BMJ Open

Screening tools for predicting mortality of adults with suspected sepsis: an international sepsis cohort validation study

Paul W Blair ,[1] Rittal Mehta,[1] Chris Kwaku Oppong,[2] Som Tin,[3] Emily Ko,[4] Ephraim L Tsalik,[4,5] Josh Chenoweth,[1] Michelle Rozo,[1] Nehkonti Adams,[6] Charmagne Beckett,[6] Christopher W Woods,[4,7] Deborah A Striegel,[1] Mark G Salvador,[1] Joost Brandsma,[1] Lauren McKean,[1] Rachael E Mahle,[4] William R Hulsey,[8] Subramaniam Krishnan,[8] Michael Prouty,[9] Andrew Letizia,[10] Anne Fox,[10] Dennis Faix,[9] James V Lawler,[11,12] Chris Duplessis,[13] Michael G Gregory,[13] Te Vantha,[3] Alex Kwame Owusu-Ofori ,[14] Daniel Ansong,[2] George Oduro,[15] Kevin L Schully,[1] Danielle V Clark[1]

To cite: Blair PW, Mehta R, Oppong CK, *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

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-067840).

Received 05 September 2022
Accepted 06 February 2023

For numbered affiliations see end of article.

Correspondence to
Dr Paul W Blair;
PBLAIR6@JHMI.EDU

ABSTRACT

Objectives We evaluated the performance of commonly used sepsis screening tools across prospective sepsis cohorts in the USA, Cambodia and Ghana.

Design Prospective cohort studies.

Setting and participants From 2014 to 2021, participants with two or more SIRS (Systemic Inflammatory Response Syndrome) criteria and suspected infection were enrolled in emergency departments and medical wards at hospitals in Cambodia and Ghana and hospitalised participants with suspected infection were enrolled in the USA. Cox proportional hazards regression was performed, and Harrell's C-statistic calculated to determine 28-day mortality prediction performance of the quick Sequential Organ Failure Assessment (qSOFA) score ≥2, SIRS score ≥3, National Early Warning Score (NEWS) ≥5, Modified Early Warning Score (MEWS) ≥5 or Universal Vital Assessment (UVA) score ≥2. Screening tools were compared with baseline risk (age and sex) with the Wald test.

Results The cohorts included 567 participants (42.9% women) including 187 participants from Kumasi, Ghana, 200 participants from Takeo, Cambodia and 180 participants from Durham, North Carolina in the USA. The pooled mortality was 16.4% at 28 days. The mortality prediction accuracy increased from baseline risk with the MEWS (C-statistic: 0.63, 95% CI 0.58 to 0.68; p=0.002), NEWS (C-statistic: 0.68; 95% CI 0.64 to 0.73; p<0.001), qSOFA (C-statistic: 0.70, 95% CI 0.64 to 0.75; p<0.001), UVA score (C-statistic: 0.73, 95% CI 0.69 to 0.78; p<0.001), but not with SIRS (0.60; 95% CI 0.54 to 0.65; p=0.13). Within individual cohorts, only the UVA score in Ghana performed better than baseline risk (C-statistic: 0.77; 95% CI 0.71 to 0.83; p<0.001).

Conclusions Among the cohorts, MEWS, NEWS, qSOFA and UVA scores performed better than baseline risk, largely driven by accuracy improvements in Ghana, while SIRS scores did not improve prognostication accuracy. Prognostication scores should be validated within the target population prior to clinical use.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This study includes two well-characterised sepsis cohorts in low-income and middle-income countries and a cohort in a high-resource setting for comparison.
⇒ The performance characteristics of five commonly used sepsis screening tools for predicting 28-day death was compared with baseline risk after adjustment for multiple comparisons.
⇒ Diagnostic testing differed at each site and mortality specifically due to sepsis could not be determined.
⇒ Enrolment was by convenience sampling within the referral hospital catchment area and may not be representative of the general population within these countries.
⇒ Sample size limitations in each of the cohorts may have led to decreased ability to identify differences between each screening tool.

