (n = 851). The curves are adjusted for baseline risk. The baseline risk model was calculated using age, sex, and HIV status.

3.5 Score adaptation to a resource-limited setting

3.5.1 Optimization of score performance on the training dataset

The training dataset (n = 595) was used to train the models and adapt the original scores to a resource-limited setting. The decision tree analysis revealed the cutoff values that led to the smallest gini impurity for each parameter of each score as shown in Table 1. The relative weights that were determined by the linear regression model and translated into a point-based scoring system are also shown in the table. The cutoff value for the clinical parameters for the detection of individuals with higher risk for a poor outcome were determined as follows: temperature <36.4°C, pulse rate >92 beats per minute, respiratory rate >28 breaths per minute, systolic blood pressure <90 mmHg, oxygen saturation <92%, and an altered consciousness, indicated by a GCS <15 or AVPU showing anything other than ‘alert’ or ‘responsive to verbal stimuli’.

Table 10: Calculation of the adapted scores

Variable	cutoff	Adapted SIRS points	Adapted MEWS points	Adapted mNEWS points	Adapted qSOFA points	Adapted UVA points
Temperature (°C)	<36.4	1	3	3		2
Pulse rate (beats/min)	>92	1	2	2		1
Respiratory rate (breaths/min)	>28	1	2	3	1	3
Systolic blood pressure (mmHg)	<90		1	1	2	3
Oxygen saturation (%)	<92			4		1
Altered consciousness (Glasgow Coma Scale)	<15				5	8
Altered consciousness (AVPU*)	P or U		3	3		
Leukocyte count (per $10^3/\mu L$)	>7.1	1				
HIV anamnesis	positive					3
Maximum Score		**4**	**11**	**16**	**8**	**21**

Table 10: Clinical variables, thresholds, and corresponding points for the calculation of the adapted scores: adapted SIRS, adapted MEWS, adapted mNEWS, adapted qSOFA, and adapted UVA. Score points are assigned to patients with suspected infection whose parameter values fall below or exceed the parameter's threshold, depending on the variable. Higher scores indicate a greater risk of in-hospital mortality. The AVPU scale assesses consciousness as Alert (A), Verbal Response (V), Pain Response (P), or Unresponsive (U).

Fig. 10 and Fig. 11 display the AUROCs for mortality prediction with 95% CI using the adapted scores in the training dataset. In the training dataset, the adapted UVA showed the highest AUROC (0.94; 95% CI: 0.93, 0.96) among the adapted scores. It was followed by the adapted qSOFA (0.91; 95% CI: 0.89, 0.93), adapted mNEWS (0.84; CI: 95% 0.81, 0.86), adapted SIRS (0.78; 95% CI: 0.75, 0.80), and adapted MEWS (0.77; 95% CI: 0.74, 0.79).

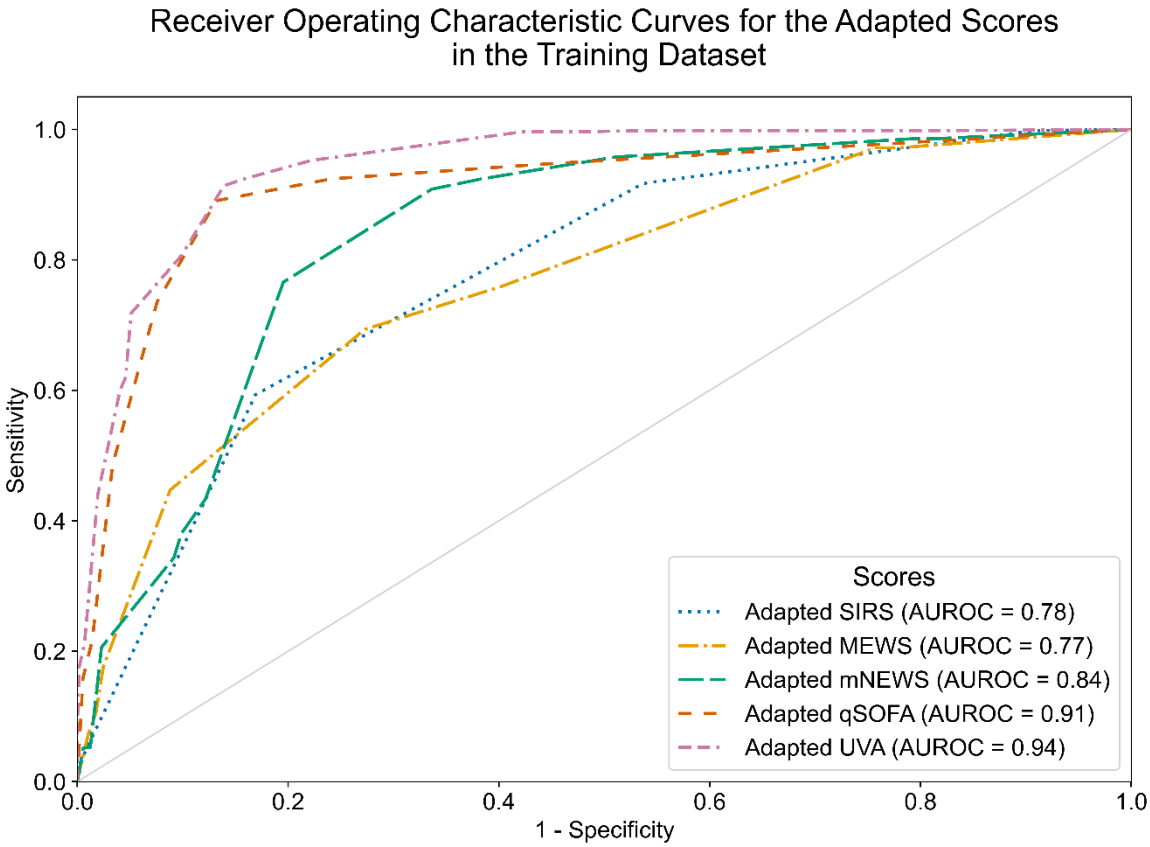


Fig. 10: Receiver operating characteristic curves for the adapted scores for mortality discrimination in the training dataset

AUROC = Area under the receiver operating characteristic curve.

Receiver operating characteristic curves for adapted SIRS, adapted MEWS, adapted mNEWS, adapted qSOFA, and adapted UVA for in-hospital mortality among patients with suspected infection calculated using the upsampled training dataset that was based on 595 cases. The curves are not adjusted for baseline risk. The grey line represents a model with no discrimination (random classification).

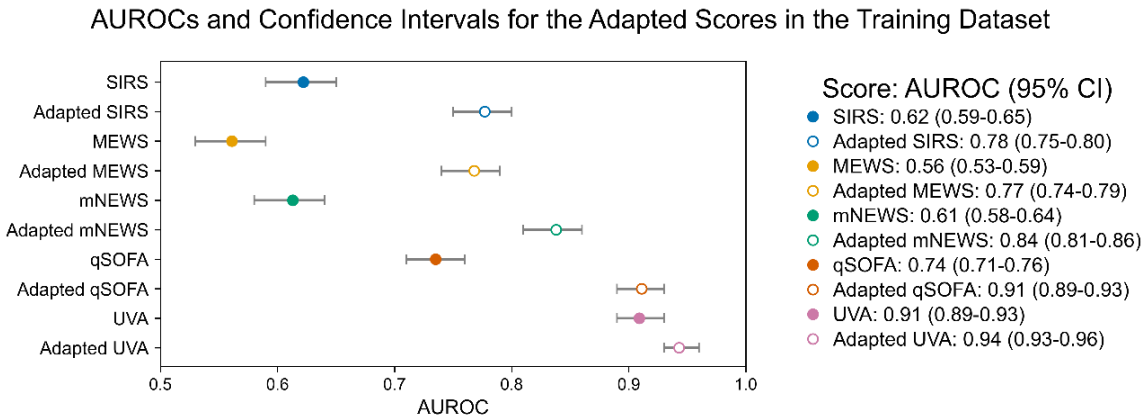


Fig. 11: Area under the receiver operating characteristic curves with 95% confidence intervals (CI) for mortality discrimination

AUROC = Area under the receiver operating characteristic curve.

Forest plot for better score comparability. AUROCS with 95% confidence intervals for SIRS, MEWS, mNEWS, qSOFA, and UVA, alongside their adapted versions, for in-hospital mortality among patients with suspected infection, calculated using the upsampled training dataset that was based on 595 cases. The area under the receiver operating characteristic curves are not adjusted for baseline risk.

Suppl. Fig. 9 to Suppl. Fig. 13 show the ROC curves for mortality prediction using the adapted scores, alongside the original score versions, using the training dataset. Model estimates of the logistic regression models for the calculation of the ORs and CIs for each one score point increase for in-hospital mortality are reported in Suppl. Table 8.

3.5.2 Evaluation of the adapted scores in the testing dataset

The adapted scores were evaluated in the testing dataset (n = 256). The calculated ORs with CIs for each score point increase and mortality are shown in Suppl. Table 9.

In Fig. 12 and Fig. 13 the AUROCs along with their 95% confidence intervals for mortality prediction using the original and adapted scores are reported. In the testing dataset, the adapted UVA also showed the highest AUROC (0.86, 95% CI: 0.76, 0.94). Similarly, the original UVA showed an AUROC of 0.86 (95% CI: 0.76, 0.94). The AUROCs for the other scores were the following: adapted mNEWS 0.85 (95% CI: 0.73, 0.95), adapted MEWS 0.84 (95% CI: 0.69, 0.96), adapted qSOFA 0.82 (95% CI: 0.72, 0.92), adapted SIRS 0.77 (95% CI: 0.67, 0.86), mNEWS 0.77 (95% CI: 0.59, 0.92), MEWS 0.74 (95% CI: 0.57, 0.88), qSOFA 0.73 95% CI: 0.64, 0.83), and SIRS 0.64 (95% CI: 0.47, 0.80).

Fig. 14 to Fig. 18 show the receiver operating characteristic curves for mortality prediction using the adapted scores, alongside the original versions of the scores in the testing dataset.

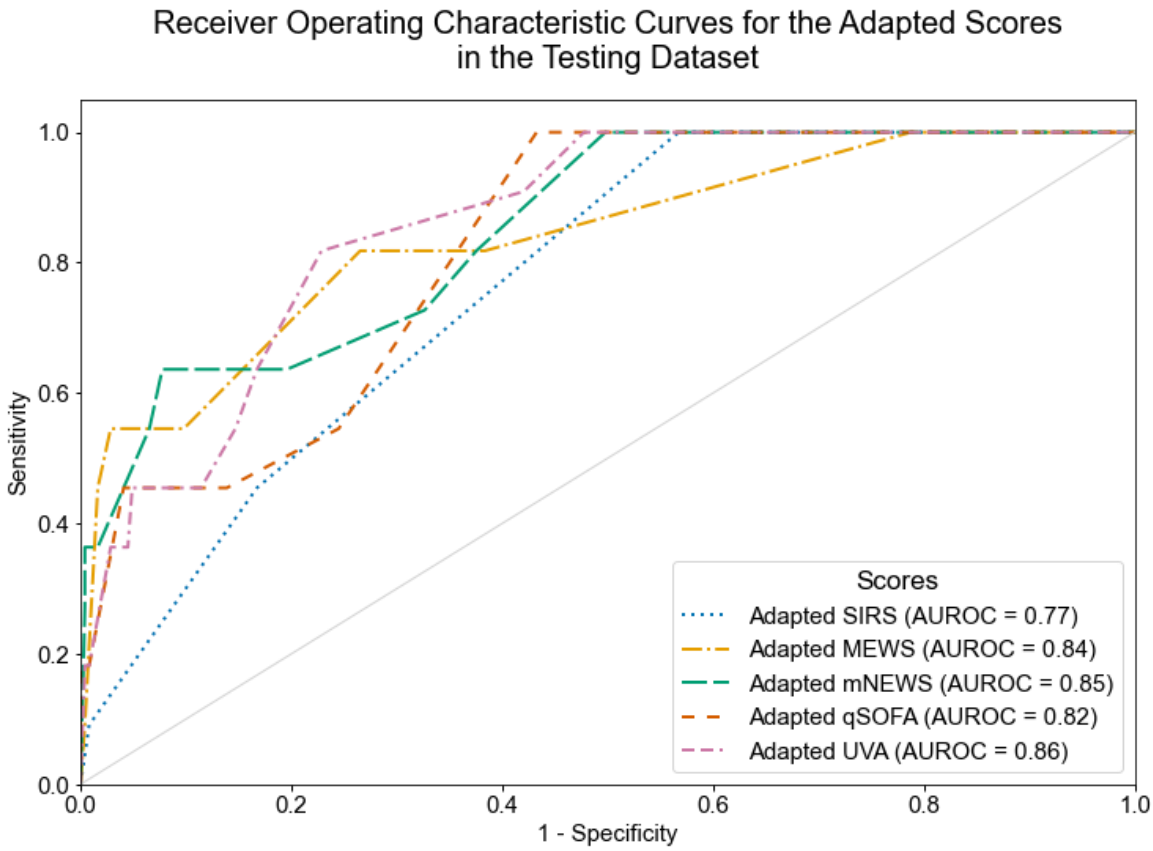


Fig. 12: Receiver operating characteristic curves for the adapted scores for mortality discrimination in the testing dataset

AUROC = Area under the receiver operating characteristic curve.

Receiver operating characteristic curves for adapted SIRS, adapted MEWS, adapted mNEWS, adapted qSOFA, and adapted UVA for in-hospital mortality among patients with suspected infection. The area under the receiver operating characteristic curves were generated using the testing dataset (n = 256) with score adaptations derived from the training dataset (n = 595). The curves are not adjusted for baseline risk. The grey line represents a model with no discrimination (random classification).

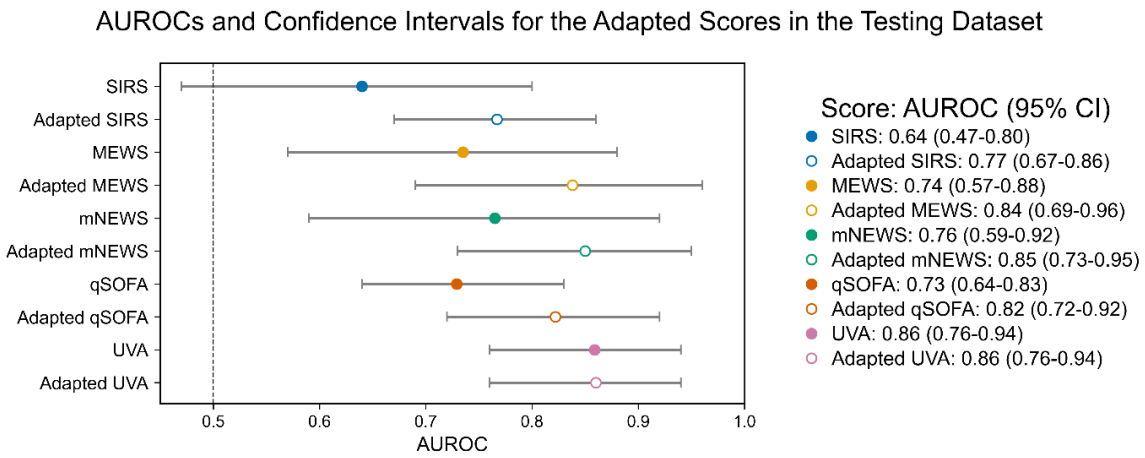


Fig. 13: Area under the receiver operating characteristic curves with 95% confidence intervals (CI) for mortality discrimination

AUROC = Area under the receiver operating characteristic curve.

Forest plot for better score comparability. Area under the receiver operating characteristic curves with 95% confidence intervals for SIRS, MEWS, mNEWS, qSOFA, and UVA, alongside their adapted versions, for in-hospital mortality among patients with suspected infection. The AUROCs were calculated using the testing dataset (n = 256) with score adaptations derived from the training dataset (n = 595). The area under the receiver operating characteristic curves are not adjusted for baseline risk. The vertical dotted line model with no discrimination (random classification).

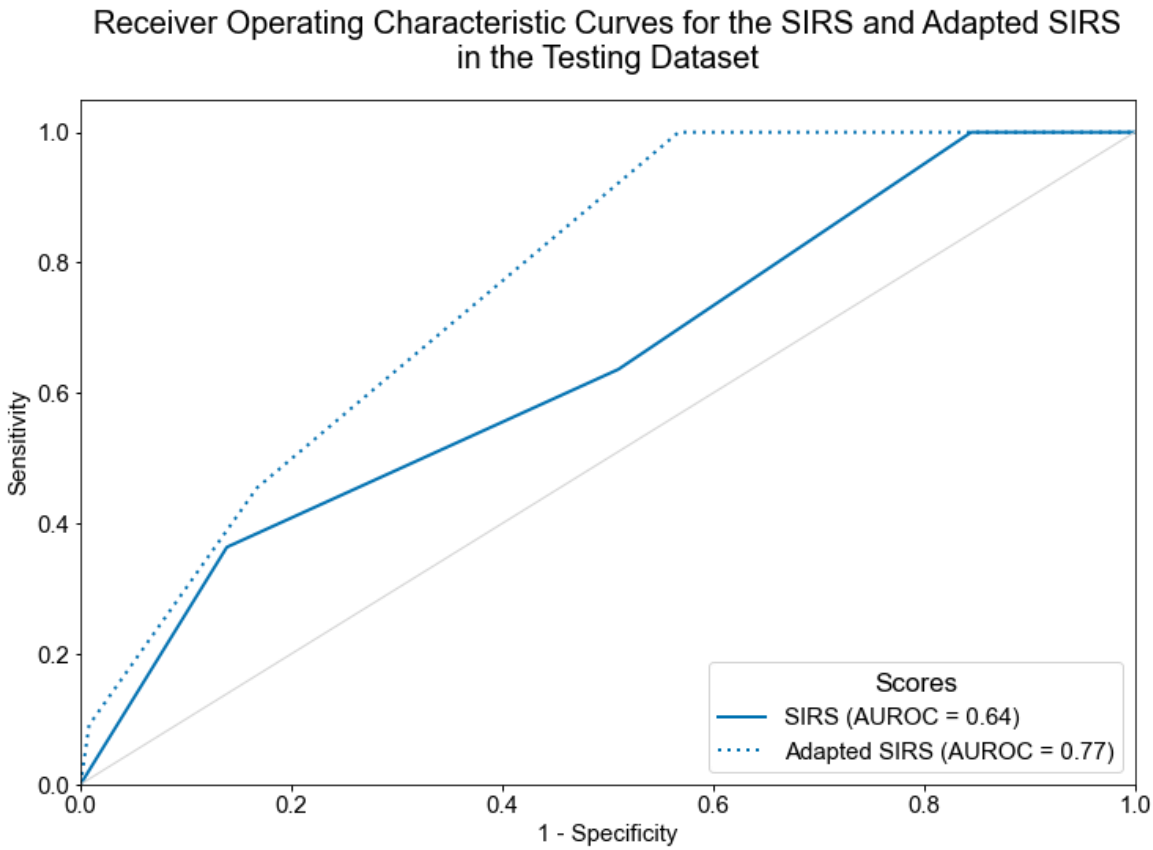


Fig. 14: Receiver operating characteristic curves for SIRS and adapted SIRS for mortality discrimination in the testing dataset

AUROC = Area under the receiver operating characteristic curve.

Receiver operating characteristic curves for SIRS and adapted SIRS for in-hospital mortality among patients with suspected infection. The AUROCs were calculated using the testing dataset (n = 256) with adapted SIRS derived from the training dataset (n = 595). The adapted SIRS did not show a significantly higher AUROC than the original SIRS (p = 0.061).

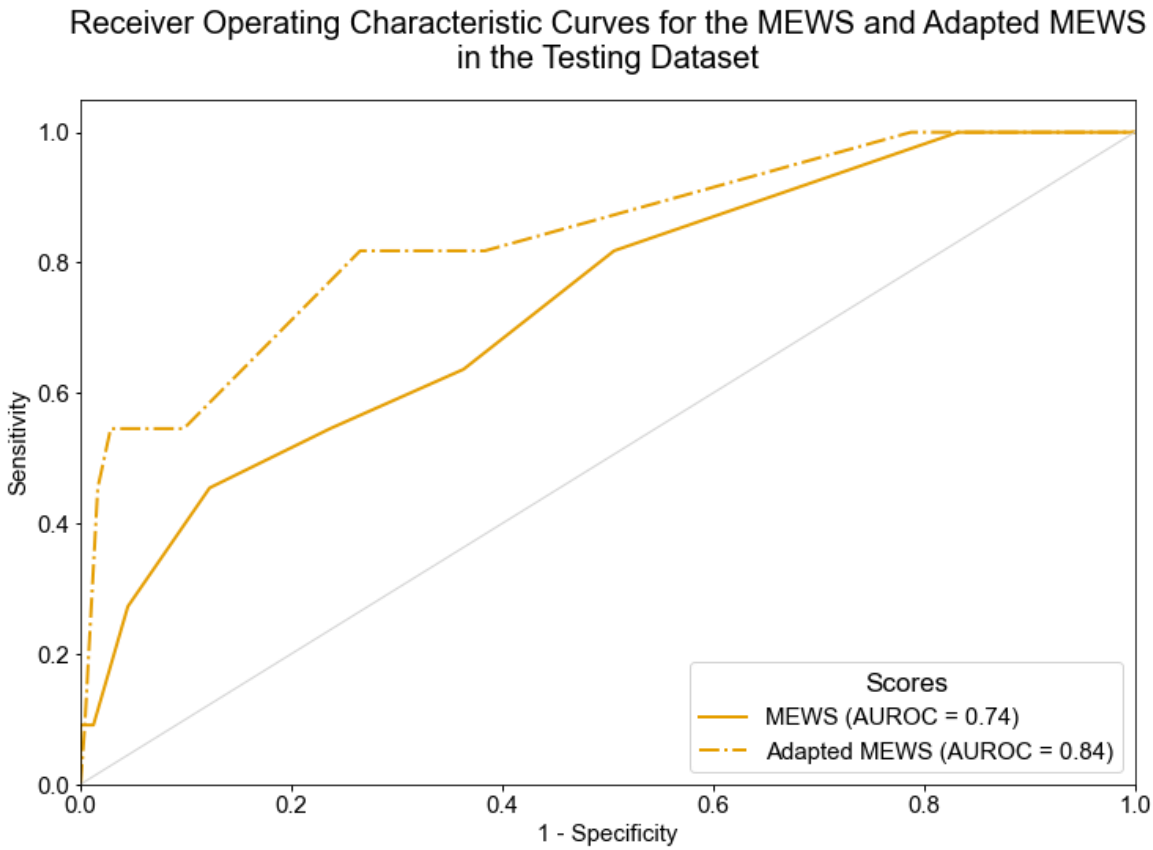


Fig. 15: Receiver operating characteristic curves for MEWS and adapted MEWS for mortality discrimination in the testing dataset

AUROC = Area under the receiver operating characteristic curve.

Receiver operating characteristic curves for MEWS and adapted MEWS for in-hospital mortality among patients with suspected infection. The AUROCs were calculated using the testing dataset (n = 256) with adapted MEWS derived from the training dataset (n = 595). The adapted MEWS showed significantly higher AUROC than the original MEWS (p = 0.017).

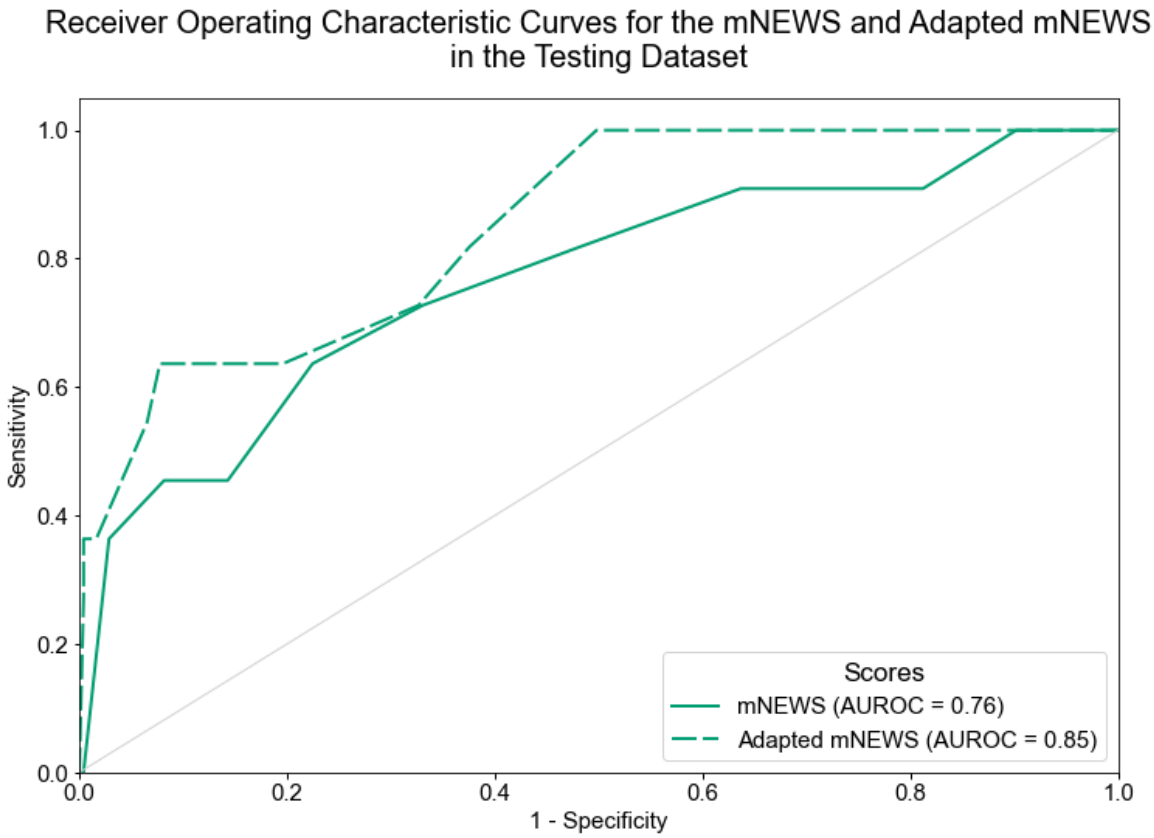


Fig. 16: Receiver operating characteristic curves for mNEWS and adapted mNEWS for mortality discrimination in the testing dataset

AUROC = Area under the receiver operating characteristic curve.

Receiver operating characteristic curves for mNEWS and adapted mNEWS for in-hospital mortality among patients with suspected infection. The AUROCs were calculated using the testing dataset (n = 256) with adapted mNEWS derived from the training dataset (n = 595). The adapted mNEWS showed a marginal significantly higher AUROC than the original mNEWS (p = 0.049).

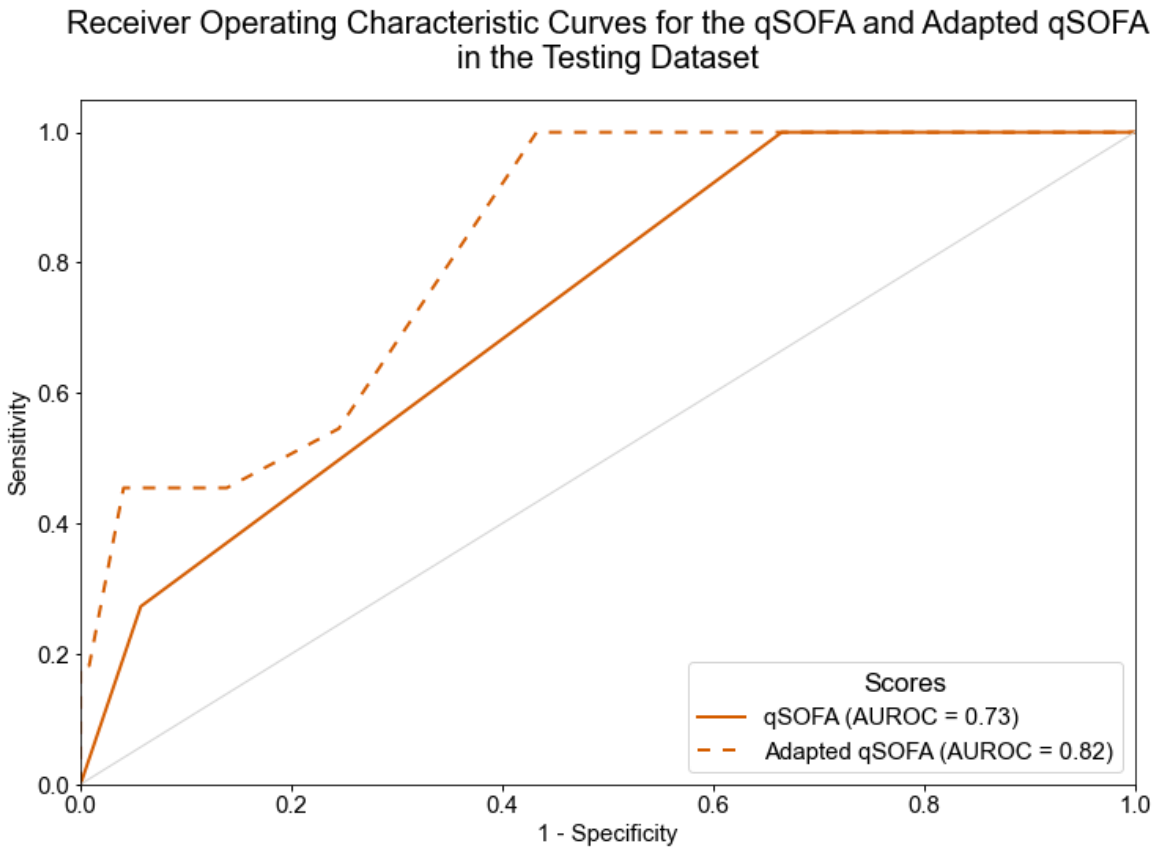


Fig. 17: Receiver operating characteristic curves for qSOFA and adapted qSOFA for mortality discrimination in the testing dataset

AUROC = Area under the receiver operating characteristic curve.

Receiver operating characteristic curves for qSOFA and qSOFA for in-hospital mortality among patients with suspected infection. The AUROCs were calculated using the testing dataset (n = 256) with adapted qSOFA derived from the training dataset (n = 595). The adapted qSOFA showed a significantly higher AUROC than the original qSOFA (p = 0.025).

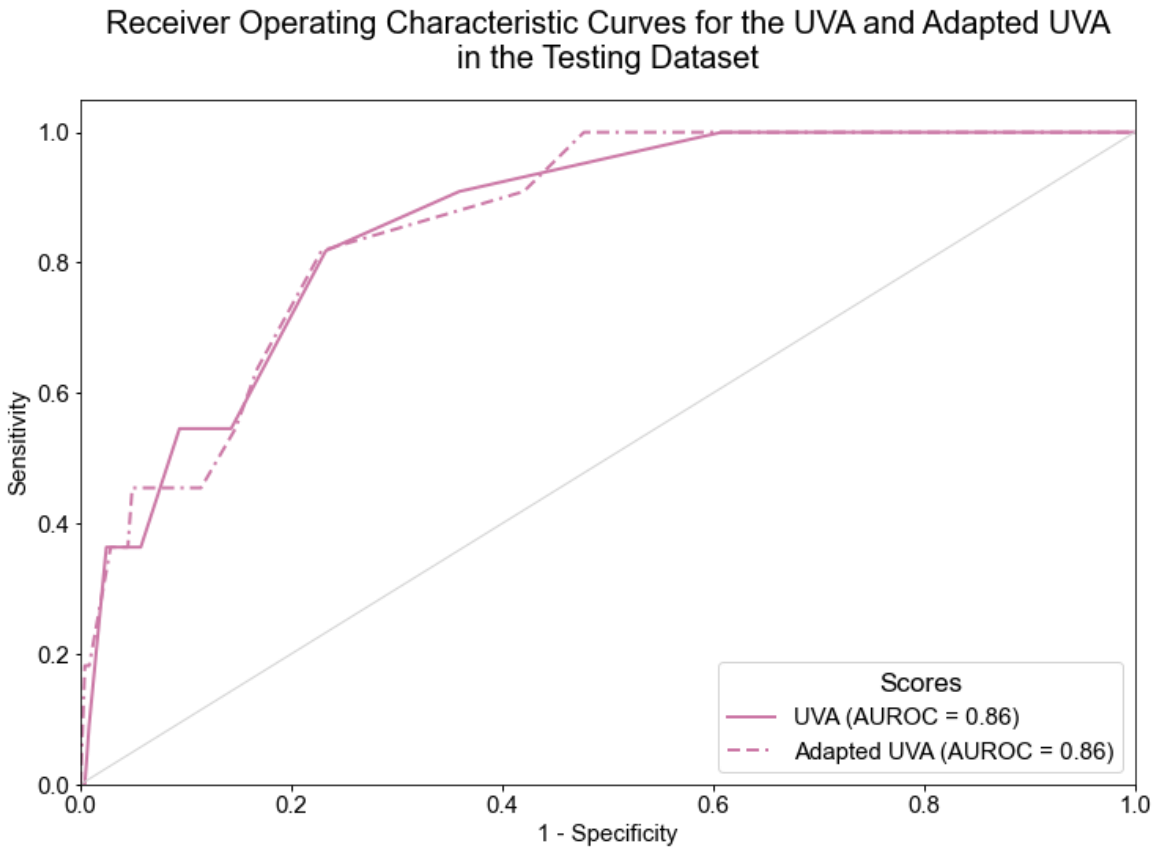


Fig. 18: Receiver operating characteristic curves for UVA and adapted UVA for mortality discrimination in the testing dataset

AUROC = Area under the receiver operating characteristic curve.

Receiver operating characteristic curves for UVA and UVA for in-hospital mortality among patients with suspected infection. The AUROCs were calculated using the testing dataset (n = 256) with adapted qSOFA derived from the training dataset (n = 595). The AUROCs of adapted and original UVA were both 0.86.

The AUROCs for mortality prediction using the adapted MEWS (0.838 vs. 0.735, p = 0.017) and adapted qSOFA (0.822 vs. 0.729, p = 0.025) were significantly higher than those using the original scores. The AUROC of the adapted mNEWS also showed a higher AUROC the original mNEWS (0.850 vs. 0.765). This difference was only marginally statistically significant (p = 0.049). The adapted SIRS showed a higher AUROC for mortality prediction compared to the original (0.767 vs. 0.640). This difference was not statistically significant (p = 0.061). There was no meaningful difference in the performance of adapted UVA and original UVA (0.860 vs. 0.859, p = 0.971) (see Table 11). In Suppl. Table 10 additional comparisons between the scores are reported.

Table 11: Comparison of the AUROCs of the scores and adapted scores.

Score 1	Score 2	AUROC 1	AUROC 2	|ΔAUROC|	p-value
SIRS	**adapted SIRS**	0.640	0.767	0.127	0.061
MEWS	**adapted MEWS**	0.735	0.838	0.103	0.017
mNEWS	**adapted mNEWS**	0.765	0.850	0.085	0.049
qSOFA	**adapted qSOFA**	0.729	0.822	0.093	0.025
UVA	**adapted UVA**	0.859	0.860	0.001	0.971

Table 11: Comparison of the area under the receiver operating characteristic curves (AUROC) of the original scores and their adapted versions in the testing dataset using DeLong's method[65]. Adapted MEWS, adapted mNEWS and adapted qSOFA showed significantly higher AUROCs for predicting mortality than their original versions. The adapted SIRS showed a higher, though not significant, AUROC than the original SIRS. The AUROCs of adapted and original UVA were similar in size.

To decide which patient needs an increase of medical care, the thresholds suggested in Table 12 can be used. A table presenting the performance metrics of more thresholds for the adapted scores is provided in the appendix (see Suppl. Table 11). In the testing dataset (n = 256), the original scores showed high sensitivities ranging from 0.91 to 1 at their standard thresholds, with low PPVs are ranging from 0.05 to 0.1. The only exception was UVA at a cutoff value of ≥4, which showed a sensitivity of 0.55 and a PPV of 0.15.

On the other hand, the adapted scores demonstrated moderate to high sensitivities, ranging from 0.82 to 1, and showed a higher specificity, which ranged from 0.43-0.77. The PPVs of the adapted scores ranged from 0.07-0.14.

Table 12: Predictive capacity of original and adapted scores.

Scores	Sensitivity	Specificity	PPV	NPV
Original Scores				
SIRS ≥2	1	0.16	0.05	1
MEWS ≥3	1	0.17	0.05	1
MEWS ≥4	0.91	0.33	0.06	0.99
MEWS ≥5	0.82	0.49	0.07	0.98
mNEWS ≥5	0.91	0.19	0.05	0.98
mNEWS ≥7	0.91	0.36	0.06	0.99
qSOFA ≥2	1	0.33	0.06	1
UVA ≥2	0.91	0.64	0.1	0.99
UVA ≥4	0.55	0.86	0.15	0.98

Adapted Scores				
Adapted SIRS ≥2	1	0.43	0.07	1
Adapted MEWS ≥4	0.82	0.73	0.12	0.99
Adapted mNEWS ≥4	0.82	0.62	0.09	0.99
Adapted qSOFA ≥1	1	0.57	0.09	1
Adapted UVA ≥5	0.82	0.77	0.14	0.99

Table 12: Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the original, as well as adapted, SIRS, MEWS, mNEWS, qSOFA, and UVA at commonly used cutoff values. Predictive capacity at more thresholds is reported in Suppl. Table 11.

4 Discussion

4.1 The need for clinical scoring systems

At our study sites, neither the evaluated scores nor other decision-making tools to identify patients at risk are regularly used in clinical practice. Thus, clinicians usually rely on their 'gut feeling' to estimate a patient's risk. Espejo et al.[67] found that for an experienced clinician, clinical gestalt can be a reliable indicator. In our study, the AUROC of a sepsis diagnosis based on the clinician's clinical gestalt as a predictor for mortality was 0.73 (95% CI: 0.67, 0.79). Though this performance is only moderate, it raises the question, whether a clinical tool to simplify clinical decision making is even needed and would perform better than the clinician's 'gut feeling'. However, even if an experienced clinician might not need a decision-making tool, for those with limited clinical experience, clinical scoring systems can provide helpful support in decision-making.

Furthermore, the implementation of a clinical score to estimate a patient's sepsis or mortality risk on a regular basis, may lead to a more thorough review of the vital signs taken. This consideration of sepsis and further attention to a patient's vital signs, may already lead to faster treatment and improved outcome. The study team actively asking the clinician about their 'gut feeling' may also have already introduced a bias, as the clinicians were reminded to consider sepsis in every patient with infection.

4.2 Original Score performance analysis

Given the considerations above, the question remains which score would be most appropriate for regular use in our setting. In this prospective cohort study of 851 patients presenting at the emergency units of two Ethiopian hospitals with suspected infection, we evaluated the ability of SIRS, MEWS, mNEWS, qSOFA, and UVA to predict in-hospital mortality.

The univariate logistic regression models confirmed the significant association of each score with in-hospital mortality. Differences in effect sizes reflect differences in the strength of association, but need to be interpreted with caution, as the ranges of the

scores vary. While qSOFA showed the highest increase in odds of death, it likely reflects the narrow scale of qSOFA, only ranging from 0-3. Scores with higher ranges are expected to show lower increase in odds of deaths per score point increase.

We reported the performances of the scores using the AUROCs of the scores, the AUROCs of the scores adjusted for baseline risk, and a sensitivity and specificity analysis at specific score cutoff values. Depending on the context, each of these methods can be useful to evaluate the scores. The AUROCs of the scores are useful single values to assess score performance[33]. Adding baseline risk to the evaluated scores can be useful when assessing score performance above baseline risk, especially, when comparing their performance across populations with different baseline risks or identifying subgroups where the scores perform best[33,68]. However, adding baseline risk to score evaluation can confound the score performance, as the results reflect the combined effect of the score and the baseline risk[33]. To identify patients at risk in a clinical setting, sensitivity and specificity analysis are most important for the identification of an optimal cutoff value for a decision making process[33].

4.2.1 AUROCs for mortality prediction using the evaluated scores

When comparing our results to those found in the literature, it is important to bear in mind that the mNEWS used in this study is an adapted version of the original NEWS and does not include a supplemental oxygen component. Comparisons of the mNEWS with performance of the original NEWS need to be interpreted cautiously.

In our cohort, SIRS and qSOFA predicted mortality poorly, with AUROCs of 0.63 (95% CI: 0.55-0.71) and 0.67 (95% CI: 0.59-0.74), respectively. MEWS (AUROC: 0.7; 95% CI: 0.61-0.79) and mNEWS (AUROC: 0.72; 95% CI: 0.63-0.81) had a higher ability to predict mortality compared to SIRS and qSOFA. mNEWS performed significantly better than SIRS, while the difference in performance between mNEWS and qSOFA was not significant. Differences between MEWS, SIRS, mNEWS, and qSOFA were also not significant. UVA (AUROC: 0.87; 95% CI: 0.82-0.92) significantly outperformed all other scores in our cohort.

This aligns with findings from other low-resource settings[33,69–71]. A prospective cohort study with 2797 participants conducted in Lao People's Democratic Republic, Malawi, Mozambique, and Zimbabwe showed that the UVA score performed better than MEWS and qSOFA in predicting mortality among febrile patients[69]. In this study, the AUROC (95% CI) for UVA, MEWS, and qSOFA were 0.82 (0.79–0.85), 0.67 (0.63–0.71), and 0.68 (0.64–0.72), respectively.

In a study conducted by Klinger et al.[33] in Rwanda, 647 patients were included with suspected infection at hospital admission. In this setting, the UVA (AUROC: 0.71; 95% CI: 0.66, 0.76) also outperformed adapted MEWS (AUROC: 0.69; 95% CI: 0.64, 0.74) and qSOFA (AUROC: 0.65; 95% CI: 0.60, 0.70).

The superiority of UVA to predict mortality was also found in a study conducted in Gabon[70]. The AUROC of UVA was 0.90 (95% CI: 0.78-1.0), outperforming qSOFA (AUROC: 0.77; 95% CI: 0.63,0.91), MEWS (AUROC: 0.72; 95% CI: 0.58, 0.87), and SIRS (AUROC: 0.70; 95% CI: 0.52–0.88).

Bonewell et al. [71] additionally evaluated the performance of the NEWS. They conducted their study in Northern Tanzania, including 597 patients at hospital admission with febrile infection. In this study UVA also had the highest discriminatory ability, with an AUROC of 0.85 (95% CI: 0.80, 0.90) The AUROCS of NEWS, MEWS, and qSOFA were 0.81 (95% CI: 0.75, 0.87), 0.75 (95% CI: 0.69, 0.82), and 0.63 (95% CI: 0.56, 0.71), respectively.

In all of these studies, UVA demonstrates superiority in predicting mortality in patients presenting to an emergency unit with suspected infection. However, the studies also show that even among assumably similar cohorts in low-resource emergency unit settings in Sub-Saharan Africa, the scores mentioned above perform differently across cohorts, e.g., the AUROC of the UVA ranged from 0.71 to 0.90[33,69–71].

In contrast to this, the UVA performed poorly in a Malawian setting[72]. In this study, 499 ICU patients were included. The AUROC for in-hospital mortality for the UVA score was only 0.54 (95% CI: 0.48, 0.60) while the Malawi Intensive Care Mortality Evaluation, a score particularly developed for this setting, showed a stronger discriminatory ability, with an AUROC of 0.72 (95% CI: 0.66, 0.78).

This shows the importance of contextualizing the use of clinical tools. While UVA seems to be a useful tool for identifying patients at risk at hospital admission in Sub-Saharan Africa, a different score might perform better in low-resource ICU settings[24]. It also demonstrates the potential advantage of a score developed for a specific population over a score derived from a different population.

Furthermore, we aimed to compare the performances of the evaluated scores in our setting to their performances in a high-income setting. There is limited data on the performance of the UVA in high-income settings. However, a study by Blair et al.[13] compares the performance of SIRS, MEWS, NEWS, qSOFA, and UVA in Cambodia, Ghana, and the United States. In this study, the pooled AUROCs for mortality prediction in a low-resource setting for SIRS, MEWS, NEWS, qSOFA, and UVA was 0.61, 0.66, 0.70, 0.71, and 0.76, respectively. Yet, in the cohort from the United States, NEWS and qSOFA performed slightly better than UVA (AUROC NEWS: 0.71 and qSOFA 0.71 vs. UVA: 0.70), demonstrating the superior predictive performance of UVA in a low-income compared to a high-income setting.

In a large study conducted in the United States, including 445,073 patients, during their non-ICU hospital stay NEWS performed best in predicting in-hospital mortality (AUROC: 0.77, 95% CI: 0.76, 0.79), followed by MEWS (AUROC: 0.73, 95% CI: 0.71, 0.74)[25]. The qSOFA showed an AUROC of 0.69 (95% CI: 0.67, 0.70), and the AUROC of SIRS was 0.65 (95% CI: 0.63, 0.66)[25]. Generally, these results showed modest performance for SIRS, MEWS, mNEWS, and qSOFA, similar to those found in our cohort.

4.2.2 AUROCs for mortality prediction using the evaluated scores adjusted for baseline risk

While SIRS and qSOFA added to the baseline risk model did not outperform the baseline risk model, MEWS, mNEWS, and UVA showed significantly higher AUROCs for mortality prediction when added to the baseline risk model, than the baseline risk model itself in our cohort. This suggests that MEWS, mNEWS, and UVA provide independent prognostic value above the baseline risk model which included age, sex, and HIV status.

These findings support the consideration of implementing MEWS, NEWS, or UVA into clinical practice at our research sites.

4.2.3 Sensitivity and specificity analysis of the original scores

To identify patients with a need of a higher level of care, sensitivity and specificity analysis are helpful to determine the optimal cutoff value of a score[33]. The sensitivity and specificity of a score at a certain cutoff value are important to assess the suitability of the particular score for decision making in that setting[33].

When looking for the optimal cutoff value of a screening tool for mortality, one may argue that sensitivity is of greater importance than specificity, because failing to identify a true case with a high risk of death may have lethal consequences. However, in settings with scarce resources, high specificity becomes equally important. Using a tool with a low specificity may lead to the provision of limited critical care resources (e.g., fluids, antibiotics, ICU beds) to patients who are not actually at risk and deprive critically ill patients of the measures they urgently need[26,33,70]. Thus, the less resources are available, the more important a high specificity and a high PPV of a score becomes.

The sensitivity and specificity of SIRS and qSOFA have been evaluated widely, but due to their poor specificity, they are not an ideal screening tool for sepsis[4]. Churpek et al.[73] found that almost half of the patients in the wards develop two out of four SIRS criteria during their stay. In the context of an ICU, over 90% of the patients showed more than two SIRS criteria at some point of their stay[45]. In our study, the SIRS criteria also only had a specificity of 18% at their common cutoff value of ≥2. The other scores were more specific than SIRS, confirming existing literature with cohorts from low-resource, as well as high-resource settings[24,74–76].