INTRODUCTION

Sepsis, a syndrome resulting from a systemic dysregulated host response to an infection, is estimated to cause 6 million deaths per year but is likely an underestimate due to limited information from low/middle-income countries (LMICs) where 87% of the world population live.[1] Despite declining age-standardised incidence and mortality, sepsis remains a major cause of health loss worldwide and has an especially high health-related burden in LMICs.[2]

Clinical sepsis guidelines developed in the Western world may not be applicable in resource-limited settings and moreover can lead to detrimental effects on sepsis care and management when applied in these conditions due to decreased access to resources to manage iatrogenesis from fluid resuscitation.[3 4] In contrast to the USA, pathogens that lead to directly lead to vascular injury are common causes of acute febrile illness in Cambodia and Ghana such as dengue virus, malaria or rickettsia and may alter empiric treatment response.[5] While early recognition and treatment of sepsis is critical, most sepsis scores or early warning systems were derived from cohorts outside of LMICs. Differences in causes of sepsis, available treatments and available resources for supportive care should affect management strategies but evidence is limited and optimal clinical scores or biomarkers for sepsis identification are unknown in these settings. Multisite international sepsis studies are essential for evaluating current and future sepsis tools to ensure effectiveness in resource-limited settings and across populations.

The most validated prognostication scores, SOFA (Sequential Organ Failure Assessment) and the APACHE IV, have been developed for prognostication but require an arterial blood gas and multiple laboratory parameters[6 7] that are not widely available in low-resource settings. The qSOFA (quick SOFA) is an abbreviated score that does not require laboratory parameters. The qSOFA is one of the most widely adopted sepsis screening tools and has largely replaced the SIRS (Systemic Inflammatory Response Syndrome) criteria as the standard abbreviated sepsis screening tool as part of the Sepsis-3 definition to identify septic patients.[6] The qSOFA and other sepsis screening tools (ie, Modified Early Warning Score (MEWS), National Early Warning Score (NEWS) and Universal Vital Assessment (UVA)) are often used clinically to identify those at risk of sepsis, but these tools have been studied for their ability to prognosticate mortality or poor composite outcomes among hospitalised adults.[8–11] Studies have evaluated these tools for predicting in-hospital mortality but the performance of these tools and the prevalence of 28-day mortality, a common metric of sepsis outcomes, have yet to be described across both high-resource and low-resource settings using similar methods.[8 12 13]

We used prospective multi-site international cohorts that are part of the Austere environments Consortium for Enhanced Sepsis Outcomes consortium to validate commonly used sepsis screening tools.[14] In contrast to APACHE IV and SOFA, these tools can be quickly performed with limited laboratory test results. We hypothesised that qSOFA may perform poorly in LMIC populations compared with the UVA score due to differences in causes of sepsis. We describe the diverse clinical characteristics, the aetiologies of suspected sepsis within these cohorts and the performance of sepsis screening tools in current clinical use for predicting mortality at 1-month post-enrolment.

METHODS

From May 2014 to November 2015, 200 participants were enrolled into a prospective observational study of sepsis at Takeo Provincial Hospital in Takeo Province Cambodia[15] (figure 1). This study was followed by a prospective study at Duke University Hospital in Durham, North Carolina, which enrolled 180 participants from December 2014 to March 2016. In Kumasi, Ghana, 187 participants were enrolled at Komfo Anokye Teaching Hospital from July 2016 to October 2017.

Hospitalised patients ≥18 years of age whose attending physician judged them to have an active infection were considered for inclusion for each of the three cohorts. Additional inclusion and exclusion criteria were required in Cambodia and Ghana but not required in the US protocol. In Cambodia and Ghana, participants were required to meet least two clinical criteria for SIRS during screening. In Cambodia and Ghana, patients were excluded if they had known malignancy, chronic renal/hepatic insufficiency, immunosuppressive conditions (except HIV) or systemic steroid usage that exceeded 20 mg/day to prevent confounding in future biomarker studies. Patients were also excluded in Cambodia and Ghana if they had a history of organ transplant, haemodynamically significant gastrointestinal bleeding, anatomic or functional asplenia, acute cardiovascular disease, general anaesthesia or surgery in the past week prior to enrolment, women who were pregnant, patients who had haemoglobin less than 70 g/L or weighed less than 35 kg.