While qSOFA was more specific than SIRS for mortality prediction, it was less sensitive, which is also consistent with previous research in both contexts[24,73–76]. Due to the low sensitivity of the qSOFA, the Surviving Sepsis Campaign (SSC) Guidelines for Management of Sepsis and Septic Shock 2021, recommend against the use of qSOFA alone as a screening tool for sepsis or septic shock[4].

In a large study conducted in the United States, including 30,677 participants, the sensitivity and specificity of SIRS, qSOFA, MEWS, and NEWS at commonly used cutoff thresholds were assessed[25]. It confirmed the results of SIRS and qSOFA mentioned above (qSOFA≥2: sensitivity: 0.69, specificity: 0.64 vs. SIRS≥2: sensitivity: 0.94, specificity: 0.12). MEWS ≥5 demonstrated a slightly higher sensitivity (0.71) and a specificity (0.65) than qSOFA, while NEWS ≥7 showed a sensitivity of 0.87 and a specificity of 0.48. When compared to our findings the sensitivity and specificity of NEWS were quite similar to the results from mNEWS, MEWS showed similar sensitivity but lower specificity in our cohort.

Blair et al.[77] also evaluated the sensitivity and specificity of the scores at different cutoff values, in both low- and high-income settings (Ghana and Cambodia vs. the United States). In the high-income cohort, qSOFA at a cutoff value of ≥2 demonstrated a sensitivity of 0.60 and a specificity of 0.72, outperforming UVA at a cutoff value of ≥2, which showed a sensitivity of 0.60 but a lower specificity of 0.58. On the other hand, Blair et al.[13] further showed that, in the low-income cohort, UVA demonstrated a higher sensitivity at the cutoff value of ≥2 than the qSOFA (UVA: 0.75 vs. qSOFA: 0.54). However, the specificity at the same cutoff value was higher for qSOFA than for UVA (qSOFA: 0.84, UVA: 0.74). NEWS ≥5, MEWS ≥4, SIRS ≥2 showed moderately high to high sensitivities but low specificities.

In our analysis, the UVA ≥2 also demonstrated the highest sensitivity in combination with an acceptable specificity, consistent with findings from other low-income cohorts[12,69–71]. Reported sensitivities and specificities of the other evaluated scores vary across previous studies in low-income settings[12,69–71].

In conclusion, AUROCs, sensitivity, and specificity of a score differ across cohorts, pointing out the importance to evaluate a score in different contexts, prior to widespread implementation[13]. Scores derived from cohorts within similar settings may perform better[13]. The decision which scores to be implemented in a specific setting should be data-driven. A context-specific weighing of sensitivity and specificity is necessary to find optimal score thresholds[33]. As in our cohorts, the UVA performed best among the scores evaluated in this study, and also performed significantly better than the clinicians

'gut feeling', the implementation of the UVA to evaluate mortality risk in infectious patients and decide who needs an increased level of clinical care at our research sites should be considered.

4.2.4 Why does score performance vary in different contexts?

Scores are developed and validated in a specific setting. Applying them to a different context, i.e., from an ICU setting to an ED setting, or from a high-income to a low-income setting, or even between distinct low-income settings, may reduce the validity of the score[13,78]. But how does our population differ from the one the scores were derived from? In the following part, possible key differences that may reduce score validity will be discussed.

Demographics

Studies from low-resource settings have shown that patients presenting at the emergency department with signs of infection tend to be comparatively younger than in high-income countries[13,24,25]. The patients in our cohort were also quite young, with a median age of 28 years. This is not surprising, considering the median age of the Ethiopian population being approximately 19 years in 2024[79]. This is more than 26 years lower than the median age in Germany and 19 years lower than in the United States in 2024[80,81]. Although sepsis mortality rates are higher in Ethiopia than in HICs, younger patients generally have a higher chance to survive sepsis, due to fewer comorbidities and faster recovery rates[1,82]. Aging increases the risk of septic shock, as comorbidities, drug use are more common, risk of malnutrition and colonization with gram negative bacteria are increased, financial resources and endocrine function decreased in elderly patients[82]. Therefore, differences in age among different populations may have an influence on score performance.

Comorbidities

Comorbidity patterns differ not only across age groups but also across regions and socioeconomic status[83–85]. In LMICs, several studies indicate that important mortality-defining comorbidities in septic patients include HIV, tuberculosis, as well as malnutrition[86–88], while in HICs, cardiovascular diseases, chronic kidney disease, cancer,

and various forms of immunodeficiency may be more important risk factors for a poor outcome[89–91].

Chronic infectious diseases among sepsis patients are associated with poorer outcomes[86,87]. The UVA score is the only of the evaluated scores that takes the HIV status of the patient into account. HIV positive patients have an increased risk for developing sepsis and are more likely to have a poor outcome[86], making HIV status an important parameter for risk stratification, as well as outcome prediction and may partly explain the good performance of UVA in a low-resource setting.

The considerable regional variation in HIV prevalence may contribute to further heterogeneity in sepsis presentation and prognosis[92]. In our cohort, 5.4% of patients self-reported, or were reported by their companion, as HIV positive. However, we tested 10.2% of the study participants to be positive for HIV. This difference possibly shows an underdiagnosis of HIV in the area. Fear of stigma may prevent patients from reporting their known HIV-positive status, especially when family members or friends who do not know about the patient's HIV status are around for the hospital admission. Individuals providing collateral history on behalf of the patient may not be informed about the patient's HIV status. 4 out of 5 patients disclose their HIV status to family members in Ethiopia[93]. The main reasons of people hiding their HIV status are fear of stigma and fear of losing family support[94]. Addressing this barrier is essential to enable earlier HIV diagnosis and treatment, and hence reduce sepsis risk. Increased discloser of HIV status, both within family and to healthcare workers, could not only improve clinical care, but also enhance accuracy of sepsis screening tools and thus lead to better management of HIV positive patients.

Other infectious comorbidities that are more common in LMICs than in HICs, such as tuberculosis, hepatitis B/C, and parasitic diseases, may also impact susceptibility to additional infections by altering or weakening the immune response or causing open wounds that can serve as entry points for bacterial infections[95–98]. In our study, 20 (2.3%) participants reported to be infected with tuberculosis (open or latent), which is several times higher than the prevalence reported by Addis Alene et al.[99] in the general

population, indicating an increased sepsis risk for patients with chronic infectious diseases.

Malaria prevalence was low in our cohort, and there was no significant difference in malaria positivity between survivors and non-survivors. Adults living in malaria regions are likely to have developed partial immunity[100]. In children, however, sepsis may be unrecognized because of the assumption that the child is solely suffering from malaria[100]. Malaria, does not only mimic sepsis, it also increases the susceptibility of a patient to a bacterial infection and thus increases the sepsis risk[101]. Because malaria transmission varies with altitude, malaria risk was low for the majority of study participants, who were living above malaria-endemic areas[102–104].

Neglected tropical diseases are widespread in Ethiopia, with trachoma, podoconiosis, leishmaniasis, onchocerciasis, lymphatic filariasis, schistosomiasis, soil-transmitted helminths, and scabies being most prevalent[105]. Lesions caused by these can also lead to secondary bacterial infections[96–98]. While newer data is lacking, Ethiopia was estimated to bear one of the highest total amounts of neglected tropical disease cases in Sub-Saharan Africa in 2012[106].

Moreover, a medical condition that influences mortality in a low-resource setting is undernutrition. Undernutrition weakens the body's immune system and thus, its ability to fight infections, making undernourished individuals more susceptible to poor outcomes[107,108]. In Ethiopia, chronic energy deficiency affects approximately 28.7% of adults[109]. While patterns of malnutrition vary globally, both under- as well as overnutrition may occur in all socioeconomic contexts[108]. In our cohort, 7.7% of the patients were estimated by our study nurses to be severely underweight, while 0.7% were estimated to be severely overweight. Overnutrition poses many health risks, one being its association with higher risk of infections[108]. The impact of overnutrition on sepsis outcomes is complex and literature on this topic somewhat contradictory. There is clinical evidence of an association of obesity with improved short-term and long-term sepsis survival[110]. However, other evidence indicates that obesity is associated with higher rates of acute respiratory distress syndrome and prolonged hospital stay in obese septic patients[111].

Non-communicable diseases like cardiovascular diseases, cancer, and diabetes mellitus may influence sepsis survival and score performance in both, low- and high-resource settings, yet likely to differing degrees[89–91,112–114]. Patients with septic shock and comorbid conditions like hypertension, obstructive coronary artery disease, peripheral artery disease, and arterial fibrillation have higher in-hospital mortality rates[89].

Cancer patients have a higher risk to develop sepsis or septic shock due to the impact of their cancer on their immunity. Although malignancies can also contribute to worse sepsis outcomes, immunotherapy may also mitigate the immune response[90,91].

Data on the impact of diabetes on sepsis outcomes is inconclusive. There is evidence from HICs supporting comparable survival rates between patients with diabetes mellitus and non-diabetic patients, while others report lower survival rates in diabetic patients[112–114]. They also found that diabetic sepsis patients have a higher risk of acute renal failure[113,114]. Although there is a lack of data from low-resource settings on the impact of diabetes as a comorbidity in sepsis patients, limited access to renal replacement therapy is likely to enhance mortality of diabetic septic patients in low-resource settings[115].

Pathogens

A retrospective cohort study conducted by Ohnuma et al.[116], using data from 201 hospitals across the United States, showed that the most common pathogens identified in blood cultures from patients with blood culture-positive sepsis were *Escherichia coli, Staphylococcus aureus, Streptococcus* species, and *Klebsiella* species.

While data on the epidemiology and frequency of pathogens causing sepsis in LMIC, remain limited, LMICs, particularly those in tropical regions, face broader ranges of infectious diseases, as mentioned in the section on comorbidities[86,105]. In addition to pathogens common in non-tropical regions, region-specific and low-resource-associated pathogens contribute to sepsis cases. These include infections caused by protozoa, such as *Plasmodium* species, tuberculosis, and viral infections that cause haemorrhagic fever, e.g. dengue[86,105,117,118]. Due to limited vaccination coverage, infection caused by *Streptococcus pneumoniae*, *Haemophilus influenzae type b* and measles virus are more frequent than in regions with high vaccination rates, potentially contributing more to the burden of sepsis, particularly in children and the elderly, than in countries with high

vaccination coverage[119,120]. Among HIV positive patients *Cryptococcus neoformans* infections are a common complication[120]. In our study cohort *Cryptococcus neoformans* was the most common cause of mortality in HIV positive patients[121]. A systematic review conducted in Sub-Saharan Africa showed that the most common pathogenic organisms in aerobic blood cultures of patients with bloodstream infections were *Mycobacterium tuberculosis*, non-typhoidal *Salmonellae* species, and *Streptococcus pneumoniae.* Malaria was less common than bacterial bloodstream infections[86].

The pathogens that are highly prevalent in LMICs, but rare in HICs, may vary markedly in their pathophysiological mechanism leading to acute organ dysfunction and presentation compared to classical bacterial sepsis[24,61,68]. Tropical diseases are often understudied and are not appropriately reflected as causes of sepsis in the current sepsis-3 definition. Whereas the current definition focuses mainly on the host response to infections, the potential direct damage through pathogens themselves or their products, is not included in this definition[8,61]. Regional variations in the prevalence of infectious diseases may partially explain differences in score performance across various low-resource settings.

Moreover, heterogeneity in antibiotic resistance patterns may further explain differences in sepsis outcomes and score performances. Resistance rates were estimated to be highest in Sub-Saharan Africa[122]. This poses a challenge to sepsis management and influences sepsis incidence and survival. Few therapeutic options for sepsis caused by resistant pathogens and different resistance patterns may raise mortality rates and contribute to performance heterogeneity of mortality predicting scores[122,123].

Health care infrastructure and clinical practice

For a mortality risk score to be clinically useful across different settings, both the required diagnostic equipment as well as comparable treatment options ensuring similar patient outcomes need to be available[13]. For example, when a score is developed in a setting where mechanical ventilation is available, a patient with respiratory failure may receive a lower score, as effective treatment options leading to enhanced survival chances are available. The same patient would have a higher mortality risk if this

treatment option were not in place. Thus, using the score developed in a high-resource setting may disguise actual risk if certain treatment options are not available.

A study conducted by Kifle et al.[115] in 51 out of 53 ICUs in Ethiopia reported that a total of 324 ICU beds was serving a population of 114 million people. That is approximately 0.3 beds per 100,000 people. In contrast, in Germany and the United States, the ICU capacity for 100,000 people was 30 and 31, respectively[124–126].

Kifle et al.[115] also reported that only 33% of the ICU beds were equipped with all of the following three non-invasive monitoring devices: pulse oximetry, electrocardiography, and blood pressure monitor. Ventilation was available for 58% of the beds, and renal dialysis for 4.8%. Strategies to prevent infection were often not available[115]. Additionally, a major challenge in critical care medicine in Sub-Saharan Africa is the shortage of essential supplies. Especially the lack of disposables and medications (i.e., oxygen, vasopressors, sedatives, and antibiotics) are the greatest challenge[127]. Regional differences in available therapeutic options and sepsis prevention strategies may result in regional variation of mortality rates[13].

Furthermore, it is important to note that treatment strategies developed in HIC also need to be evaluated in low-income settings before wide implementation[128]. For the first hour after recognition of sepsis, the SSC recommends a Hour-1 bundle including core interventions (e.g. lactate measurement, obtaining blood cultures, administration of antibiotics, fluids, and vasopressors as needed)[129]. Sasmito et al.[130] have assessed what difficulties arise when implementing the SSC Hour-1 Bundle in a middle-income country setting. They reported that lack of knowledge about the SSC Hour-1 Bundle, cost issues, lack of blood culture and lactate testing facilities, as well as lack of coordination among health workers, were the main challenges in implementation.

Even if the implementation of a sepsis management protocol is feasible, the impact must be evaluated[128]. According to the Surviving Sepsis Campaign Guidelines patients with sepsis-induced hypoperfusion or septic shock should be given at least 30 mL/kg (ideal body weight) intravenous crystalloid fluid within the first three hours of rescucitation[131]. In a retrospective cohort study conducted by Kuttab et al.[132], involving 509 patients in the United States, failure to carry out this intervention was associated

with an increased odds of in-hospital mortality, delayed resolution of hypotension, and increased length of ICU stay, irrespective of comorbidities. However, in low-resource settings, higher mortality rates were reported in adults and children when applying a protocol including early resuscitation with administration of intravenous fluid and vasopressors, compared to routine care[128,133,134]. In these cohorts, patients may have been particularly vulnerable to fluid overload, which potentially results in pulmonary oedema and respiratory failure. If ventilator support is not available, fluid administration should therefore be carried out with caution[133]. This highlights that interventions that have been shown to be effective in high-income countries may not be advantageous in limited-resource settings, when critical supportive measures may not be available.

Additionally, the availability of healthcare facilities varies across low- and high-income settings, further limiting timely treatment. Hendrix et al.[135] found, that in Ethiopia, approximately 18% of the population has to walk longer than two hours to reach a public health care facility. Regional disparities are huge, with best coverage in Addis Abeba, Dire Dawa, and the Harari Region[135]. In areas with lower health care access, like the Somali Region, 86% of the population is not able to reach a public health centre within one hour walking, but accessibility improves when motorized transport is available and used, leaving 56% of the population unable to reach a public health centre within one hour[136]. Patient referral and delays within the hospital further prolongs time to treatment[137,138]. Hence, treatment delays due to long travel times remain a significant challenge in Ethiopia, especially in rural areas. In contrast, a study conducted in the United States showed that the average travel time to an emergency department was 17.3 minutes by car[139]. Delayed presentation due to long travel times may result in further disease progression. Thus, patients may present at the EOPD in a more severe condition, leading to lower chances of survival.

Physiological vital signs

It is widely known that populations adapt themselves to their living conditions. Moore et al.[140] explains that this context, the term “adaptation” can have two different meanings. On the one hand, it can refer to any trait that improves survival and reproduction rates in the specific environment and lead to natural selection, while on

the other hand, it can be used to describe acclimatization processes of individuals moving from one habitat to another. Both forms of adaptation can occur in inhabitants of high altitudes[140].

The lower environmental oxygen levels in those areas are a challenge for adequate organ oxygen supply[140]. Despite this stressor, about 500.3 million people live at ≥1500 m above sea level (a.s.l.), 81.6 million ≥2500 m a.s.l., and 14.4 million ≥3500 m a.s.l., mainly in the Andes, the Tibetan Plateau, and in highlands in Ethiopia[141]. The largest population residing above 1500m a.s.l. and 2500 m a.s.l. lives in Ethiopia, while the largest population residing above 3500m a.s.l. is located in China[141]. These areas have been inhabited for thousands of years, giving the possibility for genetic adaptation to emerge[142–144]. Moore et al.[140] found that permanent residents living at high altitudes ≥2,500 m a.s.l. show O_2-transport characteristics that differ from migrants in those areas and populations living at sea-level. Tibetans, Andeans, and Highland Ethiopians were found to have higher blood flow and distinct circulatory distribution patterns to ensure O_2 supply and increase efficiency of O_2 usage compared to lowlanders[140].

Additionally, deviations in standard values of vital parameters such as heart rate, blood pressure level, and oxygen saturation among populations living at different heights have been reported[145–150]. Generally, the evidence mostly supports the presence of lower blood pressure, elevated pulse rate, and decreased oxygen saturation in inhabitants residing at high altitudes[145,147,148]. Nevertheless, some papers report higher blood pressure in highland compared to lowland populations[149,150].

To evaluate the average values of vital parameters in Ethiopia among populations living at different heights, an observational study has been conducted by Früh et al.[11]. The study was conducted at three different sites: in Asella (2400 m a.s.l.), Adama (1620 m a.s.l.), and Semara (400 m a.s.l.). Vital signs were assessed, and a questionnaire was used to collect data of each of the 612 healthy volunteers included in the statistical analysis. Participants with acute or chronic disease were excluded beforehand. In Asella, the study site at highest altitude, the mean systolic blood pressure was significantly lower compared to the two other study sites (106.5 ± 13 mmHg in Asella, 114.6 ± 15 mmHg in Adama, and 111.1 ±1 mmHg in Semara). At the same site, the mean respiratory rate was

observed to be 22.1 (± 3) breaths per minute. This mean respiratory rate value at high altitude surpassed the qSOFA threshold of 22 breaths per minute. In Adama and Semara, the mean respiratory rate was significantly lower (20.5 ± 3, 21.2 ± 3 breaths/ min in Adama, and Semara, respectively). The authors found a high rate of positive qSOFA scores among the healthy appearing individuals: 28.1% (55/196) in Asella, in 8.3% (19/ 230) in Adama, and 15.1% (28/ 186) in Semara. According to the authors, these unexpected high proportions of participants showing a qSOFA score ≥2 indicate its limited specificity among healthy individuals. In our study in the EOPD among infected patients, the qSOFA positive rate was 66%, with a specificity of only 35%, at the commonly used cutoff value of ≥2.

Früh et al.[11] argue that altitude related deviations in vital signs from standard values, such as higher respiratory rate and lower systolic blood pressure at high altitudes, may contribute to poor performance of the qSOFA score in those settings. In Semara, however, their lowest study site, the presence of positive qSOFA scores among healthy individuals suggests that other factors beyond altitude also influence vital sign distribution. The recalibration of scores within a specific population (e.g., using correcting factors for altitude) may enhance score performance[11].

In summary, clinical scores tend to perform better when applied in settings with characteristics similar to those of the derivation cohort setting[13,78]. The superior performance of UVA in our setting compared to the other evaluated scores is therefore not surprising, considering it was developed using data from Sub-Saharan Africa.

Although many challenges arise with the transferability of scores designed for one population to another, it is of course common practice and contributes to the quality of global health care provision. Needless to say, in HICs there are more financial resources to carry out research activities. Using the gained knowledge in a low-income setting can be beneficial. In fact, the evaluation of data collected in different parts of the world leads to scientific progress. However, it is essential to evaluate the acquired insights in different settings[13,78].

4.3 Score adaptation

Considering the loss in performance when using a score in a new context, the need to adapt a score to specific contexts becomes self-explanatory. Given that our data was highly imbalanced, with a mortality rate of only 4.5% in the complete dataset, we upsampled our data using SMOTENC[66] to generate synthetic samples by interpolation, instead of duplicating existing ones, improving minority class representation. As synthetic samples may not always perfectly reflect real data and can lead to overfitting of models, the findings in the score adaptation analysis need to be interpreted with caution. The performances of the adapted scores were lower in the testing dataset than in the training dataset, indicating potential overfitting and the need for a larger dataset. Therefore, the adapted scores that we have developed are not to be used directly in a clinical context but need to be further evaluated and possibly refined within bigger cohorts.

Our adapted mortality risk prediction models suggest that the adjustment of scores to a specific context can enhance their predictive ability. For four scores, the adjustment led to larger AUROCS: SIRS (0.64 → 0.77), MEWS (0.74 → 0.84), mNEWS (0.76 → 0.85), qSOFA (0.73 → 0.82), while the adaptation of the UVA did not lead to a larger AUROC (0.86 → 0.86). For MEWS, mNEWS, and qSOFA the differences between the AUROCs of the adapted and original score versions were significant. For SIRS, in contrast, significance was not reached, despite a 0.13 increase in AUROC. This may reflect high variability in SIRS performance across subsets of the data, leading to wide confidence intervals. The limited sample size of the evaluation dataset (n = 256) and percentage of positive cases (4.3%) likely contributed to this statistical uncertainty.

Notably, the UVA was the only score that did not show a higher AUROC after adaptation. As the only score developed in a low-resource context, this suggests that the score was already well adapted to our setting.

Although score adaptation to specific contexts may lead to better score performance, it comes at the price of variability of mortality prediction scores. Clinicians may struggle to understand score values assigned to patients who are referred from one region to another. In research, the comparability of scores used in different contexts may pose a

challenge. To address this, standardised risk groupings (e.g., low, medium, high) based on local score distributions may possibly improve compatibility of mortality risk across different regions and prediction models[12].

4.4 Further research and implementation

Further studies are needed to evaluate the original scores in a wider range of low-resource settings. As mentioned above, our adapted scores also need to be further evaluated within the setting in which they were developed in. Adjustments of our models to optimize score performances may be necessary, due to the small sample size. The process of score adaptation to regional contexts should be evaluated in further studies.

In addition to the adaptation of scores from HICs, the development of new context-specific scores should be considered. Large electronic databases would facilitate the development of new scores or the adaptation of existing ones. However, the implementation of electronic databases may be limited by resource constraints.

Given that the UVA score demonstrated the highest performance in a low-resource setting among the evaluated scores, and it being the only of the evaluated scores, that includes HIV as a mortality risk-modifying comorbidity, one may assume that incorporating common mortality-affecting comorbidities would also further enhance score performance in high-income settings. This may be a valuable focus for future research in HICs.

A digital decision support may be useful to calculate scores fast and effortlessly, especially when score adaptation may lead to more complex models. Its feasibility, however, depends on the availability and reliability of digital infrastructure at the given location.

In addition to the development of clinical decision-making tools to identify patients at risk to ensure rapid treatment, the implementation of the developed tools is crucial, but can be challenging and further research on score implementation success as well as challenges, in LMIC, are needed[70]. For the widespread implementation of a new tool into clinical use, clinical scientists, public health leaders, and clinical staff need to work

together. Initial as well as regular clinical trainings to successfully implement the tool into clinical workflows are crucial.

4.5 What else needs to be done?

Even after the implementation of an optimal screening tool, huge challenges will remain in sepsis management. Investments in ICU capacity, laboratory equipment, staff training, as well as treatment options are needed[27,127,151]. In addition, reliable supply chains must be established to ensure consistent access to diagnostics, medication, and supportive care, including oxygen supply[28]. Furthermore, local sepsis management strategies are needed[151].

However, the shortage of health care workers is likely the most critical issue to be addressed. A retrospective cohort study from a resource-limited setting in Uganda reported the impact of a low staff-to-patient ratio on patient status assessment. The median daily monitoring frequency for patients admitted to medical wards was 1.1 for blood pressure and 0.5 for respiratory rate[29]. A better staff-patient ratio would facilitate earlier risk recognition and timely treatment[26].

Although infrastructure, resources, and staff may be limited, in the majority of low-income settings, important interventions known to improve sepsis outcome, such as timely administration of antibiotics and oxygen, careful fluid management, as well as monitoring of treatment response, are possible to implement[152].

4.6 Public health strategies to improve sepsis outcomes

Within the last decade, much has already been done to improve sepsis awareness and clinical management. In 2017, at the 70th World Health Assembly, the World Health Organisation (WHO) committed to action by recognising sepsis as a global health threat. The *WHA70.7* resolution was adopted, emphasizing several key future priorities in order to improve sepsis treatment and infection prevention[153]:

- Raising awareness of sepsis and its possible manifestations;
- Supporting research to understand the burden of sepsis, to improve prevention, early diagnosis, and clinical management of sepsis;

- Strengthening prevention programs and increase vaccine access;
- Increasing access to life-saving supplies and management.

Since, the WHO has, among other measures to improve sepsis outcomes, published the *Global Report on the Epidemiology and Burden of Sepsis* in 2020, published fact sheets, started working on Guidelines on the Management of Sepsis in Adults and Children, continued the global infection prevention campaign *SAVE LIVES: Clean Your Hands*, and provided resources for healthcare worker training on sepsis prevention[151,154]. They also focused on supporting member states in developing and improving national action plans, as well as strengthening collaborations with other organisations and stakeholders[155,156].

To improve sepsis outcomes on a continental, national, and local level, the African Sepsis Alliance was founded in 2016. Its main objectives are to raise awareness and educate about sepsis, as well as to develop and implement a sepsis strategy in Africa[157]. These goals were also recognised in the *Kampala Declaration: Commitment to improve care for sepsis and severely ill patients in Africa*[158], which was published by the African Sepsis Alliance, Federation of Critical Care Nurses, Global Sepsis Alliance, and World Federation of Critical Care Nurses in October 2017 in Kampala, Uganda. It aligns with the WHO goals and focuses on the following goals:

- Improving sepsis prevention and management by expanding access to vaccinations, improving sanitation and nutrition, increasing ICU capacities, and enhancing the frequencies of vital signs monitoring;
- Encouraging countries to develop national sepsis management strategies;
- Promoting sepsis research and development of guidelines for recognition and management of severely ill patients;
- Promoting education on sepsis among health care workers, policy makers, and the public;
- Encouraging cooperation across countries;
- Strengthening collaborations between health care workers, public, and political stakeholders.

To reach these goals, the main focus should lie in the improvement of identification, monitoring, and timely response to patients with severe illness, as well as enhancement of supply and access to oxygen[159]. National actions plan for sepsis that address specific needs of each country are essential and can be realised when national health ministries collaborate with international organisations, such as the Global Sepsis Alliance, African Sepsis Alliance, Africa Centres for Disease Control and Prevention, African Union, and the World Health Organisation[159,160]. The strengthening of health systems will not only save lives in the short term but also benefits the whole population on the long run[159,160].

In 2023/2024, under the leadership of the Global Sepsis Alliance, 70 members and partner organisations developed the *2030 Global Agenda for Sepsis*[161], setting the path for future sepsis collaboration, research, and management. It aims to reduce sepsis incidence by 25%, reduce cost-per-patient by 20%, and improve sepsis survival rates by 20% by 2030.

To enable future progress in sepsis survival, all stakeholders need to work together and ensure that public health policies are put into practice.

4.7 Strengths and limitations of this study

Our study has several strengths and limitations. First, a strength of this study is the prospective study design, which allows for better control over data quality and standardization during data collection.

Second, as this study was conducted at two study sites, there was an increased sample diversity, improving robustness of the results. While this is a strength of this study, it may also be a limitation, as we did not take into account differences between the populations at the two study sites and assumed that one score for both sites is a good choice, even though differences between the two cohorts and clinical measures available may exist (e.g. socioeconomic factors, comorbidities, adaptation to altitude, differing risk factors for developing sepsis etc.). We did, however, control for data collection heterogeneity by rotating the study nurse team between the two sites.

Additionally, several data have been missing and assumed to be in normal range for the score evaluation analysis, as well as imputed for the adaptation of the scores. Although

missing values are likely to occur in clinical practice, they may have affected score performance and score adaptation.

Moreover, 25 (2.9%) patients were included in the study but discharged without the knowledge of our study team. Thus, the end-of-treatment worksheet was not filled. Therefore, we were not able to determine the outcome for these 25 patients. They were retrospectively excluded, which may have introduced a bias.

Another limitation of this study is the sample size. A larger sample size would have been desirable due to the high imbalance of our dataset. This is particularly important for the score adaptation analysis, as a larger dataset would have increased the reliability and accuracy of our models.

We decided to use in-hospital mortality as the primary endpoint, as it is commonly used in low-income setting studies, because of limited post-discharge follow-up and outcomes after discharge that may be influenced by external factors[33,70–72]. This approach was also suggested by Moore et al.[12] for the evaluation of UVA. However, this may have led to a bias, as some patients may have been discharged not because they had recovered, but due to financial constraints or a preference to die at home.

Furthermore, we only created scores with a binary cutoff value. In contrast, NEWS and MEWS use a graded approach, in which different weights are assigned based on how far a measured value deviates from the expected norm. Similarly to SIRS, their approach is bidirectional. Values that surpass or fall short of the normal range get score points added. Using a bidirectional approach to adapt scores to a specific setting could further increase score performance.

In this study, we tested the validity of the scores in predicting in-hospital mortality and adapted those to local circumstances. However, not all of these scores were designed to predict mortality. SIRS, for example, was created to identify systemic inflammatory response that may indicate infection or sepsis, while MEWS and NEWS were developed to identify patients at risk for deterioration. In contrast, qSOFA was created to identify patients at risk of poor outcomes, including death outside the ICU and the UVA was developed to help clinicians categorise patients according to their risk of dying in the hospital. Thus, the varying intentions behind the scores analysed in this study need to

be taken into account when using them to predict patient outcomes such as in-hospital mortality.

Lastly, while this study focused on score evaluation and adaptation of the scores, it did not examine whether the scores could differentiate between various types of infections or types of organ dysfunction.

5 Conclusion

Sepsis and septic shock are life-threatening conditions with particularly high mortality rates in LMICs[1,15]. Early diagnosis and treatment can improve sepsis outcomes and therefore, a prompt identification of patients at risk of a poor outcome is crucial[2–4,31,32]. In resource-scarce settings identifying patients who are most at risk could ensure that limited resources are allocated to those most in need[26,33].

This study aimed to evaluate the performance of tools for identifying patients at risk of in-hospital mortality in a resource-limited setting. Four out of the five evaluated scores were derived from cohorts in HICs and had rarely been evaluated in low-resource settings[5–8]. Only one of the evaluated scores, the UVA, had been developed in a cohort in Sub-Saharan Africa[12]. In our cohort, the UVA score showed the highest performance in in-hospital mortality prediction for patients presenting at the EOPD with infections. This is consistent with previous research from LMIC[13,25,33,69]. Data on UVA performance in high-resource settings is limited, and has not shown superior performance of UVA compared to other clinical risk prediction scores[71]. This shows, that score performance can vary across different populations, which highlights the importance of score validation when using a score in a new context. Furthermore, our results support our hypothesis, that risk prediction scores developed within contextually similar cohorts show higher performance compared to those derived from different settings. The implementation of the UVA into clinical practice at our research sites should be considered.

Our adapted scores for SIRS, MEWS, mNEWS, and qSOFA outperformed their original versions. These improvements were significant for MEWS, mNEWS, and qSOFA, while for SIRS the improvement of performance did not reach significance. However, we were not able to further improve the mortality prediction performance of the UVA, which was designed for a similar cohort and already showed good performance before adaptation, suggesting that it was tailored to the most important characteristics of our cohort. Thus, our findings support our second hypothesis that regional adaptation is feasible and can improve score performance. Nevertheless, the potential score improvement may be very limited when scores were developed in settings similar to the

adaptation context. To validate our findings and refine our models, further research and larger sample sizes are needed.

After identifying or developing the optimal screening tool, implementation into clinical practice is an essential step to improve sepsis outcomes. Health care workers need to be trained, and the results of training, as well as score usage, need to be evaluated consistently[27,153,158]. Additionally, investments into health care capacities, strengthening of supply chains, and most importantly, an increase of health care workforce are needed to improve sepsis outcomes further[26–29,153,158].

Global health policymakers have been recognising the need to improve sepsis outcomes by making it a global health priority[152,153,157]. Only through a united approach and global collaboration we will be able to get closer to the global sepsis alliance's ambitious vision of a world free of sepsis[162].

6 References

1. 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;395:200–11. doi:10.1016/S0140-6736(19)32989-7.

2. Kim H Il, Park S. Sepsis: Early recognition and optimized treatment. *Tuberc Respir Dis (Seoul)*. 2018;82:6–14. doi:10.4046/trd.2018.0041.

3. Seymour CW, Gesten F, Prescott HC, Friedrich ME, Iwashyna TJ, Phillips GS, et al. Time to Treatment and Mortality during Mandated Emergency Care for Sepsis. *N Engl J Med*. 2017;376:2235–44. doi:10.1056/nejmoa1703058.

4. 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. *Intensive Care Med*. 2021;47:1181–247. doi:10.1007/s00134-021-06506-y.

5. Bone RC, Balk RA, Cerra FB, Dellinger RP, Fein AM, Knaus WA, et al. Definitions for Sepsis and Organ Failure and Guidelines for the Use of Innovative Therapies in Sepsis. *Chest*. 1992;101:1644–55. doi:10.1378/chest.101.6.1644.

6. Subbe CPP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. *QJM*. 2001;94:521–6. doi:10.1093/QJMED/94.10.521.

7. Royal College of Physicians. National Early Warning Score (NEWS) National Early Warning Score (NEWS) Report of a working party. vol. 17. 2012.

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;315:801. doi:10.1001/jama.2016.0287.

9. Abraham E. New definitions for sepsis and septic shock continuing evolution but with much still to be done. *JAMA - J Am Med Assoc*. 2016;315:757–9. doi:10.1001/jama.2016.0290.

10. McGloughlin S, Richards GA, Nor MBM, Prayag S, Baker T, Amin P. Sepsis in tropical regions: Report from the task force on tropical diseases by the World Federation of Societies of Intensive and Critical Care Medicine. *J Crit Care*. 2018;46:115–8. doi:10.1016/j.jcrc.2017.12.018.

11. Früh J, Fuchs A, Tufa TB, Früh L, Hurissa Z, Orth HM, et al. Variation of vital signs with potential to influence the performance of qSOFA scoring in the Ethiopian general population at different altitudes of residency: A multisite cross-sectional study. *PLoS One*. 2021;16 doi:10.1371/journal.pone.0245496.

12. Moore CC, Hazard R, Saulters KJ, Ainsworth J, Adakun SA, Amir A, et al. Derivation and validation of a universal vital assessment (UVA) score: a tool for

predicting mortality in adult hospitalised patients in sub-Saharan Africa. *BMJ Glob Heal*. 2017;2:e344. doi:10.1136/bmjgh-2017-000344.

13. Blair PW, Mehta R, Oppong C, Som T, Ko E, Tsalik E, et al. The performance of screening tools for predicting mortality across multi-site international sepsis cohorts. 2022 doi:10.1101/2022.08.27.22279297.

14. Fleischmann C, Scherag A, Adhikari NKJ, Hartog CS, Tsaganos T, Schlattmann P, et al. Assessment of global incidence and mortality of hospital-treated sepsis current estimates and limitations. *Am J Respir Crit Care Med*. 2016;193:259–72. doi:10.1164/rccm.201504-0781OC.

15. Reinhart K, Daniels R, Kissoon N, Machado FR, Schachter RD, Finfer S. Recognizing Sepsis as a Global Health Priority — A WHO Resolution. *N Engl J Med*. 2017;377:414–7. doi:10.1056/nejmp1707170.

16. Riviello ED, Sugira V, Twagirumugabe T. Sepsis research and the poorest of the poor. *Lancet Infect Dis*. 2015;15:501–3. doi:10.1016/S1473-3099(15)70148-9.

17. Fleischmann-Struzek C, Mellhammar L, Rose N, Cassini A, Rudd KE, Schlattmann P, et al. Incidence and mortality of hospital- and ICU-treated sepsis: results from an updated and expanded systematic review and meta-analysis. *Intensive Care Med*. 2020;46:1552–62. doi:10.1007/s00134-020-06151-x.

18. Gaieski DF, Edwards JM, Kallan MJ, Carr BG. Benchmarking the Incidence and Mortality of Severe Sepsis in the United States*. *Crit Care Med*. 2013;41:1167–74. doi:10.1097/CCM.0b013e31827c09f8.

19. Kwizera A, Dünser M, Nakibuuka J. National intensive care unit bed capacity and ICU patient characteristics in a low income country. *BMC Res Notes*. 2012;5:475. doi:10.1186/1756-0500-5-475.

20. Blair PW, Okello S, Wailagala A, Ayebare RR, Olebo DF, Kayiira M, et al. Long-term Mortality Among Hospitalized Adults with Sepsis in Uganda: a Prospective Cohort Study. *MedRxiv*. 2023:2023.09.14.23295526. doi:10.1101/2023.09.14.23295526.

21. Cohen J, Vincent JL, Adhikari NKJ, Machado FR, Angus DC, Calandra T, et al. Sepsis: A roadmap for future research. *Lancet Infect Dis*. 2015;15:581–614. doi:10.1016/S1473-3099(15)70112-X.

22. African Institute for Development Policy (AFIDEP). A Call to Action: Sepsis is Africa's Neglected Silent Killer. Nairobi, Kenya: 2019. [cited 2025 May 29]. Available from: https://www.mendeley.com/catalogue/6c7823de-6ccb-3a77-ae4e-de4c1713c81b/?utm_source=desktop&utm_medium=1.19.8&utm_campaign=open_catalog&userDocumentId=%7B44a32609-50f8-4c9a-a2d5-5a86194050cc%7D.

23. Liang L, Moore B, Soni A. National Inpatient Hospital Costs: The Most Expensive Conditions by Payer, 2017: Statistical Brief #261. Rockville, MD, USA: Agency

for Healthcare Research and Quality (US). 2020. [cited 2020 Oct 3]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32833416.

24. Adegbite BR, Edoa JR, Ndzebe Ndoumba WF, Dimessa Mbadinga LB, Mombo-Ngoma G, Jacob ST, et al. A comparison of different scores for diagnosis and mortality prediction of adults with sepsis in Low-and-Middle-Income Countries: a systematic review and meta-analysis. *EClinicalMedicine*. 2021;42:101184. doi:10.1016/j.eclinm.2021.101184.

25. Churpek MM, Snyder A, Han X, Sokol S, Pettit N, Howell MD, et al. Quick Sepsis-related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores for Detecting Clinical Deterioration in Infected Patients outside the Intensive Care Unit. *Am J Respir Crit Care Med*. 2017;195:906–11. doi:10.1164/rccm.201604-0854OC.

26. Mar Minn M, Aung NM, Kyaw DZ, Zaw TT, Chann PN, Khine HE, et al. The comparative ability of commonly used disease severity scores to predict death or a requirement for ICU care in patients hospitalised with possible sepsis in Yangon, Myanmar. *Int J Infect Dis*. 2021;104:543–50. doi:10.1016/j.ijid.2021.01.047.

27. Rudd KE, Kissoon N, Limmathurotsakul Di, Bory S, Mutahunga B, Seymour CW, et al. The global burden of sepsis: Barriers and potential solutions. *Crit Care*. 2018;22:123305059. doi:10.1186/s13054-018-2157-z.

28. Morton B, Stolbrink M, Kagima W, Rylance J, Mortimer K. The Early Recognition and Management of Sepsis in Sub-Saharan African Adults: A Systematic Review and Meta-Analysis. *Int J Environ Res Public Heal 2018, Vol 15, Page 2017*. 2018;15:2017. doi:10.3390/IJERPH15092017.

29. Asiimwe SB, Okello S, Moore CC. Frequency of Vital Signs Monitoring and its Association with Mortality among Adults with Severe Sepsis Admitted to a General Medical Ward in Uganda. *PLoS One*. 2014;9:e89879. doi:10.1371/journal.pone.0089879.

30. Papali A, McCurdy MT, Calvello EJB. A 'three delays' model for severe sepsis in resource-limited countries. *J Crit Care*. 2015;30:861.e9-861.e14. doi:10.1016/j.jcrc.2015.04.003.

31. Kalil AC, Johnson DW, Lisco SJ, Sun J. Early Goal-Directed Therapy for Sepsis: A Novel Solution for Discordant Survival Outcomes in Clinical Trials. *Crit Care Med*. 2017;45:607–14. doi:10.1097/CCM.0000000000002235.

32. Liu VX, Fielding-Singh V, Greene JD, Baker JM, Iwashyna TJ, Bhattacharya J, et al. The Timing of Early Antibiotics and Hospital Mortality in Sepsis. *Am J Respir Crit Care Med*. 2017;196:856–63. doi:10.1164/rccm.201609-1848OC.

33. Klinger A, Mueller A, Sutherland T, Mpirimbanyi C, Nziyomaze E, Niyomugabo J-P, et al. Predicting mortality in adults with suspected infection in a Rwandan hospital: an evaluation of the adapted MEWS, qSOFA and UVA scores. *BMJ Open*.

2021;11:e040361. doi:10.1136/bmjopen-2020-040361.

34. Rojas-Garcia P, van der Pol S, van Asselt ADI, Postma MJ, Rodríguez-Ibeas R, Juárez-Castelló CA, et al. Diagnostic Testing for Sepsis: A Systematic Review of Economic Evaluations. *Antibiotics*. 2021;11:27. doi:10.3390/antibiotics11010027.

35. Adhikari NKJ, Fowler RA, Bhagwanjee S, Rubenfeld GD. Critical care and the global burden of critical illness in adults. *Lancet*. 2010;375:1339–46. doi:10.1016/S0140-6736(10)60446-1.

36. Fleischmann-Struzek C, Mellhammar L, Rose N, Cassini A, Rudd KE, Schlattmann P, et al. Incidence and mortality of hospital- and ICU-treated sepsis: results from an updated and expanded systematic review and meta-analysis. *Intensive Care Med*. 2020;46:1552–62. doi:10.1007/s00134-020-06151-x.

37. Fleischmann-Struzek C, Rudd K. Challenges of assessing the burden of sepsis. *Medizinische Klin - Intensivmed Und Notfallmedizin*. 2023;118:68–74. doi:10.1007/s00063-023-01088-7.

38. Geroulanos S, Douka ET. Historical perspective of the word 'sepsis'. *Intensive Care Med*. 2006;32:2077. doi:10.1007/s00134-006-0392-2.

39. Stiefel M, Shaner A, Schaefer SD. The Edwin Smith Papyrus: The birth of analytical thinking in medicine and otolaryngology. *Laryngoscope*. 2006;116:182–8. doi:10.1097/01.mlg.0000191461.08542.a3.

40. Hernandez Botero JS, Florian Perez MC. The History of Sepsis from Ancient Egypt to the XIX Century. *Sepsis - An Ongoing Signif Chall*. 2012 doi:10.5772/51484.

41. Balk RA, Bone RC. The Septic Syndrome: Definition and Clinical Implications. *Crit Care Clin*. 1989;5:1–8. doi:10.1016/S0749-0704(18)30447-0.

42. Balk RA. Systemic inflammatory response syndrome (SIRS): Where did it come from and is it still relevant today? *Virulence*. 2014;5:20–6. doi:10.4161/viru.27135.

43. Castillo Lara RA, Soria Ruiz M, Martínez Carrera O. Definitions for Sepsis and Organ Failure. *JAMA J Am Med Assoc*. 1993;270:939. doi:10.1001/jama.1993.03510080043022.

44. Vincent JL, Opal SM, Marshall JC, Tracey KJ. Sepsis definitions: Time for change. *Lancet*. 2013;381:774–5. doi:10.1016/S0140-6736(12)61815-7.

45. Sprung CL, Sakr Y, Vincent JL, Le Gall JR, Reinhart K, Ranieri VM, et al. An evaluation of systemic inflammatory response syndrome signs in the Sepsis Occurrence in Acutely ill Patients (SOAP) study. *Intensive Care Med*. 2006;32:421–7. doi:10.1007/s00134-005-0039-8.

46. Levy MM, Fink MP, Marshall JC, Abraham E, Angus D, Cook D, et al. 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. Crit.

Care Med., vol. 31, 2003, p. 1250–6 doi:10.1097/01.CCM.0000050454.01978.3B.

47. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin*. 2017;67:93–9. doi:10.3322/caac.21388.

48. Marshall JC. The PIRO (predisposition, insult, response, organ dysfunction) model: Toward a staging system for acute illness. *Virulence*. 2014;5:27–35. doi:10.4161/viru.26908.

49. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, et al. Assessment of Clinical Criteria for Sepsis. *JAMA*. 2016 doi:10.1001/jama.2016.0288.

50. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, et al. Assessment of clinical criteria for sepsis for the third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA - J Am Med Assoc*. 2016;315:762–74. doi:10.1001/jama.2016.0288.

51. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med*. 1996;22:707–10. doi:10.1007/BF01709751/METRICS.

52. 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 - J Am Med Assoc*. 2016;315:775–87. doi:10.1001/jama.2016.0289.

53. Singer M. The new sepsis consensus definitions (Sepsis-3): the good, the not-so-bad, and the actually-quite-pretty. *Intensive Care Med*. 2016;42:2027–9. doi:10.1007/s00134-016-4600-4.

54. Antcliffe DB, Burrell A, Boyle AJ, Gordon AC, McAuley DF, Silversides J. Sepsis subphenotypes, theragnostics and personalized sepsis care. *Intensive Care Med*. 2025;51:756–68. doi:10.1007/s00134-025-07873-6.

55. DeMerle KM, Kennedy JN, Chang CCH, Delucchi K, Huang DT, Kravitz MS, et al. Identification of a hyperinflammatory sepsis phenotype using protein biomarker and clinical data in the ProCESS randomized trial. *Sci Rep*. 2024;14:6234. doi:10.1038/s41598-024-55667-5.

56. Kruisselbrink R, Kwizera A, Crowther M, Fox-Robichaud A, O'Shea T, Nakibuuka J, et al. Modified Early Warning Score (MEWS) Identifies Critical Illness among Ward Patients in a Resource Restricted Setting in Kampala, Uganda: A Prospective Observational Study. *PLoS One*. 2016 doi:10.1371/JOURNAL.PONE.0151408.

57. Sridhar S, Schmid A, Biziyaremye F, Hodge S, Patient N, Wilson K. Implementation of a pediatric early warning score to improve communication and nursing empowerment in a rural district hospital in Rwanda. *Glob Heal Sci Pract*. 2020;8:838–45. doi:10.9745/GHSP-D-20-00075.