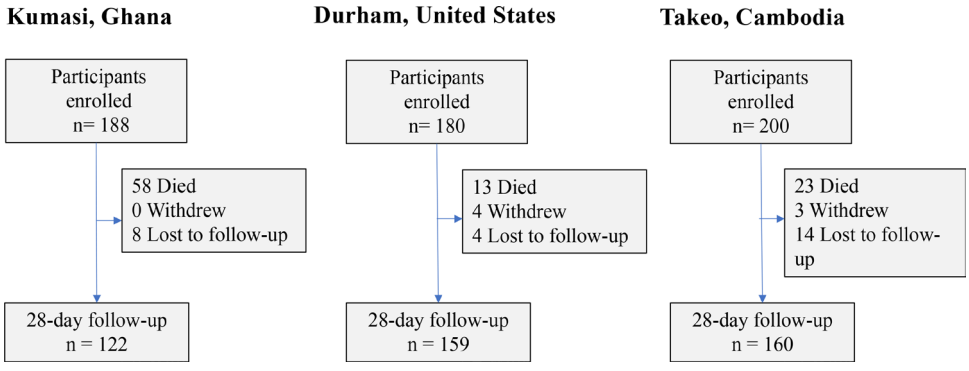


Figure 1 Enrolment flow diagram across cohorts.

Table 1 Baseline demographic characteristics stratified by sites

Characteristic	Total (n=567)	Takeo, Cambodia (n=200)	Durham, USA (n=180)	Kumasi, Ghana (n=187)	P value*
Female gender, no. (%)	243 (42.9%)	64 (32.0%)	81 (45.0%)	98 (52.4%)	<0.001
Age, years, median (IQR)	50 (36–63)	50 (36–62)	52.5 (40 – 63)	46 (35–63)	0.151
Medical history†, no. (%)					
Cancer	44 (9.9%)	0 (0.0%)	44 (24.4%)	0 (0.0%)	<0.001
Cardiovascular	202 (41.4%)	22 (18.2%)	118 (65.6%)	62 (33.2%)	<0.001
Dermatological	15 (3.1%)	1 (0.8%)	14 (7.8%)	0 (0.0%)	<0.001
Endocrine	126 (25.8%)	6 (5.0%)	74 (41.1%)	46 (24.6%)	<0.001
Gastrointestinal	76 (15.6%)	4 (3.3%)	66 (36.7%)	6 (3.2%)	<0.001
Genitourinary or reproductive	34 (7.0%)	1 (0.8%)	33 (18.3%)	0 (0.0%)	<0.001
HIV	26 (4.7%)	12 (6.2%)	8 (4.5%)	6 (3.2%)	0.388
Neurological	62 (12.7%)	1 (0.8%)	44 (24.4%)	17 (9.1%)	<0.001
Other	206 (42.2%)	48 (39.7%)	151 (83.9%)	7 (3.7%)	<0.001
Psychiatric	143 (29.3%)	41 (33.9%)	78 (43.3%)	24 (12.8%)	<0.001
Renal	41 (8.4%)	0 (0.0%)	41 (22.8%)	0 (0.0%)	<0.001
Respiratory	89 (18.2%)	7 (5.8%)	76 (42.2%)	6 (3.2%)	<0.001
Rheumatological	29 (5.9%)	1 (0.8%)	28 (15.6%)	0 (0.0%)	<0.001
Surgery	27 (5.5%)	0 (0.0%)	22 (12.2%)	5 (2.7%)	<0.001
Baseline scores, no. (%)					
MEWS (≥4)	315 (57.8%)	81 (40.7%)	105 (65.6%)	129 (69.3%)	<0.001
NEWS score (≥5)	324 (61.6%)	90 (47.9%)	98 (64.5%)	136 (73.1%)	<0.001
qSOFA (≥2)	139 (25.4%)	22 (11.1%)	48 (29.6%)	69 (37.1%)	<0.001
SIRS (≥2)	447 (81.8%)	125 (68.3%)	157 (89.2%)	165 (88.2%)	<0.001
UVA (≥2)	199 (37.8%)	47 (25.8%)	68 (42.8%)	84 (45.4%)	<0.001
Baseline scores (median (IQR))					
MEWS	4 (3–6)	3 (2–5)	1 (0–4)	1 (1–2)	<0.001
NEWS	6 (3–8)	4 (2–7)	7 (3–9)	6 (4–8)	<0.001
qSOFA	1 (1–2)	1 (0–1)	1 (0–2)	1 (1–2)	<0.001
SIRS	2 (2–3)	2 (1–3)	3 (2–3)	3 (2–3)	<0.001
UVA	1 (0–3)	1 (0–2)	1 (0–4)	1 (0–4)	<0.001