58. Badr MN, Khalil NS, Mukhtar AM. Effect of national early warning scoring system implementation on cardiopulmonary arrest, unplanned ICU admission, emergency surgery, and acute kidney injury in an emergency hospital, Egypt. *J Multidiscip Healthc*. 2021;14:1431–42. doi:10.2147/JMDH.S312395,.

59. Le Gall JR, Klar J, Lemeshow S, Saulnier F, Alberti C, Artigas A, et al. The logistic organ dysfunction system: A new way to assess organ dysfunction in the intensive care unit. *JAMA*. 1996;276:802–10. doi:10.1001/jama.276.10.802.

60. Fuchs A, Tufa TB, Hörner J, Hurissa Z, Nordmann T, Bosselmann M, et al. Clinical and microbiological characterization of sepsis and evaluation of sepsis scores. *PLoS One*. 2021;16 doi:10.1371/JOURNAL.PONE.0247646.

61. Schultz MJ, Dunser MW, Dondorp AM, Adhikari NKJ, Iyer S, Kwizera A, et al. Current challenges in the management of sepsis in ICUs in resource-poor settings and suggestions for the future. *Intensive Care Med*. 2017;43:612–24. doi:10.1007/s00134-017-4750-z.

62. College of Health Sciences Arsi University. Arsi University Information Leaflet. n.d.. [Internet]. [cited 2020 Oct 12]. Available from: https://arsiun.edu.et/images/arsi uni 2.pdf.

63. Adama Hospital Medical College. About Adama Hospital Medical College. n.d.. [Internet]. [cited 2025 Mar 21]. Available from: https://www.adamahmc.edu.et/.

64. Smith GB, Prytherch DR, Jarvis S, Kovacs C, Meredith P, Schmidt PE, et al. A Comparison of the Ability of the Physiologic Components of Medical Emergency Team Criteria and the U.K. National Early Warning Score to Discriminate Patients at Risk of a Range of Adverse Clinical Outcomes∗. *Crit Care Med*. 2016;44:2171–81. doi:10.1097/CCM.0000000000002000.

65. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics*. 1988;44:837. doi:10.2307/2531595.

66. Jiang K, Lu J, Xia K. A Novel Algorithm for Imbalance Data Classification Based on Genetic Algorithm Improved SMOTE. *Arab J Sci Eng*. 2016;41:3255–66. doi:10.1007/s13369-016-2179-2.

67. Espejo T, Nieves-Ortega R, Amsler L, Riedel HB, Balestra G, Rosin C, et al. Clinical Gestalt to Predict Bacterial Infection and Mortality in Emergency Department Patients: A Prospective Observational Study. *J Gen Intern Med*. 2025 doi:10.1007/s11606-025-09440-7.

68. Rudd KE, Seymour CW, Aluisio AR, Augustin ME, Bagenda DS, Beane A, et al.

Association of the Quick Sequential (Sepsis-Related) Organ Failure Assessment (qSOFA) Score With Excess Hospital Mortality in Adults With Suspected Infection in Low- and Middle-Income Countries. *JAMA*. 2018;319:2202. doi:10.1001/jama.2018.6229.

69. Lal S, Luangraj M, Keddie SH, Ashley EA, Baerenbold O, Bassat Q, et al. Predicting mortality in febrile adults: comparative performance of the MEWS, qSOFA, and UVA scores using prospectively collected data among patients in four health-care sites in sub-Saharan Africa and South-Eastern Asia. *EClinicalMedicine*. 2024;77:102856. doi:10.1016/j.eclinm.2024.102856.

70. Schmedding M, Adegbite BR, Gould S, Beyeme JO, Adegnika AA, Grobusch MP, et al. A Prospective Comparison of Quick Sequential Organ Failure Assessment, Systemic Inflammatory Response Syndrome Criteria, Universal Vital Assessment, and Modified Early Warning Score to Predict Mortality in Patients with Suspected Infection in Gabon. *Am J Trop Med Hyg*. 2019;100:202–8. doi:10.4269/ajtmh.18-0577.

71. Bonnewell JP, Rubach MP, Madut DB, Carugati M, Maze MJ, Kilonzo KG, et al. Performance Assessment of the Universal Vital Assessment Score vs Other Illness Severity Scores for Predicting Risk of In-Hospital Death Among Adult Febrile Inpatients in Northern Tanzania, 2016-2019. *JAMA Netw Open*. 2021;4:e2136398. doi:10.1001/jamanetworkopen.2021.36398.

72. Prin M, Pan S, Kadyaudzu C, Li G, Charles A. ICU Risk Stratification Models Feasible for Use in Sub-Saharan Africa Show Poor Discrimination in Malawi: A Prospective Cohort Study. *World J Surg*. 2019;43:2357–64. doi:10.1007/s00268-019-05078-9.

73. Churpek MM, Zadravecz FJ, Winslow C, Howell MD, Edelson DP. Incidence and Prognostic Value of the Systemic Inflammatory Response Syndrome and Organ Dysfunctions in Ward Patients. *Am J Respir Crit Care Med*. 2015;192:958–64. doi:10.1164/rccm.201502-0275OC.

74. Herwanto V, Shetty A, Nalos M, Chakraborty M, McLean A, Eslick GD, et al. Accuracy of Quick Sequential Organ Failure Assessment Score to Predict Sepsis Mortality in 121 Studies Including 1,716,017 Individuals. *Crit Care Explor*. 2019;1:e0043. doi:10.1097/CCE.0000000000000043.

75. Fernando SM, Tran A, Taljaard M, Cheng W, Rochwerg B, Seely AJE, et al. Prognostic Accuracy of the Quick Sequential Organ Failure Assessment for Mortality in Patients With Suspected Infection. *Ann Intern Med*. 2018;168:266–75. doi:10.7326/M17-2820.

76. Serafim R, Gomes JA, Salluh J, Póvoa P. A Comparison of the Quick-SOFA and Systemic Inflammatory Response Syndrome Criteria for the Diagnosis of Sepsis and Prediction of Mortality. *Chest*. 2018;153:646–55. doi:10.1016/j.chest.2017.12.015.

77. Blair PW, Mehta R, Oppong CK, Tin S, Ko E, Tsalik EL, et al. Screening tools for predicting mortality of adults with suspected sepsis: an international sepsis cohort validation study. *BMJ Open*. 2023;13:e067840. doi:10.1136/bmjopen-2022-067840.

78. Guan G, Lee CMY, Begg S, Crombie A, Mnatzaganian G. The use of early warning system scores in prehospital and emergency department settings to predict clinical deterioration: A systematic review and meta-analysis. *PLoS One*. 2022;17:e0265559. doi:10.1371/journal.pone.0265559.

79. Database Earth. Population of Ethiopia 1950-2024 & Future Projections. 2025. [Internet]. [cited 2025 May 26]. Available from: https://database.earth/population/ethiopia.

80. Database Earth. Median Age of Germany 1950-2025 & Future Projections. 2025. [Internet]. [cited 2025 Jun 28]. Available from: https://database.earth/population/germany/median-age.

81. Database Earth. Median Age of United States of America 1950-2025 & Future Projections. 2025. [Internet]. [cited 2025 Jun 28]. Available from: https://database.earth/population/united-states-of-america/median-age.

82. Nasa P, Juneja D, Singh O. Severe sepsis and septic shock in the elderly: An overview. *World J Crit Care Med*. 2012;1:23. doi:10.5492/WJCCM.V1.I1.23.

83. Kyu HH, Maddison ER, Henry NJ, Mumford JE, Barber R, Shields C, et al. The global burden of tuberculosis: results from the Global Burden of Disease Study 2015. *Lancet Infect Dis*. 2018;18:261–84. doi:10.1016/S1473-3099(17)30703-X.

84. Carter A, Zhang M, Tram KH, Walters MK, Jahagirdar D, Brewer ED, et al. Global, regional, and national burden of HIV/AIDS, 1990–2021, and forecasts to 2050, for 204 countries and territories: the Global Burden of Disease Study 2021. *Lancet HIV*. 2024;11:e807–22. doi:10.1016/S2352-3018(24)00212-1.

85. Chowdhury SR, Chandra Das D, Sunna TC, Beyene J, Hossain A. Global and regional prevalence of multimorbidity in the adult population in community settings: a systematic review and meta-analysis. *EClinicalMedicine*. 2023;57 doi:10.1016/J.ECLINM.2023.101860.

86. Lewis JM, Feasey NA, Rylance J. Aetiology and outcomes of sepsis in adults in sub-Saharan Africa: A systematic review and meta-analysis. *Crit Care*. 2019;23:1–11. doi:doi.org/10.1186/s13054-019-2501-y.

87. Chaka W, Berger C, Huo S, Robertson V, Tachiona C, Magwenzi M, et al. Presentation and outcome of suspected sepsis in a high-HIV burden, high antiretroviral coverage setting. *Int J Infect Dis*. 2020;96:276–83. doi:10.1016/j.ijid.2020.04.004.

88. Rello J, Leblebicioglu H. Sepsis and septic shock in low-income and middle-income countries: need for a different paradigm. *Int J Infect Dis*. 2016;48:120–2.

doi:10.1016/j.ijid.2016.04.017.

89. Arnautovic J, Mazhar A, Souther B, Mikhijan G, Boura J, Huda N. Cardiovascular Factors Associated with Septic Shock Mortality Risks. *Spartan Med Res J*. 2018;3 doi:10.51894/001c.6516.

90. Nazer L, Lopez-Olivo MA, Cuenca JA, Awad W, Brown AR, Abusara A, et al. All-cause mortality in cancer patients treated for sepsis in intensive care units: a systematic review and meta-analysis. *Support Care Cancer*. 2022;30:10099–109. doi:10.1007/s00520-022-07392-w.

91. Nates JL, Pène F, Darmon M, Mokart D, Castro P, David S, et al. Septic shock in the immunocompromised cancer patient: a narrative review. *Crit Care*. 2024;28:1–16. doi:10.1186/s13054-024-05073-0.

92. Carter A, Zhang M, Tram KH, Walters MK, Jahagirdar D, Brewer ED, et al. Global, regional, and national burden of HIV/AIDS, 1990–2021, and forecasts to 2050, for 204 countries and territories: the Global Burden of Disease Study 2021. *Lancet HIV*. 2024;11:e807–22. doi:10.1016/S2352-3018(24)00212-1.

93. Kassaw C, Sisay D, Awulachew E, Endashaw Hareru H. Perceived Stigma and Its Association with Gender and Disclosure Status among People Living with HIV/AIDS and Attending Antiretroviral Therapy Clinics in Ethiopia: A Systematic Review and Meta-Analysis. *AIDS Res Treat*. 2022;2022 doi:10.1155/2022/3246249.

94. Kabriku PC, Ansah EW, Hagan JE. Positive Status Disclosure and Sexual Risk Behavior Changes among People Living with HIV in the Northern Region of Ghana. *Infect Dis Rep*. 2023;15:255–66. doi:10.3390/idr15030026.

95. Lewis JM, Feasey NA, Rylance J. Aetiology and outcomes of sepsis in adults in sub-Saharan Africa: a systematic review and meta-analysis. *Crit Care*. 2019;23:212. doi:10.1186/s13054-019-2501-y.

96. Deribe K, Meribo K, Gebre T, Hailu A, Ali A, Aseffa A, et al. The burden of neglected tropical diseases in Ethiopia, and opportunities for integrated control and elimination. *Parasit Vectors*. 2012;5:240. doi:10.1186/1756-3305-5-240.

97. World Health Organization. Scabies. 2023. [Internet]. [cited 2025 Jul 25]. Available from: https://www.who.int/news-room/fact-sheets/detail/scabies.

98. Costa CHN, Chang KP, Costa DL, Cunha FVM. From Infection to Death: An Overview of the Pathogenesis of Visceral Leishmaniasis. *Pathogens*. 2023;12 doi:10.3390/PATHOGENS12070969,.

99. Alene KA, Python A, Weiss DJ, Elagali A, Wagaw ZA, Kumsa A, et al. Mapping tuberculosis prevalence in Ethiopia using geospatial meta-analysis. *Int J Epidemiol*. 2023;52:1124–36. doi:10.1093/ije/dyad052.

100. Li Q, Liu T, Lv K, Liao F, Wang J, Tu Y, et al. Malaria: past, present, and future. *Signal Transduct Target Ther 2025 101*. 2025;10:1–31. doi:10.1038/s41392-025-

02246-3.

101. Church J, Maitland K. Invasive bacterial co-infection in African children with Plasmodium falciparum malaria: A systematic review. *BMC Med*. 2014;12 doi:10.1186/1741-7015-12-31,.

102. Bishop RA, Litch JA. Malaria at high altitude. *J Travel Med*. 2000;7:157–8. doi:10.2310/7060.2000.00049/2/JTM7-0157.PDF.GIF.

103. Haile M, Lemma H, Weldu Y. Population Movement as a Risk Factor for Malaria Infection in High-Altitude Villages of Tahtay–Maychew District, Tigray, Northern Ethiopia: A Case–Control Study. *Am J Trop Med Hyg*. 2017;97:726. doi:10.4269/AJTMH.17-0129.

104. World Health Organization. Malaria - Ethiopia. Geographical Distribution of Malaria Cases. 2024. [Internet]. [cited 2025 Jul 25]. Available from: https://www.who.int/emergencies/disease-outbreak-news/item/2024-DON542.

105. Semahegn A, Manyazewal T, Getachew E, Fekadu B, Assefa E, Kassa M, et al. Burden of neglected tropical diseases and access to medicine and diagnostics in Ethiopia: a scoping review. *Syst Rev*. 2023;12:140. doi:10.1186/s13643-023-02302-5.

106. Deribe K, Meribo K, Gebre T, Hailu A, Ali A, Aseffa A, et al. The burden of neglected tropical diseases in Ethiopia, and opportunities for integrated control and elimination. *Parasit Vectors*. 2012;5:240. doi:10.1186/1756-3305-5-240.

107. Baek MS, Kwon YS, Kang SS, Shim D, Yoon Y, Kim JH. Association of malnutrition status with 30-day mortality in patients with sepsis using objective nutritional indices: a multicenter retrospective study in South Korea. *Acute Crit Care*. 2024;39:127–37. doi:10.4266/ACC.2023.01613.

108. Asiimwe SB, Amir A, Vittinghoff E, Muzoora CK. Causal impact of malnutrition on mortality among adults hospitalized for medical illness in sub-Saharan Africa: What is the role of severe sepsis? *BMC Nutr*. 2015;1 doi:10.1186/S40795-015-0023-9.

109. Dagne S, Menber Y, Wassihun Y, Dires G, Abera A, Adane S, et al. Chronic Energy Deficiency and Its Determinant Factors among Adults Aged 18-59 Years in Ethiopia: A Cross-Sectional Study. *J Nutr Metab*. 2021;2021 doi:10.1155/2021/8850241.

110. Trivedi V, Bavishi C, Jean R. Impact of obesity on sepsis mortality: A systematic review. *J Crit Care*. 2015;30:518–24. doi:10.1016/j.jcrc.2014.12.007.

111. Gong MN, Bajwa EK, Thompson BT, Christiani DC. Body mass index is associated with the development of acute respiratory distress syndrome. *Thorax*. 2010;65:44–50. doi:10.1136/thx.2009.117572.

112. Du Q, Yin X, Zhao H, Li J, Zhang J. Association between comorbid diabetes mellitus and mortality of patients with sepsis: A meta-analysis. *Int J Diabetes Dev*

Ctries. 2024;44:128–36. doi:10.1007/s13410-023-01225-0.

113. Jiang L, Cheng M. Impact of diabetes mellitus on outcomes of patients with sepsis: an updated systematic review and meta-analysis. *Diabetol Metab Syndr*. 2022;14:39. doi:10.1186/s13098-022-00803-2.

114. Wang Z, Ren J, Wang G, Liu Q, Guo K, Li J. Association Between Diabetes Mellitus and Outcomes of Patients with Sepsis: A Meta-Analysis. *Med Sci Monit*. 2017;23:3546–55. doi:10.12659/MSM.903144.

115. Kifle F, Boru Y, Tamiru HD, Sultan M, Walelign Y, Demelash A, et al. Intensive Care in Sub-Saharan Africa: A National Review of the Service Status in Ethiopia. *Anesth Analg*. 2022;134:930–7. doi:10.1213/ANE.0000000000005799.

116. Ohnuma T, Chihara S, Costin B, Treggiari M, Bartz RR, Raghunathan K, et al. Epidemiology, Resistance Profiles, and Outcomes of Bloodstream Infections in Community-Onset Sepsis in the United States*. *Crit Care Med*. 2023;51:1148–58. doi:10.1097/CCM.0000000000005870.

117. Kiya GT, Mekonnen Z, Melaku T, Tegene E, Gudina EK, Cools P, et al. Prevalence and mortality rate of sepsis among adults admitted to hospitals in sub-Saharan Africa: a systematic review and meta-analysis. *J Hosp Infect*. 2024;144:1–13. doi:10.1016/j.jhin.2023.11.012.

118. Boillat-Blanco N, Mbarack Z, Samaka J, Mlaganile T, Mamin A, Genton B, et al. Prognostic value of quickSOFA as a predictor of 28-day mortality among febrile adult patients presenting to emergency departments in Dar es Salaam, Tanzania. *PLoS One*. 2018;13:e0197982. doi:10.1371/journal.pone.0197982.

119. Chen W, Du M, Deng J, Liu M, Liu J. Global, regional, and national trends of measles burden and its vaccination coverage among children under 5 years old: An updated systematic analysis from the Global Burden of Disease Study 2021. *Int J Infect Dis*. 2025;156:107908. doi:10.1016/j.ijid.2025.107908.

120. Wahl B, O'Brien KL, Greenbaum A, Majumder A, Liu L, Chu Y, et al. Burden of Streptococcus pneumoniae and Haemophilus influenzae type b disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000–15. *Lancet Glob Heal*. 2018;6:e744–57. doi:10.1016/S2214-109X(18)30247-X.

121. Tufa TB, Orth HM, Wienemann T, Jensen B-EO, Mackenzie CR, Boulware DR, et al. Disseminated Cryptococcosis Is a Common Finding among Human Immunodeficiency Virus-Infected Patients with Suspected Sepsis and Is Associated with Higher Mortality Rates. *J Fungi*. 2023;9:836. doi:10.3390/jof9080836.

122. Murray CJL, Ikuta KS, Sharara F, Swetschinski L, Robles Aguilar G, Gray A, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet*. 2022;399:629–55. doi:10.1016/S0140-6736(21)02724-0.

123. Kumar NR, Balraj TA, Kempegowda SN, Prashant A. Multidrug-Resistant Sepsis: A Critical Healthcare Challenge. *Antibiotics*. 2024;13:46. doi:10.3390/ANTIBIOTICS13010046.

124. Statistisches Bundesamt. Bevölkerungsstand: Amtliche Einwohnerzahl Deutschlands 2023 - Statistisches Bundesamt. *Stat Bundesamt*. 2024. [Internet]. [cited 2025 Jul 7]. Available from: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/_inhalt.html.

125. Statistisches Bundesamt. Medical facilities, hospital beds and movement of patient - German Federal Statistical Office. 2024. [Internet]. [cited 2025 Jul 7]. Available from: https://www.destatis.de/EN/Themes/Society-Environment/Health/Hospitals/Tables/gd-hospitals-laender.html.

126. Kempker JA, Stearns E, Peterson EN, Waller LA. U.S. Adult Critical Care Beds Per Capita: A 2021 County-Level Cross-Sectional Study. *Crit Care Explor*. 2023;5:e0868. doi:10.1097/CCE.0000000000000868.

127. Spencer SA, Adipa FE, Baker T, Crawford AM, Dark P, Dula D, et al. A health systems approach to critical care delivery in low-resource settings: a narrative review. *Intensive Care Med*. 2023;49:772–84. doi:10.1007/s00134-023-07136-2.

128. Andrews B, Semler MW, Muchemwa L, Kelly P, Lakhi S, Heimburger DC, et al. Effect of an Early Resuscitation Protocol on In-hospital Mortality Among Adults With Sepsis and Hypotension: A Randomized Clinical Trial. *JAMA*. 2017;318:1233–40. doi:10.1001/JAMA.2017.10913.

129. Levy MM, Evans LE, Rhodes A. The Surviving Sepsis Campaign Bundle: 2018 update. *Intensive Care Med*. 2018;44:925–8. doi:10.1007/s00134-018-5085-0.

130. Sasmito P, Pranata S, Pamungkas RA, Emaliyawati E, Arifani N. Challenges of implementing the hour-1 sepsis bundle: a qualitative study from a secondary hospital in Indonesia. *Acute Crit Care*. 2024;39:545–53. doi:10.4266/acc.2023.01473.

131. 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:E1063–143. doi:10.1097/CCM.0000000000005337.

132. Kuttab HI, Lykins JD, Hughes MD, Wroblewski K, Keast EP, Kukoyi O, et al. Evaluation and Predictors of Fluid Resuscitation in Patients with Severe Sepsis and Septic Shock. *Crit Care Med*. 2019;47:1582–90. doi:10.1097/CCM.0000000000003960.

133. Andrews B, Muchemwa L, Kelly P, Lakhi S, Heimburger DC, Bernard GR. Simplified severe sepsis protocol: A randomized controlled trial of modified early goal-directed therapy in Zambia. *Crit Care Med*. 2014;42:2315–24.

doi:10.1097/CCM.0000000000000541.

134. Maitland K, Kiguli S, Opoka RO, Engoru C, Olupot-Olupot P, Akech SO, et al. Mortality after Fluid Bolus in African Children with Severe Infection. *N Engl J Med*. 2011;364:2483–95. doi:10.1056/nejmoa1101549.

135. Hendrix N, Warkaye S, Tesfaye L, Woldekidan MA, Arja A, Sato R, et al. Estimated travel time and staffing constraints to accessing the Ethiopian health care system: A two-step floating catchment area analysis. *J Glob Health*. 2023;13:04008. doi:10.7189/jogh.13.04008.

136. Hierink F, Oladeji O, Robins A, Muñiz MF, Ayalew Y, Ray N. A geospatial analysis of accessibility and availability to implement the primary healthcare roadmap in Ethiopia. *Commun Med*. 2023;3:140. doi:10.1038/s43856-023-00372-z.

137. Assefa EM, Berhane Y. Delays in emergency obstetric referrals in Addis Ababa hospitals in Ethiopia: a facility-based, cross-sectional study. *BMJ Open*. 2020;10:e033771. doi:10.1136/bmjopen-2019-033771.

138. Kamangira B, Ayele G, Dube P, Melaku K, Vushoma E. Maternal near miss hospitalizations in the Borana Zone, Ethiopia: A facility-based longitudinal cross-sectional study. *J Public Health Res*. 2024;13:22799036241238664. doi:10.1177/22799036241238665.

139. Tolpadi A, Elliott MN, Waxman D, Becker K, Flow-Delwiche E, Lehrman WG, et al. National travel distances for emergency care. *BMC Health Serv Res*. 2022;22:388. doi:10.1186/s12913-022-07743-7.

140. Moore LG. Measuring high-altitude adaptation. *J Appl Physiol*. 2017;123:1371–85. doi:10.1152/japplphysiol.00321.2017.

141. Tremblay JC, Ainslie PN. Global and country-level estimates of human population at high altitude. *Proc Natl Acad Sci*. 2021;118:e2102463118. doi:10.1073/pnas.2102463118.

142. Aldenderfer M. Peopling the Tibetan Plateau: Insights from Archaeology. *High Alt Med Biol*. 2011;12:141–7. doi:10.1089/ham.2010.1094.

143. Rademaker K, Hodgins G, Moore K, Zarrillo S, Miller C, Bromley GRM, et al. Paleoindian settlement of the high-altitude Peruvian Andes. *Science (80-)*. 2014;346:466–9. doi:10.1126/science.1258260.

144. Tymko MM, Tremblay JC, Bailey DM, Green DJ, Ainslie PN. The impact of hypoxaemia on vascular function in lowlanders and high altitude indigenous populations. *J Physiol*. 2019;597:5759–76. doi:10.1113/JP277191.

145. Naeije R. Physiological adaptation of the cardiovascular system to high altitude. *Prog Cardiovasc Dis*. 2010;52:456–66. doi:10.1016/j.pcad.2010.03.004.

146. Caffrey D, Jaime Miranda J, Gilman RH, Davila-Roman VG, Cabrera L, Dowling

R, et al. A cross-sectional study of differences in 6-min walk distance in healthy adults residing at high altitude versus sea level. *Extrem Physiol Med*. 2014;3 doi:10.1186/2046-7648-3-3.

147. Aryal N, Weatherall M, Bhatta YKD, Mann S. Blood Pressure and Hypertension in Adults Permanently Living at High Altitude: A Systematic Review and Meta-Analysis. *High Alt Med Biol*. 2016;17:185–93. doi:10.1089/ham.2015.0118.

148. Shrestha S, Shrestha A, Shrestha S, Bhattarai D. Blood Pressure in Inhabitants of High Altitude of Western Nepal. *J Nepal Med Assoc*. 2012;52:154–8. doi:10.31729/jnma.371.

149. Aryal N, Weatherall M, Bhatta YKD, Mann S. Blood pressure and hypertension in people living at high altitude in Nepal. *Hypertens Res*. 2019;42:284–91. doi:10.1038/s41440-018-0138-x.

150. Brook RD. The Environment and Blood Pressure. *Cardiol Clin*. 2017;35:213–21. doi:10.1016/j.ccl.2016.12.003.

151. World Health Organization. Global report on the epidemiology and burden of sepsis: current evidence, identifying gaps and future directions. Geneva, Switzerland: World Health Organization. 2020. [cited 2025 Apr 30]. Available from: https://www.who.int/publications/i/item/9789240010789.

152. Keeley AJ, Nsutebu E. Improving sepsis care in Africa: an opportunity for change? *Pan Afr Med J*. 2021;40 doi:10.11604/pamj.2021.40.204.30127.

153. Seventieth World Health Assembly Executive Board. Improving the prevention, diagnosis and clinical management of sepsis. Geneva, Switzerland: 2017 Available from: https://apps.who.int/gb/ebwha/pdf_files/WHA70/A70_13-en.pdf.

154. World Health Organization. World Hand Hygiene Day 2019. *World Heal Organ*. 2019. [Internet]. [cited 2025 Jun 30]. Available from: https://www.who.int/campaigns/world-hand-hygiene-day/2019.

155. Seventieth World Health Assambly (World Health Organization). Progress Report - Improving the prevention, diagnosis and clinical management of sepsis. Geneva, Switzerland: 2020 Available from: https://www.alz.co.uk/sites/default/files/pdfs/WAM Campaign Report 2019.pdf.

156. Saito H. Sepsis in Global Health: Current global strategies to fight against sepsis. *Acute Med Surg*. 2025;12:e70045. doi:10.1002/ams2.70045.

157. African Sepsis Alliance. About us. *African Sepsis Alliance*. n.d.. [Internet]. [cited 2025 May 10]. Available from: https://www.africansepsisalliance.org/about.

158. African Sepsis Alliance T, African Federation of Critical Care Nurses T, Global Sepsis Alliance T, World Federation of Critical Care Nurses T. The Kampala Declaration: Commitment to improve care for sepsis and severely ill patients in Africa. *Connect World Crit Care Nurs*. 2017;11 doi:10.1891/1748-6254.11.3.47.

159. Nsutebu E, Rylance J, Appiah JA, Grobusch MP, Williams G, Kissoon N, et al. COVID-19 reinforces the need to improve sepsis care resources in Africa. *Infection*. 2021;49:791–3. doi:10.1007/s15010-020-01571-9.

160. Keeley AJ, Nsutebu E. Improving sepsis care in Africa: an opportunity for change? *Pan Afr Med J*. 2021;40 doi:10.11604/pamj.2021.40.204.30127.

161. Global Sepsis Alliance. 2030 Agenda Changes: Making the Next Success Story in Global Health. Global Sepsis Alliance. 2024 Available from: https://static1.squarespace.com/static/58a7025b8419c215b30b2df3/t/67d1879b8dbfb06f8c3be697/1741784989573/2030_Global_Agenda_for_Sepsis_Online.pdf.

162. Global Sepsis Alliance. Our Vision. n.d.. [Internet]. [cited 2025 Jul 24]. Available from: https://globalsepsisalliance.org/.

Disclaimer: DeepL (https://www.deepl.com/de/) was used in some parts of this thesis to support translation. All text was subsequently reviewed by me and adapted to ensure accuracy and understandability.

Appendix

Suppl. Table 1: Sepsis – 2 Possible signs of systemic inflammation in response to infection

Diagnostic criteria for sepsis Infection* documented or suspected, and some of the following:
General variables
Fever (core temperature >38.3°C)
Hypothermia (core temperature <36°C)
Heart rate >90 beats/min or >2 SD above the normal value for age
Tachypnea
Altered mental status
Significant edema or positive fluid balance (>20 mL/kg over 24 hrs)
Hyperglycemia (plasma glucose >120 mg/dL or 7.7 mmol/L) in the absence of diabetes
Inflammatory variables
Leukocytosis (WBC count >12,000 /μL^{-1})
Leukopenia (WBC count <4000 /μL^{-1})
Normal WBC count with >10% immature forms
Plasma C-reactive protein >2 SD above the normal value
Plasma procalcitonin >2 SD above the normal value
Hemodynamic variables
Arterial hypotension (SBP <90 mmHg, MAP <70, or an SBP decrease >40 mm Hg in adults or <2 SD below normal for age)
SvO_2 >70%
Cardiac index >3.5 $L \cdot min^{-1} \cdot m^{-2}$
Organ dysfunction variables
Arterial hypoxemia (PaO_2/FiO_2 < 300)
Acute oliguria (urine output <0.5mL/kg/hour for at least 2 hours despite adequate fluid resuscitation)
Creatinine increase >0.5mg/dL
Coagulation abnormalities (INR >1.5 or aPTT >60 s)
Ileus (absent bowel sounds)
Thrombocytopenia (platelet count <100,000 µL–1)
Hyperbilirubinemia (plasma total bilirubin >4mg/dL or 70 µmol/L)
Tissue perfusion variables
Hyperlactatemia (>1 mmol/L)
Decreased capillary refill or mottling

Suppl. Table 1: List of possible signs of systemic inflammation in response to infection according to the Sepsis-2 definitions[46]. *Infection defined as a pathologic process induced by a microorganism; WBC = white blood cell; SD = standard deviation; SvO_2 = Mixed venous oxygen saturation; Cardiac index >3.5 $L \cdot min^{-1} \cdot m^{-2}$ = Cardiac output adjusted for body surface area, 3.5 liters of blood per minute per square meter of body surface area; PaO_2 = Partial pressure of oxygen in arterial blood; FiO_2 = Fraction of inspired oxygen.

Informed Consent Form for sepsis II study

Participant CONSENT FORM

Name: _____ **Study-No.:**_____

Statement of person obtaining informed consent:

I have fully explained this research to _____ and have given sufficient information to enable the prospective participant to make an informed decision on whether to participate or not.

NAME: _____

DATE: _____ SIGNATURE: _____

Statement of person giving consent:

I have read the information on the study "**Sepsis: recognition, treatment, causing pathogens and impact of antimicrobial resistance on treatment outcome Arsi Zone, Oromia state, central Ethiopia"** or have had it translated into a language I understand. I have also talked it over with the interviewer to my satisfaction. I understand that my participation is voluntary (optional) and not a requirement for further treatment of my illness. For this study I explicitly allow the investigator's delegate to obtain 2 sets or 40ml blood or appropriate body fluids for the investigation. Additionally, I also understand explicitly that 10ml blood samples taken will be transported abroad or preserved for further investigations but nonetheless every sample obtained will be discarded and not used for any other purpose than consented to within this study. I understand no monetary benefits will arise by participating in this study except get the result for treatment option. I was also informed that HIV test will be done from my blood. I know enough about the purpose, methods, risks and benefits of the research study to judge that I want to take part in it and without being externally forced to do so I consent in participating in this study with my signature below. I understand that I may freely stop being part of this study at any time.

Name: _____ DATE: _____ SIGNATURE / THUMBPRINT: _____

WITNESSES NAME: _____

Suppl. Fig. 1: Participant consent form - English version

Informed Consent Form for sepsis II study (Afan Oromo Version)

Ragaa Uunkaa Hirmaatottaa

Name: _____ **Study-No.:**_____

Jechaa nama ragaa fudhatee:

Ani guutummaatti waa`ee qorannoo kanaa _____ yoo fedha isaa ykn ishee ta`e akka hirmaataniif oddeeffanno gahaa ta`e kenneera.

Maqaa: _____

Guyyaa: _____ Mallattoo: _____

Jechaa namaa ragaa kennuu:

Ani oddeeffannoo waa`ee qorannoo "**Weerarrinsa jarmii(seepsisi): hubachuu, yaalu, jarmoota sababa tahanii fi walbar´insi isaan dawwaa faana qabafi fiirii isaan yaala kennamu irratti qaban Godina Arsii, Naannoo, Oromiyaa, Giddigaleessa Itiyoophiyaatti ilaalu**"" dubbiiseera ykn akkaataa ani hubachuu danda`uutti naaf dubbifameera. Dabalatanis haasaa taasisuudhan, oddeeffannoo quubsa ta`e argadheera. Qorrannoo kana keessatti hirmaachunis feedhii irrati qofaa akka ta`e sirritti hubadheera. Qorannoo kanaafis kan oolu dhiigni hanag 40ml akka naarraa barbaachisu hubadheera. Dabalatas, dhiigni 10ml ta´u narraa fudhateme akka qorannoo dabalataaf taa´u ykn gara biyya Jarmanitti geessamuu akka danda`an natt beeksifameera. Dhiiga fudhatame irraas HIV akka ilaalamu natti himameera. Walumaagalatti ani oddeeffannoo gaha ta`e waa`ee kaayyoo, miidhaa fi faayidaa qorannoo kana dhiibbaa tokko malee hubadhee irrattii hirmaachuuf feedhii qabachukoo mallattookoo armaan gadiittin mirkannessa. Yeroo barbaaddetti dhaabuu akkan danda`us hubadheera.

Maqaa: _____Guyyaa: _____Mallattoo/ashaara qubaa: _____

Maqaa nama ragaa ta`ee: _____

Mallattoo maatii (nama of hinbeekneef ykn umurii waggaa <18fi): _____

Suppl. Fig. 2: Participant consent form - Oromo version

CRF Sepsis II study (SEPRESIST) at ATH – *Inclusion/Baseline*

1. **Patient Name:** _____
2. **Age:** _____
3. **Gender:** ☐ female ☐ male
4. **Address/residence** (home visit if not reached by phone):

   _____

   _____

   _____
5. **Telephone numbers:**

   Tel. number 1: _____ Owner of tel:_____

   Tel. number 2: _____ Owner of tel:_____

   Tel. number 3: _____ Owner of tel:_____
6. **Date and Time of admission** (international):

   _____/_____/_____ (DD/MM/YYYY), time _____:_____
7. **Date and Time of study inclusion** (international):

   _____/_____/_____ (DD/MM/YYYY), time _____:_____
8. **What are your chief complaints?**

   ☐ Neurological symptoms ☐ Yes ☐ No

   (e.g. Headache, stiffness of neck, altered consciousness etc.)

   ☐ Respiratory symptoms (e.g. shortness of breath, cough) ☐ Yes ☐ No

   ☐ Cardiac symptoms (chest pain, palpitations) ☐ Yes ☐ No

   ☐ Abdominal symptoms (e.g. abd. pain, nausea/vomiting, diarrhea) ☐ Yes ☐ No

   ☐ Fever / chills ☐ Yes ☐ No

   ☐ Skin alterations (redness, swelling, pain) ☐ Yes ☐ No

   ☐ other, specify:_____
9. **Diagnosis of sepsis according to clinician?** ☐ Yes ☐ No
10. **Clinical diagnosis at admission according to clinician:**

    ☐ Respiratory tract infection ☐ Urinary tract infection

    ☐ Surgical site infection ☐ Abdominal infection

    ☐ Skin infection/Abscess ☐ Osteomyelitis

    ☐ Meningitis/Encephalitis ☐ Focus unclear

    ☐ Other, specify:_____

Suppl. Fig. 3: Worksheet for baseline data

1. **Do you regularly take any medication?** □ No □ Yes, specify:

   Drug 1_____

   Drug 2_____

   Drug 3_____

   Drug 4_____

   Drug 5_____

2. **Did you see a health professional (doctor/health officer/nurse)?**

   1. in the last week? □ No □ Yes
   2. in the last 3 month? □ No □ Yes

3. **Did you take any drug for your current illness**? □ No □ Yes, specify:_____

   □ Bought informally □ Prescribed by health professional (including pharmacist)

4. **Did you take any antibiotic for your current illness? If yes, what is the duration?**

   □ No □ I don't know □ Yes, specify: □ Yes, duration:

   _____

5. **Have you been inpatient in a hospital in the past 90 days?**

   □ No □ Yes, specify reason:_____

   _____

Do you consume alcohol regularly? (to select yes: >14 units (♂) and >7 units (♀) per week)

□ No □ Yes

7. **Do you regularly smoke cigarettes?** □ No □ Yes
8. **Do you regularly consume khat?** □ No □ Yes
9. **Do you have any chronic diseases?** □ No □ Yes, specify

   □ Diabetes mellitus □ Art. Hypertension □ Asthma / COPD

   □ Heart disease □ Psychiatric disorder □ Neurologic disorder

   □ Other, specify:_____

10. **Do or did you have any chronic infections?** □ No □ Yes, specify below

    □ Tuberculosis if yes: □ Current (treated) □ current (untreated) □ History of TB

    □ Hepatitis

    □ HIV, if yes: on ART □ yes □ no

Suppl. Fig. 4: Worksheet for medical history

CRF Sepsis II study (SEPRESIST) at ATH – *Inclusion/Baseline*

For referred patients only

1. **Data of transferring health facility (thf):**
   - a.) **Name and town of thf:**_____
   - b.) **Date and time of admission at thf:** _____/_____/_____ (DD/MM/YYYY), time _____:_____
   - c.) **Diagnosis at thf:**_____

2. **Treatment at transferring health facility (thf):**
   - a.) **How long did you have symptoms before you came to the thf?**

     _____days
   - b.) **Did you receive treatment at the thf?**

     ☐ Yes ☐ No
   - c.) **Did you receive antibiotics at the thf?**

     ☐ Yes ☐ No ☐ Unsure

     1. **Name:**_____ **Dosage:** _____mg **Application:** ☐i.v. ☐p.o. ☐i.m.
     2. **Name:**_____ **Dosage:** _____mg **Application:** ☐i.v. ☐p.o. ☐i.m.
     3. **Name:**_____ **Dosage:**_____mg **Application:** ☐i.v. ☐p.o. ☐i.m.
   - d.) **Did you receive fluids at the thf?**

     ☐ Yes ☐ No ☐ Unsure
   - e.) **Did you receive other medication at the thf?**

     ☐ Yes ☐ No ☐ Unsure

     4. **Name:**_____ **Dosage:** _____mg **Application:** ☐i.v. ☐p.o. ☐i.m.
     5. **Name:**_____ **Dosage:** _____mg **Application:** ☐i.v. ☐p.o. ☐i.m.
     6. **Name:**_____ **Dosage:**_____mg **Application:** ☐i.v. ☐p.o. ☐i.m.

Suppl. Fig. 5: Worksheet for referral patients at inclusion

CRF Sepsis II study (SEPRESIST) at ATH – *Clinical Assessment - BASELINE*

Name of study nurse: _____

Date of assessment: ____/_____/_____ (DD/MM/YYYY, international calendar)

1. **Constitution by approximation:**
   a. ☐ massive underweight
   b. ☐ regular (moderate over- or underweight)
   c. ☐ massive overweight

2. **Open wounds?** ☐ No ☐ Yes, acute wound ☐ Yes, chronic wound

3. **Vital parameters**

Pulse rate	_____/min	
Resp.rate	_____/min	
BP systolic/diastolic	_____/_____mmHg	
Temp (tympanic)	_____°C	
SpO_2	_____%	
GCS	_____points	
AVPU score	☐ **A**lert	☐ responds to **V**erbal stimuli
	☐ Responds to **P**ainful stimuli	☐ **U**nresponsive
Lactate	_____mmol/l	
Blood glucose	_____mg/dl	

4. **Samples taken:**

a.	EDTA (ca. 5ml)	☐ No	☐ Yes
b.	Serum (ca. 10ml)	☐ No	☐ Yes
c.	Blood cultures (**2 pairs**)	☐ No	☐ Yes
d.	Other samples for microbiology?	☐ No	☐ Yes, specify:_____
			_____

Suppl. Fig. 6: Worksheet for clinical assessment at inclusion

CRF Sepsis II study (SEPRESIST) at ATH – *Clinical Assessment – 24h-Follow-Up*

Name of study nurse: _____

1. **Date of assessment:** ____/_____/_____ (DD/MM/YYYY, international calendar)

2. **Admission status:**

   ☐ Outpatient ☐ Admitted at EOPD ☐ Admitted to ICU

   ☐ Admitted to regular ward (☐ Int.Medicine ☐ Surgery ☐ Gynecology/Obstetrics)

3. **Vital parameters**

Pulse rate	_____/min
Resp.rate	_____/min
BP systolic/diastolic	_____/_____mmHg
Temp (tympanic)	_____℃
SpO_2	_____%
GCS	_____points
AVPU score	☐ **A**lert ☐ responds to **V**erbal stimuli
	☐ Responds to **P**ainful stimuli ☐ **U**nresponsive
Lactate	_____mmol/l
Blood glucose	_____mg/dl

4. **Clinical diagnosis at 24h follow up according to clinician:**

   ☐ Respiratory tract infection ☐ Urinary tract infection

   ☐ Surgical site infection ☐ Abdominal infection

   ☐ Skin infection/Abscess ☐ Osteomyelitis

   ☐ Meningitis/Encephalitis ☐ Focus unclear

   ☐ Other, specify:_____

5. **Fluid management**

   a. Administered within first 4 hours after admission: _____ml

   b. Administered within first 24 hours after admission: _____ml

Suppl. Fig. 7: Worksheet for clinical assessment after 24 hours

CRF Sepsis II study (SEPRESIST) at ATH – *End-of-treatment assessment form*

Study nurse name: _____

1. **Treatment information**

☐ Never admitted (treated as outpatient)

☐ Treated as inpatient on regular ward, never treated on ICU

☐ Int. medicine ☐ Surgery ☐ Gynecology/Obstetrics

☐ Treated (at least part-time) as inpatient on ICU

2. **Hospital discharge**

a. Deceased in hospital?

☐ No ☐ Yes, date of death: ____/____/____ (DD/MM/YYYY), time_____:_____

b. If deceases, reason for death:

☐ Septic shock ☐ Respiratory Failure

☐ Cardiac failure ☐ Kidney failure

☐ Other, specify:_____

c. Date of discharge: Date: _____/_____/2019 time_____:_____

d. Reason for discharge:

☐ cured or well enough to be discharged

☐ financial reasons

☐ hopeless condition

☐ referral, specify: _____

☐ Other, specify: _____

3. **Blood culture results:**

☐ no blood cultures taken

☐ blood cultures taken (2 pairs); specify:

☐ blood cultures sterile after 6 days of culturing

☐ blood cultures positive (____ out of 2 pairs) after _____ hrs of culture time

Specify result: _____

Suppl. Fig. 8: Worksheet end of treatment

Suppl. Table 2: Number and percentage of missing values

Variable	Number of missing values	Percentage of missing values, %
Study site	0	0.0
Age	0	0.0
Sex	0	0.0
Symptoms and Clinical Diagnosis at Admission		
Symptoms at admission		
Fever, chills	0	0.0
Abdominal symptoms	0	0.0
Neurological symptoms	0	0.0
Respiratory symptoms	0	0.0
Cardiac symptoms	0	0.0
Skin alterations	1	0.1
Other symptoms	0	0.0
Clinical diagnosis at admission	2	0.2
Sepsis diagnosis according to treating clinician	37	4.3
Medical history		
On any medication for current illness at admission	5	0.6
On any antibiotic for current illness at admission	96	11.3
Saw a health professional within last 3 months (excluding last week)	29	3.4
Saw a health professional within last week	4	0.5
Was admitted to a hospital within last 90 days	2	0.2
Positive HIV anamnesis	3	0.4
Non-infectious chronic diseases	1	0.1
Infectious chronic diseases	3	0.4
Regular Medication	1	0.1
Substance use		
Alcohol consumption	3	0.4
Khat consumption	2	0.2
Cigarette consumption	1	0.1
Physical constitution		
Open wound	20	2.4
Body weight	22	2.6
Vital signs		
Heart rate	3	0.4
Respiratory rate	0	0.0

Variable	Number of missing values	Percentage of missing values, %
Systolic blood pressure	8	0.9
Diastolic blood pressure	8	0.9
Temperature	1	0.1
Oxygen saturation	62	7.3
Oxygen application	375	44.1
GCS	4	0.5
AVPU	15	1.8
Laboratory results		
Blood culture results	0	0.0
Bloodfilm results	12	1.4
HIV	10	1.2
White Blood Cell count per $10^3/\mu L$	12	1.4
Lymphocyte count per $10^3/\mu L$	30	3.5
Mid-sized cells per $10^3/\mu L$	30	3.5
Granulocyte count per $10^3/\mu L$	29	3.4
Hemoglobin concentration per g/dL	10	1.2
Red Blood Cell count per $10^6/\mu L$	10	1.2
Mean Corpuscular Volume per fL	11	1.3
Platelet count per $10^3/\mu L$	10	1.2
Lactate concentration per mmol/L	245	28.8
Blood glucose concentration per mg/dL	215	25.3
Creatinine concentration per mg/dL	13	1.5
Mortality		
In-hospital mortality	0	0.0
28-days mortality	58	6.8

Suppl. Table 2: Percentage of missing data for each variable. GCS = Glasgow Coma Scale; AVPU = Alert, response to verbal stimuli, response to painful stimuli, unresponsive score; HIV = human immunodeficiency virus.

Suppl. Table 3: Comparison of survivors and non-survivors

Variable	Total (n = 875)	Survivors (n = 836)	Non-Survivors (n = 39)	*p*-value
Study site				0.575
Asella, n (%)	569 (66.9)	542 (66.7)	27 (71.1)	
Adama, n (%)	282 (33.1)	271 (33.3)	11 (29)	
Age, median (IQR)	28 (22–40)	28 (21–40)	33.5 (25–41)	0.063
Male sex, n (%)	456 (53.6)	437 (53.8)	19 (50)	0.65
Symptoms and Clinical Diagnosis at Admission				
Symptoms at admission				
Fever, chills, n (%)	573 (67.3)	554 (68.1)	19 (50)	0.02
Abdominal symptoms, n (%)	380 (44.7)	370 (45.5)	10 (26.3)	0.02
Neurological symptoms, n (%)	366 (43)	344 (42.3)	22 (57.9)	0.058
Respiratory symptoms, n (%)	309 (36.3)	288 (35.4)	21 (55.3)	0.013
Cardiac symptoms, n (%)	90 (10.6)	81 (10)	9 (23.7)	0.007
Skin alterations, n (%)	74 (8.7)	68 (8.4)	6 (15.8)	0.113
Other symptoms, n (%)	48 (5.6)	46 (5.7)	2 (5.3)	>0.999
Clinical diagnosis at admission, n (%)				0.011
Abdominal infection, n (%)	256 (30.2)	251 (31)	5 (13.2)	
Respiratory tract infection, n (%)	216 (25.4)	206 (25.4)	10 (26.3)	
Meningitis/ encephalitis, n (%)	105 (12.4)	94 (11.6)	11 (29)	
Urinary tract infection, n (%)	69 (8.1)	69 (8.5)	0 (0.0)	
Skin infection/ abscess, n (%)	31 (3.7)	30 (3.7)	1 (2.6)	
Infective Endocarditis, n (%)	11 (1.3)	9 (1.1)	2 (5.3)	
Focus unclear, n (%)	9 (1.1)	9 (1.1)	0 (0.0)	
Surgical site infection, n (%)	5 (0.6)	5 (0.6)	0 (0.0)	
Puerperal sepsis, n (%)	4 (0.5)	4 (0.5)	0 (0.0)	
Osteomyelitis, n (%)	1 (0.1)	1 (0.1)	0 (0.0)	
Other, n (%)	142 (16.7)	133 (16.4)	9 (23.7)	
Sepsis diagnosis according to treating clinician, n (%)	342 (42)	310 (39.9)	32 (86.5)	**<0.001**
Medical history				
On any medication for current illness at admission, n (%)	139 (16.4)	132 (16.3)	7 (18.4)	0.735
On any antibiotic for current illness at admission, n (%)	95 (12.6)	90 (12.5)	5 (14.3)	0.793

Variable	Total (n = 875)	Survivors (n = 836)	Non-Survivors (n = 39)	*p*-value
Saw a health professional within last 3 months (excluding last week), n (%)	152 (18.5)	141 (18)	11 (29)	0.089
Saw a health professional within last week, n (%)	240 (28.3)	225 (27.8)	15 (39.5)	0.119
Was admitted to a hospital within last 90 days, n (%)	49 (5.8)	46 (5.7)	3 (7.9)	0.477
Non-infectious chronic diseases				0.216
No known non-infectious chronic disease, n (%)	759 (89.3)	729 (89.8)	30 (79)	
Known non-infectious chronic disease, n (%)	91 (10.7)	83 (10.2)	8 (21.1)	
Heart Disease, n (%)	35 (4.1)	32 (3.9)	3 (7.9)	
Arterial Hypertension, n (%)	17 (2.0)	16 (2)	1 (2.6)	
Diabetes Mellitus, n (%)	17 (2.0)	15 (1.9)	2 (5.3)	
Asthma/COPD, n (%)	15 (1.8)	14 (1.7)	1 (2.6)	
Other, n (%)	4 (0.5)	3 (0.4)	1 (2.6)	
Psychiatric disorder, n (%)	3 (0.4)	3 (0.4)	0 (0.0)	
Infectious chronic diseases				0.022
No known infectious chronic disease, n (%)	779 (91.9)	747 (92.2)	32 (84.2)	
Known infectious chronic disease, n (%)	69 (8.1)	63 (7.8)	6 (15.8)	
HIV, n (%)	46 (5.4)	41 (5.1)	5 (13.2)	
Tuberculosis, n (%)	18 (2.1)	18 (2.2)	0 (0.0)	
Tuberculosis and HIV, n (%)	2 (0.2)	2 (0.2)	0 (0.0)	
Viral hepatitis, n (%)	3 (0.4)	2 (0.2)	1 (2.6)	
Regular medication, n (%)	98 (11.5)	90 (11.1)	8 (21.1)	0.060
Substance use				
Alcohol consumption, n (%)	28 (3.3)	23 (2.8)	5 (13.2)	0.006
Khat consumption, n (%)	44 (5.2)	43 (5.3)	1 (2.6)	0.716
Cigarette consumption, n (%)	23 (2.7)	22 (2.7)	1 (2.6)	>0.999
Physical constitution				
Open wound, n (%)	36 (4.3)	30 (3.8)	6 (15.8)	**<0.001**
Body weight				0.317
Regular, n (%)	759 (91.6)	730 (92.2)	29 (78.4)	
Massive underweight, n (%)	64 (7.7)	57 (7.2)	7 (18.9)	
Massive overweight, n (%)	6 (0.7)	5 (0.6)	1 (2.7)	

Variable	Total (n = 875)	Survivors (n = 836)	Non-Survivors (n = 39)	*p*-value
Vital signs				
Temperature (°C), median (IQR)	37.9 (37–38.6)	37.9 (37–38.6)	38 (36.1–38.6)	0.218
Pulse rate (beats/min), median (IQR)	102.5 (88–118)	102 (88–117.5)	110 (92–125)	0.081
Respiratory rate (breaths/min), median (IQR)	26 (24–31.5)	26 (24–30)	32 (25–36)	**<0.001**
Systolic blood pressure (mmHg), median (IQR)	100 (90–110)	100 (90–110)	90 (80–110)	0.099
Diastolic blood pressure (mmHg), median (IQR)	60 (60–70)	60 (60–70)	60 (50–70)	0.351
Oxygen saturation (%), median (IQR)	95 (92–97)	95 (92–97)	92.5 (89–96)	0.009
Altered consciousness (Glasgow Coma Scale), median (IQR)	15 (15–15)	15 (15–15)	14 (13–15)	**<0.001**
Altered consciousness (AVPU)				0.001
Alert, n (%)	779 (93.2)	757 (94.5)	22 (62.9)	
Responds to verbal stumuli, n (%)	27 (3.2)	24 (3)	6 (17.1)	
Responds to painful stumuli, n (%)	21 (2.5)	15 (1.9)	4 (11.4)	
Unresponsive, n (%)	9 (1.1)	5 (0.6)	3 (8.6)	
Laboratory results				
WBC x10^3/µl (IQR)	8.6 (5.55–12.5)	8.5 (5.5–12.4)	11 (7.2–15.4)	0.037
Lymph x10^3/µl (IQR)	1.3 (0.9–1.9)	1.3 (0.9–1.9)	1 (0.7–1.6)	0.012
Mid x10^3/µl (IQR)	0.6 (0.4–0.9)	0.6 (0.4–0.9)	0.7 (0.4–0.8)	0.632
Gran x10^3/µl (IQR)	6.3 (3.8–9.9)	6.2 (3.8–9.8)	9 (6–13)	0.009
Hgb g/dL (IQR)	13.6 (11.4–15.2)	13.6 (11.4–15.2)	12.4 (11–14.7)	0.14
RBC x10^6/µl (IQR)	4.59 (3.97–5.07)	4.595 (3.97–5.08)	4.3 (3.51–4.94)	0.079
MCV fL (IQR)	91.95 (88.18–95.82)	91.9 (88.1–95.8)	92 (88.9–96.8)	0.517
PLT x10/µl (IQR)	239 (170–318)	240 (174–320)	162 (88–222)	**<0.001**
Creatinine, mg/dl (IQR)	0.6 (0.5–0.8)	0.6 (0.5–0.8)	0.7 (0.6–1.4)	0.002

Variable	Total (n = 875)	Survivors (n = 836)	Non-Survivors (n = 39)	*p*-value
Lactate (IQR)	2.3 (1.5–4.07)	2.3 (1.5–3.7)	4.6 (2.8–7)	**<0.001**
Blood glucose (IQR)	110 (96–128)	110 (96–127)	130 (112–165)	**0.001**
Blood culture results				**<0.001**
No blood culture taken, n (%)	243 (28.6)	242 (29.8)	1 (2.6)	
Sterile, n (%)	524 (61.6)	497 (61.1)	27 (71.1)	
Positive, n (%)	84 (9.9)	74 (9.1)	10 (26.3)	
Bloodfilm results				0.894
Negative, n (%)	826 (98.5)	789 (98.4)	37 (100)	
Positive, n (%)	13 (1.6)	13 (1.6)	0 (0.0)	
P. falciparum, n (%)	6 (0.7)	6 (0.8)	0 (0.0)	
P. vivax/ovale, n (%)	4 (0.5)	4 (0.5)	0 (0.0)	
Spirochetes, n (%)	3 (0.4)	3 (0.4)	0 (0.0)	
HIV, n (%)	86 (10.2)	77 (9.6)	9 (24.3)	0.004
Mortality				
In-hospital mortality, n (%)	38 (4.5)	0 (0.0)	38 (100.0)	-
28-day mortality, n (%)	56 (7.1)	18 (2.4)	38 (100.0)	**<0.001**

Suppl. Table 3: Comparison of Survivors and Non-Survivors. P-values written in bold indicate significance after Bonferroni adjustment for 48 comparisons($p<0.00104$). n = number; IQR = interquartile range, AVPU = Alert, response to verbal stimuli, response to painful stimuli, unresponsive score; COPD = chronic obstructive pulmonary disease; HIV = human immunodeficiency virus; WBC = white blood count; Lymph= lymphocytes; Mid = mid-sized white blood cells; Gran = granulocytes; Hgb = Haemoglobin; RBC = red blood cell count; MCV = mean corpuscular volume, PLT = platelets.

Suppl. Table 4: Logistic regression model estimates

	Coefficient	Standard Error	Odds Ratio (95% CI)	*p*-value
Model 1 - SIRS				
Intercept	-4.472	0.56	-	**<0.001**
SIRS*	0.528	0.19	1.70 (1.18, 2.45)	**0.005**
Model 2 - MEWS				
Intercept	-4.923	0.49	-	**<0.001**
MEWS*	0.334	0.07	1.40 (1.21, 1.61)	**<0.001**
Model 3 - mNEWS				
Intercept	-5.370	0.58	-	**<0.001**
mNEWS*	0.277	0.06	1.32 (1.17, 1.48)	**<0.001**
Model 4 - qSOFA				
Intercept	-5.423	0.65	-	**<0.001**
qSOFA*	1.250	0.30	3.49 (1.93, 6.30)	**<0.001**
Model 5 - UVA				
Intercept	-4.743	0.35	-	**<0.001**
UVA*	0.572	0.07	1.77 (1.54, 2.04)	**<0.001**

Suppl. Table 4: Logistic regression model estimates reporting the odds ratios with 95% confidence intervals for each score point increase. Separate models were built for SIRS, MEWS, mNEWS, qSOFA, and UVA as continuous predictors using the complete dataset (n = 851) to predict in-hospital mortality. P-values written in bold indicate significance. *per one point increase in score points. CI = confidence interval.

Suppl. Table 5: Discrimination comparison of the original scores

Score 1	Score 2	AUROC 1	AUROC 2	|ΔAUROC|	*p*-value
SIRS	MEWS	0.632	0.701	0.069	0.136
SIRS	mNEWS	0.632	0.724	0.092	**0.033**
SIRS	qSOFA	0.632	0.667	0.035	0.505
SIRS	UVA	0.632	0.874	0.242	**<0.001**
MEWS	mNEWS	0.701	0.724	0.023	0.467
MEWS	qSOFA	0.701	0.667	0.034	0.383
MEWS	UVA	0.701	0.874	0.173	**<0.001**
mNEWS	qSOFA	0.724	0.667	0.057	0.094
mNEWS	UVA	0.724	0.874	0.15	**0.001**
qSOFA	UVA	0.667	0.874	0.207	**<0.001**

Suppl. Table 5: Comparison of the area under the receiver operating characteristic curves of the original scores in the complete dataset (n = 851) using DeLong's method[65]. P-values written in bold indicate significance. AUROC = area under the receiver operating characteristic curve.

Suppl. Table 6: Logistic regression model estimates including baseline risk

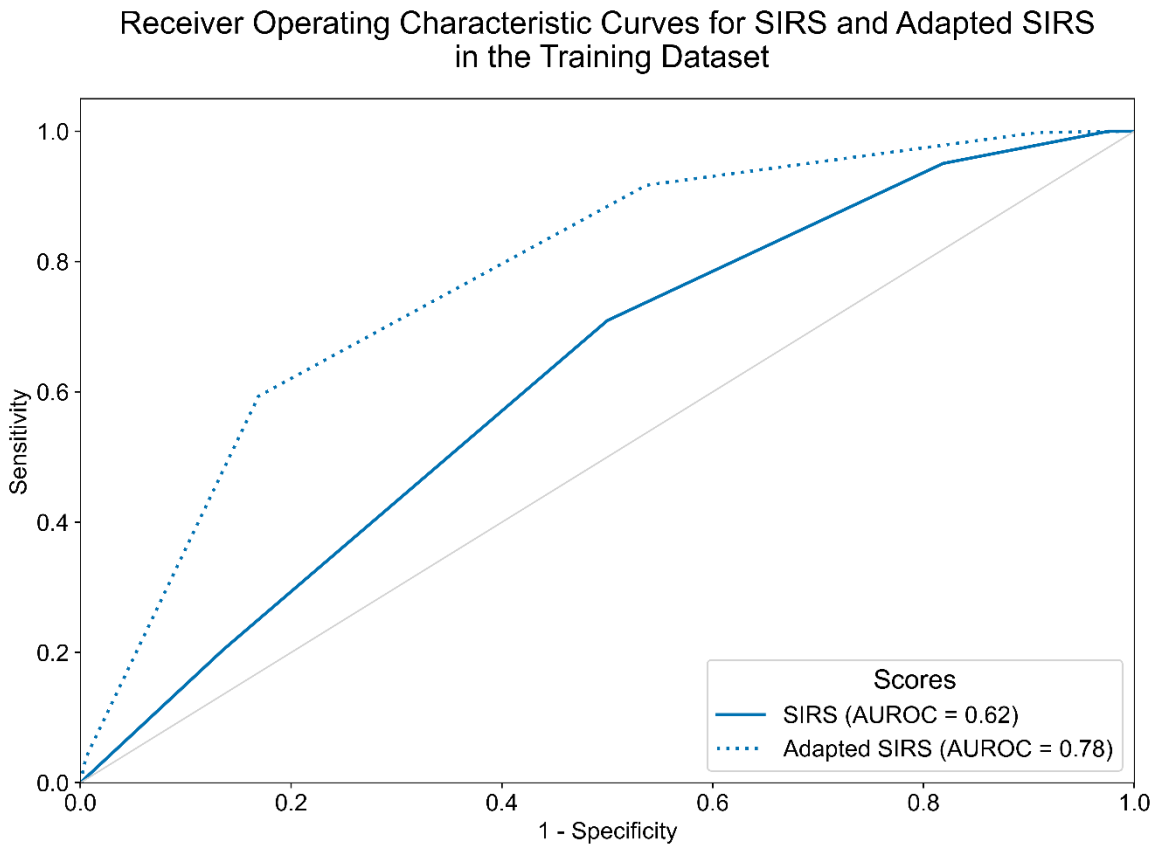
	Coefficient	Standard Error	Odds Ratio (95% CI)	*p*-value
Model 1 - baseline				
Intercept	-3.741	0.44	-	**<0.001**
Age	0.013	0.01	1.01 (1.00, 1.03)	0.167
Sex (male vs. female)	0.097	0.34	1.10 (0.57, 2.13)	0.775
HIV (positive vs. negative)	1.066	0.40	2.90 (1.32, 6.40)	**0.008**
Model 2 - baseline + SIRS				
Intercept	-5.230	0.71	-	**<0.001**
Age	0.015	0.01	1.01 (1.00, 1.04)	0.107
Sex (male vs. female)	0.046	0.34	1.05 (0.54, 2.04)	0.893
HIV (positive vs. negative)	1.054	0.41	2.87 (1.29, 6.37)	**0.010**
SIRS*	0.538	0.19	1.71 (1.19, 2.47)	**0.004**
Model 3 - baseline + MEWS				
Intercept	-5.665	0.65	-	**<0.001**
Age	0.017	0.01	1.02 (1.00, 1.04)	0.081
Sex (male vs. female)	-0.148	0.35	0.86 (0.44, 1.71)	0.671
HIV (positive vs. negative)	1.044	0.41	2.84 (1.26, 6.40)	**0.012**
MEWS*	3.343	0.07	1.41 (1.22, 1.63)	**<0.001**
Model 4 - baseline + mNEWS				
Intercept	-5.756	0.68	-	**<0.001**
Age	0.012	0.01	1.01 (0.99, 1.03)	0.203
Sex (male vs. female)	-0.080	0.35	0.92 (0.47, 1.82)	0.818
HIV (positive vs. negative)	0.856	0.42	2.35 (1.04, 5.33)	**0.040**
mNEWS*	0.260	0.06	1.30 (1.15, 1.46)	**<0.001**
Model 5 - baseline + qSOFA				
Intercept	-5.807	0.74	-	**<0.001**
Age	0.012	0.01	1.01 (0.99, 1.03)	0.234
Sex (male vs. female)	-0.039	0.34	0.96 (0.49, 1.88)	0.910
HIV (positive vs. negative)	0.891	0.41	2.44 (1.09, 5.47)	**0.031**
qSOFA*	1.174	0.30	3.24 (1.79, 5.85)	**<0.001**
Model 6 - baseline + UVA				
Intercept	-4.690	0.52	-	**<0.001**
Age	0.001	0.01	1.00 (0.98, 1.02)	0.939
Sex (male vs. female)	-0.189	0.38	0.83 (0.40, 1.73)	0.616
HIV (positive vs. negative)	-0.199	0.49	0.82 (0.31, 2.14)	0.685
UVA*	0.586	0.08	1.80 (1.56, 2.09)	**<0.001**

Suppl. Table 6: Logistic regression model estimates reporting the odds ratios with 95% confidence intervals for each score point increase. Separate models were built for SIRS, MEWS, mNEWS, qSOFA, and UVA as continuous predictors, each added to a baseline risk model, in the complete dataset (n = 851) to predict in-hospital mortality. The baseline risk model was calculated using age, sex, and HIV status. P-values written in bold indicate significance. *per one point increase in score points. CI = confidence interval. HIV = human immunodeficiency virus.

Suppl. Table 7: Discrimination comparison of the original scores adjusted for baseline risk

Score 1	Score 2	AUROC 1	AUROC 2	|ΔAUROC|	*p*-value
Baseline Risk	Baseline Risk + SIRS	0.637	0.684	0.047	0.225
Baseline Risk	Baseline Risk + MEWS	0.637	0.767	0.13	**0.009**
Baseline Risk	Baseline Risk + mNEWS	0.637	0.754	0.117	**0.024**
Baseline Risk	Baseline Risk + qSOFA	0.637	0.709	0.072	0.134
Baseline Risk	Baseline Risk + UVA	0.637	0.872	0.235	**<0.001**
Baseline Risk + SIRS	Baseline Risk + MEWS	0.684	0.767	0.083	**0.039**
Baseline Risk + SIRS	Baseline Risk + mNEWS	0.684	0.754	0.07	0.101
Baseline Risk + SIRS	Baseline Risk + qSOFA	0.684	0.709	0.025	0.612
Baseline Risk + SIRS	Baseline Risk + UVA	0.684	0.872	0.188	**<0.001**
Baseline Risk + MEWS	Baseline Risk + mNEWS	0.767	0.754	0.013	0.607
Baseline Risk + MEWS	Baseline Risk + qSOFA	0.767	0.709	0.058	0.122
Baseline Risk + MEWS	Baseline Risk + UVA	0.767	0.872	0.105	**<0.001**
Baseline Risk + mNEWS	Baseline Risk + qSOFA	0.754	0.709	0.045	0.152
Baseline Risk + mNEWS	Baseline Risk + UVA	0.754	0.872	0.118	**<0.001**
Baseline Risk + qSOFA	Baseline Risk + UVA	0.709	0.872	0.163	**<0.001**

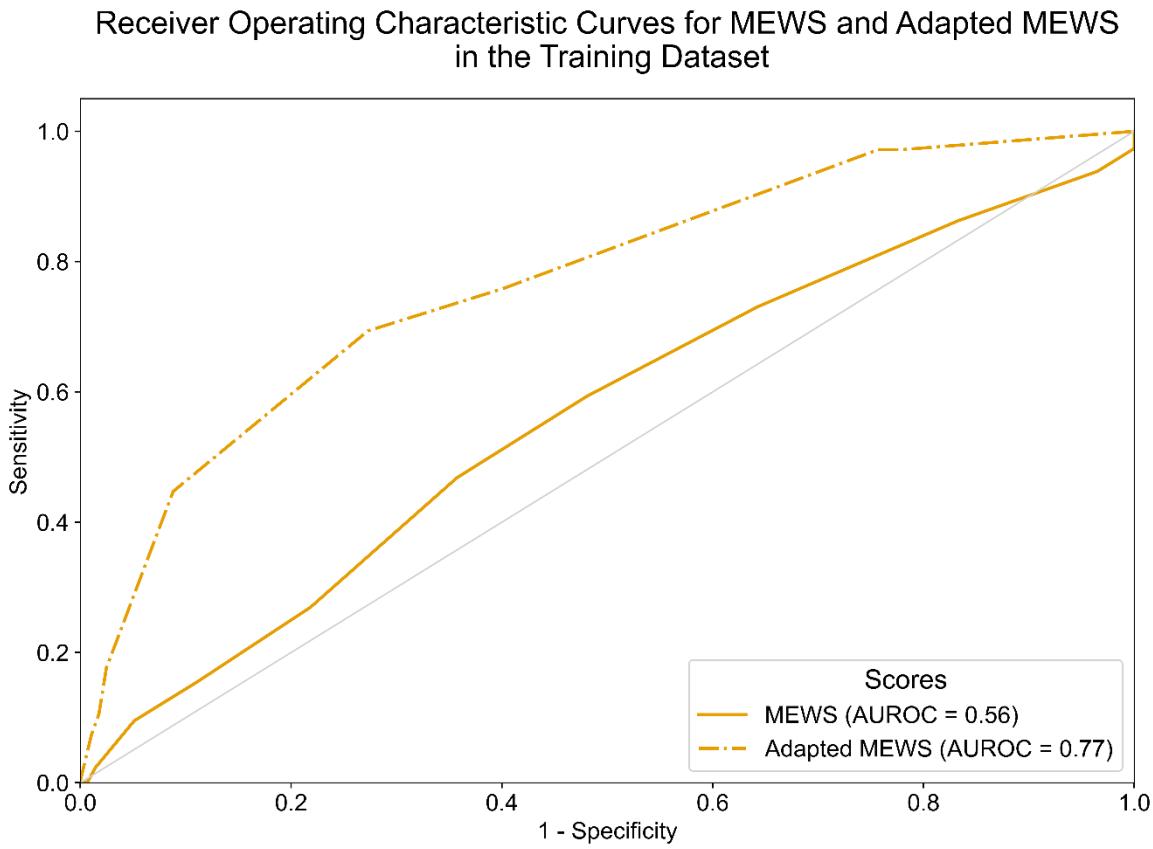
Suppl. Table 7: Comparison of the area under the receiver operating characteristic curves of the original scores, each added to a baseline risk model in the complete dataset (n = 851). The baseline risk model included age, sex, and HIV status. The p-values were calculated using DeLong's method[65]. P-values written in bold indicate significance. AUROC = area under the receiver operating characteristic curve.



Suppl. Fig. 9: Receiver operating characteristic curves for SIRS and adapted SIRS for mortality discrimination in the training dataset

AUROC = Area under the receiver operating characteristic curve.

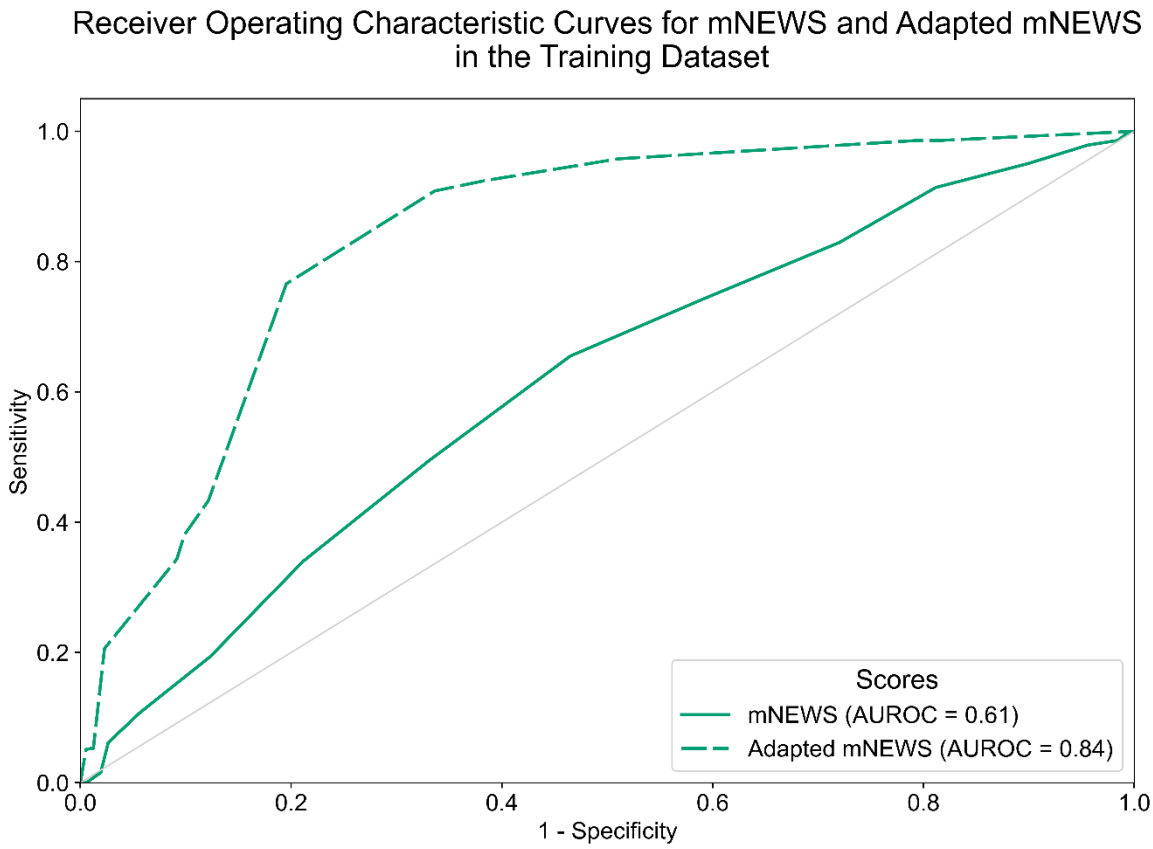
Receiver operating characteristic curves for SIRS and adapted SIRS for in-hospital mortality among patients with suspected infection. The AUROCs were calculated using the training dataset (n = 595) with adapted SIRS derived from the same dataset.



Suppl. Fig. 10: Receiver operating characteristic curves for MEWS and adapted MEWS for mortality discrimination in the training dataset

AUROC = Area under the receiver operating characteristic curve.

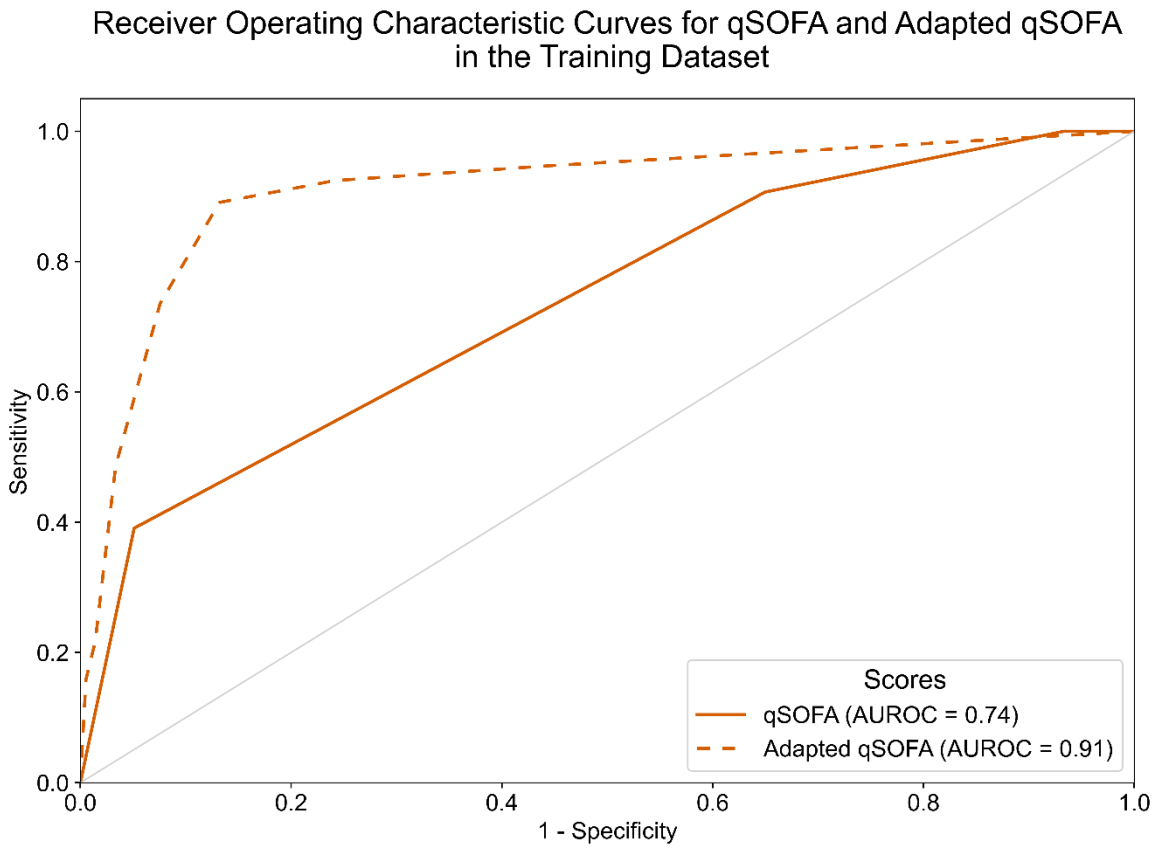
Receiver operating characteristic curves for MEWS and adapted MEWS for in-hospital mortality among patients with suspected infection. The AUROCs were calculated using the training dataset (n = 595) with adapted MEWS derived from the same dataset.



Suppl. Fig. 11: Receiver operating characteristic curves for mNEWS and adapted mNEWS for mortality discrimination in the training dataset

AUROC = Area under the receiver operating characteristic curve.

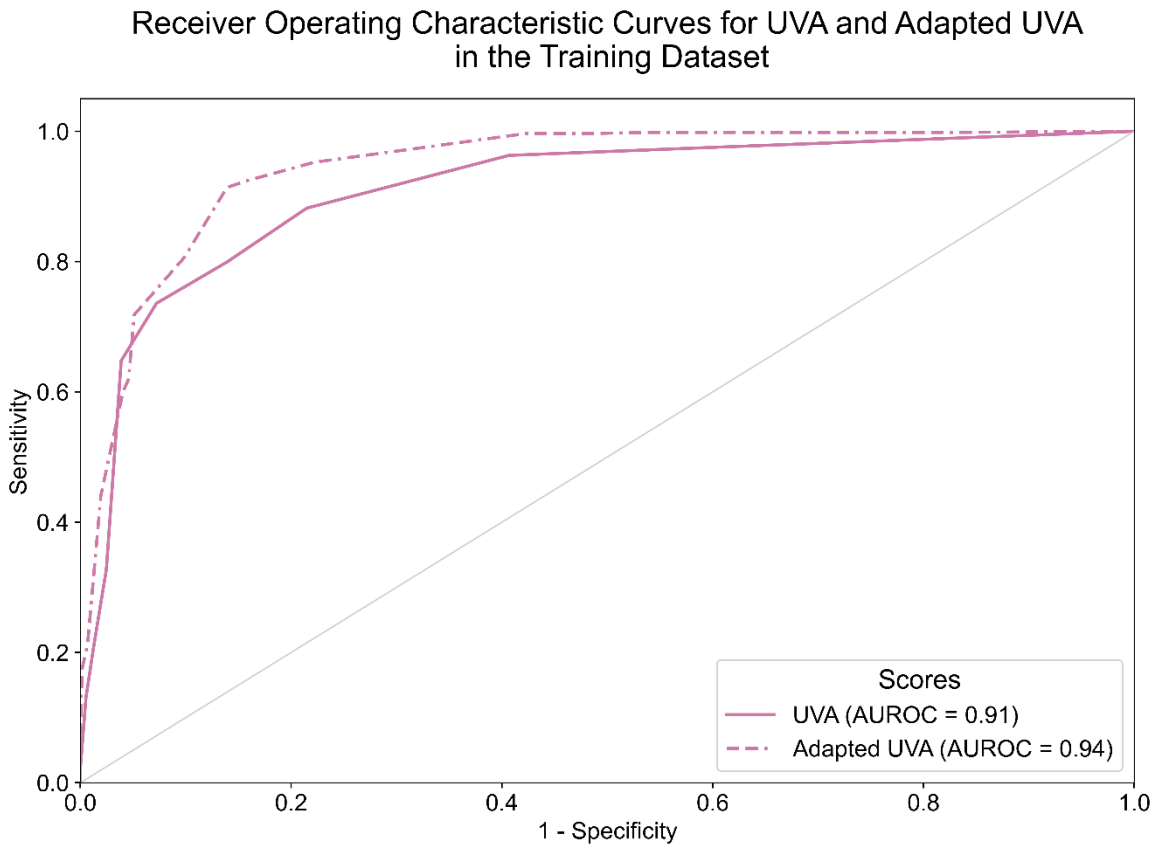
Receiver operating characteristic curves for mNEWS and adapted mNEWS for in-hospital mortality among patients with suspected infection. The AUROCs were calculated using the training dataset (n = 595) with adapted mNEWS derived from the same dataset.



Suppl. Fig. 12: Receiver operating characteristic curves for qSOFA and adapted qSOFA for mortality discrimination in the training dataset

AUROC = Area under the receiver operating characteristic curve.

Receiver operating characteristic curves for qSOFA and adapted qSOFA for in-hospital mortality among patients with suspected infection. The AUROCs were calculated using the training dataset (n = 595) with adapted qSOFA derived from the same dataset.



Suppl. Fig. 13: Receiver operating characteristic curves for UVA and adapted UVA for mortality discrimination in the training dataset

AUROC = Area under the receiver operating characteristic curve.

Receiver operating characteristic curves for UVA and adapted UVA for in-hospital mortality among patients with suspected infection. The AUROCs were calculated using the training dataset (n = 595) with adapted UVA derived from the same dataset.

Suppl. Table 8: Logistic regression model estimates of original and adapted scores in the training dataset

	Coefficient	Standard Error	Odds Ratio (95% CI)	*p*-value
Model 1 - SIRS				
Intercept	-1.435	0.20	-	**<0.001**
SIRS*	0.541	0.07	1.71 (1.50, 1.97)	**<0.001**
Model 2 - MEWS				
Intercept	-0.401	0.14	-	**0.004**
MEWS*	0.083	0.03	1.09 (1.03, 1.14)	**0.002**
Model 3 - mNEWS				
Intercept	-1.057	0.18	-	**<0.001**
mNEWS*	0.137	0.02	1.15 (1.10, 1.20)	**<0.001**
Model 4 - qSOFA				
Intercept	-3.15	0.25	-	**<0.001**
qSOFA*	1.587	0.12	4.89 (3.88, 6.16)	**<0.001**
Model 5 - UVA				
Intercept	-2.684	0.16	-	**<0.001**
UVA*	0.788	0.04	2.20 (2.03, 2.39)	**<0.001**
Model 6 - adapted SIRS				
Intercept	-2.918	0.21	-	**<0.001**
SIRS*	1.381	0.09	3.98 (3.33, 4.76)	**<0.001**
Model 7 - adapted MEWS				
Intercept	-1.913	0.15	-	**<0.001**
MEWS*	0.589	0.04	1.80 (1.66, 1.96)	**<0.001**
Model 8 - adapted mNEWS				
Intercept	-2.468	0.17	-	**<0.001**
mNEWS*	0.479	0.03	1.62 (1.52, 1.71)	**<0.001**
Model 9 - adapted qSOFA				
Intercept	-2.37	0.14	-	**<0.001**
qSOFA*	0.836	0.04	2.31 (2.12, 2.51)	**<0.001**
Model 10 - adapted UVA				
Intercept	-3.739	0.22	-	**<0.001**
UVA*	0.539	0.03	1.71 (1.62, 1.82)	**<0.001**

Suppl. Table 8: Logistic regression model estimates reporting the odds ratios with 95% confidence intervals for each score point increase. Separate models were built for SIRS, MEWS, mNEWS, qSOFA, and UVA, as well as their adapted version, as continuous predictors, using the training dataset (n = 595) to predict in-hospital mortality. P-values written in bold indicate significance. *per one point increase in score points. CI = confidence interval.

Suppl. Table 9: Logistic regression model estimates of original and adapted scores in the testing dataset

	Coefficient	Standard Error	Odds Ratio (95% CI)	*p*-value
Model 1 - SIRS				
Intercept	-4.831	1.12	-	**<0.001**
SIRS*	0.63	0.36	1.88 (0.93, 3.80)	**0.080**
Model 2 - MEWS				
Intercept	-5.537	0.97	-	**<0.001**
MEWS*	0.415	0.14	1.52 (1.16, 1.97)	**0.002**
Model 3 - mNEWS				
Intercept	-6.493	1.27	-	**<0.001**
mNEWS*	0.382	1.12	1.47 (1.55, 1.86)	**0.002**
Model 4 - qSOFA				
Intercept	-6.611	1.34	-	**<0.001**
qSOFA*	1.763	0.59	5.83 (1.85, 18.36)	**0.003**
Model 5 - UVA				
Intercept	-4.662	0.64	-	**<0.001**
UVA*	0.536	0.13	1.71 (1.33, 2.20)	**<0.001**
Model 6 - adapted SIRS				
Intercept	-5.769	1.06	-	**<0.001**
SIRS*	1.257	0.41	3.52 (1.58, 7.82)	**0.002**
Model 7 - adapted MEWS				
Intercept	-6.255	0.97	-	**<0.001**
MEWS*	0.839	0.19	2.31 (1.60, 3.35)	**<0.001**
Model 8 - adapted mNEWS				
Intercept	-5.836	0.90	-	**<0.001**
mNEWS*	0.493	0.11	1,64 (1.31, 2.05)	**<0.001**
Model 9 - adapted qSOFA				
Intercept	-4.17	0.52	-	**<0.001**
qSOFA*	0.492	0.12	1.64 (1.29, 2.08)	**<0.001**
Model 10 - adapted UVA				
Intercept	-4.825	0.66	-	**<0.001**
UVA*	0.304	0.71	1.36 (1.18, 1.56)	**<0.001**

Suppl. Table 9: Logistic regression model estimates reporting the odds ratios with 95% confidence intervals for each score point increase. Separate models were built for SIRS, MEWS, mNEWS, qSOFA, and UVA, as well as their adapted version, as continuous predictors, using the testing dataset (n = 256) to predict in-hospital mortality. P-values written in bold indicate significance. *per one point increase in score points. CI = confidence interval.

Suppl. Table 10: Discrimination comparison of the original and the adapted scores

Score 1	Score 2	AUROC 1	AUROC 2	|ΔAUROC|	*p*-value
SIRS	MEWS	0.640	0.735	0.095	0.317
SIRS	mNEWS	0.640	0.765	0.125	0.174
SIRS	qSOFA	0.640	0.729	0.089	0.402
SIRS	UVA	0.640	0.859	0.219	0.026
SIRS	adapted SIRS	0.640	0.767	0.127	0.061
SIRS	adapted MEWS	0.640	0.838	0.198	**0.024**
SIRS	adapted mNEWS	0.640	0.85	0.21	**0.01**
SIRS	adapted qSOFA	0.640	0.822	0.182	0.104
SIRS	adapted UVA	0.640	0.86	0.22	**0.027**
MEWS	mNEWS	0.735	0.765	0.03	0.346
MEWS	qSOFA	0.735	0.729	0.006	0.93
MEWS	UVA	0.735	0.859	0.124	0.08
MEWS	adapted SIRS	0.735	0.767	0.032	0.653
MEWS	adapted MEWS	0.735	0.838	0.103	**0.017**
MEWS	adapted mNEWS	0.735	0.85	0.115	**0.032**
MEWS	adapted qSOFA	0.735	0.822	0.087	0.289
MEWS	adapted UVA	0.735	0.86	0.125	0.093
mNEWS	qSOFA	0.765	0.729	0.036	0.625
mNEWS	UVA	0.765	0.859	0.094	0.204
mNEWS	adapted SIRS	0.765	0.767	0.002	0.97
mNEWS	adapted MEWS	0.765	0.838	0.073	**0.029**
mNEWS	adapted mNEWS	0.765	0.85	0.085	**0.049**
mNEWS	adapted qSOFA	0.765	0.822	0.057	0.534
mNEWS	adapted UVA	0.765	0.86	0.095	0.225
qSOFA	UVA	0.729	0.859	0.13	**0.004**
qSOFA	adapted SIRS	0.729	0.767	0.038	0.544
qSOFA	adapted MEWS	0.729	0.838	0.109	0.095
qSOFA	adapted mNEWS	0.729	0.85	0.121	**0.029**
qSOFA	adapted qSOFA	0.729	0.822	0.093	**0.025**
qSOFA	adapted UVA	0.729	0.86	0.131	**0.003**
UVA	adapted SIRS	0.859	0.767	0.092	0.076
UVA	adapted MEWS	0.859	0.838	0.021	0.7
UVA	adapted mNEWS	0.859	0.85	0.009	0.836
UVA	adapted qSOFA	0.859	0.822	0.037	0.292
UVA	adapted UVA	0.859	0.86	0.001	0.971
adapted SIRS	adapted MEWS	0.767	0.838	0.071	0.178
adapted SIRS	adapted mNEWS	0.767	0.85	0.083	0.072
adapted SIRS	adapted qSOFA	0.767	0.822	0.055	0.409
adapted SIRS	adapted UVA	0.767	0.86	0.093	0.069
adapted MEWS	adapted mNEWS	0.838	0.85	0.012	0.703
adapted MEWS	adapted qSOFA	0.838	0.822	0.016	0.834
adapted MEWS	adapted UVA	0.838	0.86	0.022	0.712

adapted mNEWS	adapted qSOFA	0.85	0.822	0.028	0.675
adapted mNEWS	adapted UVA	0.85	0.86	0.01	0.84
adapted qSOFA	adapted UVA	0.822	0.86	0.038	0.21

Suppl. Table 10: Comparison of the area under the receiver operating characteristic curves of the original scores and the adapted in the testing dataset (n = 256) using DeLong's method[65]. P-values written in bold indicate significance. AUROC = area under the receiver operating characteristic curve.

Suppl. Table 11: Predictive capacity of original and adapted scores

Score	Sensitivity	Specificity	PPV	NPV
Original Scores				
SIRS ≥2	1	0.16	0.05	1
MEWS ≥3	1	0.17	0.05	1
MEWS ≥4	0.91	0.33	0.06	0.99
MEWS ≥5	0.82	0.49	0.07	0.98
mNEWS ≥5	0.91	0.19	0.05	0.98
mNEWS ≥7	0.91	0.36	0.06	0.99
qSOFA ≥2	1	0.33	0.06	1
UVA ≥2	0.91	0.64	0.1	0.99
UVA ≥4	0.55	0.86	0.15	0.98
Adapted Scores				
Adapted SIRS ≥2	1	0.43	0.07	1
Adapted SIRS ≥3	0.45	0.83	0.11	0.97
Adapted MEWS ≥3	0.82	0.62	0.09	0.99
Adapted MEWS ≥4	0.82	0.73	0.12	0.99
Adapted MEWS ≥5	0.55	0.9	0.2	0.98
Adapted MEWS ≥6	0.55	0.97	0.46	0.98
Adapted mNEWS ≥3	1	0.5	0.08	1
Adapted mNEWS ≥4	0.82	0.62	0.09	0.99
Adapted mNEWS ≥5	0.73	0.67	0.09	0.98
Adapted mNEWS ≥6	0.64	0.8	0.13	0.98
Adapted mNEWS ≥7	0.64	0.89	0.21	0.98
Adapted mNEWS ≥8	0.64	0.92	0.27	0.98
Adapted mNEWS ≥9	0.55	0.93	0.27	0.98
Adapted qSOFA ≥1	1	0.57	0.09	1
Adapted qSOFA ≥2	0.55	0.76	0.09	0.97
Adapted qSOFA ≥3	0.45	0.86	0.13	0.97
Adapted qSOFA ≥4	0.45	0.91	0.19	0.97
Adapted qSOFA ≥5	0.45	0.91	0.19	0.97
Adapted UVA ≥3	1	0.52	0.09	1
Adapted UVA ≥4	0.91	0.58	0.09	0.99
Adapted UVA ≥5	0.82	0.77	0.14	0.99
Adapted UVA ≥6	0.64	0.83	0.15	0.98
Adapted UVA ≥7	0.55	0.85	0.14	0.98
Adapted UVA ≥8	0.45	0.89	0.15	0.97
Adapted UVA ≥9	0.45	0.92	0.2	0.97

Suppl. Table 11: Sensitivity, specificity, positive predictive value, and negative predictive value of the original, as well as adapted, SIRS, MEWS, mNEWS, qSOFA, and UVA. For the original scores common cutoffs were chosen. PPV = positive predictive value; NPV = negative predictive value.

Acknowledgements

First, I would like to thank Prof. Dr. Lüdde, Director of the Department of Gastroenterology, Hepatology, and Infectious Diseases for allowing me to carry out my doctoral research in his department.

Furthermore, I would like to express my sincere gratitude to Prof. Dr. Häusinger for supervising and supporting the beginning of my dissertation, to Prof. Dr. Feldt for taking over and supervising my dissertation after Prof. Dr. Häussinger retired, advising, and encouraging throughout the project, as well as to Prof. Dr. Pfeffer, my second supervisor. I would also like to thank Dr. Fuchs, who also supervised and encouraged me throughout the project.

Additionally, I would like to thank Dr. Tafese, for his supervision and support during my time in Asella. My thanks also go to the other colleagues, I had the pleasure of working with in Asella. Special thanks go to our study team, Kdist, Admas, Daniel, Asegid, Kasim, Kelil, and Tesfa, who collected clinical data, worked in the laboratory, and entered data. I would also like to thank Marlen, Yannik, and Urs, my German colleagues, and Teddy who encouraged me during and after my stay in Ethiopia to do my best. You all gave me a very warm welcome in Ethiopia.

I also want to thank all the study participants for taking the time to answer questionnaires and participating in this study.

Furthermore, I would like to thank my family, who believed in me, always encouraged me, and accompanied me during this time.

My sincere thanks go to everyone I lived with, while working on this thesis, for their emotional support, encouragement, and remarkable tolerance of my mood swings and grumpiness throughout the thesis-writing period.

Above all, I thank my father in heaven for giving me strength to work on this project, and for blessing me with the opportunity to finish this work.