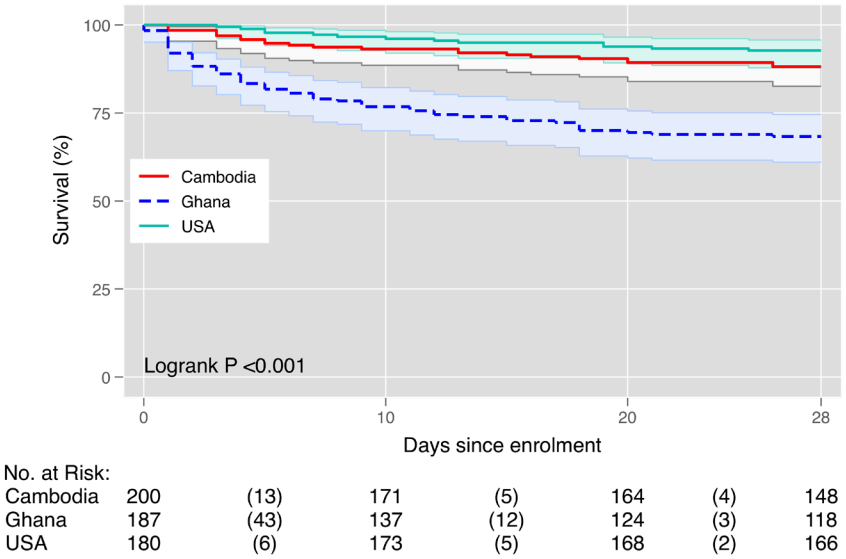
*Categorical parameters compared with χ^2 test and numeric parameters compared with Kruskal-Wallis test. Not adjusted for multiple comparisons.
†There were 79 subjects without comorbidity information in the Cambodia cohort.
MEWS, Modified Early Warning Score; NEWS, National Early Warning Score; qSOFA, quick Sequential Organ Failure Assessment; SIRS, Systemic Inflammatory Response Syndrome; UVA, Universal Vital Assessment.

Hospital physicians who deemed their patients too ill to participate could defer enrolment.

Study procedures

Following informed consent, study team members conducted a detailed medical history, including prior medications and physical examination. Responses were recorded on a standardised case report form and included demographics, medical history, physical examination findings and admission diagnoses. Study specific procedures conducted in Cambodia were described in detail by Rozo *et al*.[16] Similar enrolment and study procedures were followed in Kumasi, Ghana and in Durham, North Carolina, USA. Blood was collected at the time of enrolment, then at 6 hours later, and at 24 hours later. In Ghana and Cambodia, standardised clinical tests included a peripheral venous blood gas with lactate, complete blood count, complete metabolic panel, optional HIV screening with consent (Alere Determine HIV1/2, Abbott, OK, USA), malaria rapid diagnostic tests (SD Bioline Ag. P.f./Pan, Abbott, OK, USA) and aerobic blood cultures (one aerobic bottle, Bactec 9050, BD, NJ, USA) as part of study procedures in Ghana and Cambodia. Microbiological results were available if collected through routine clinical care across cohorts. Additional molecular testing and next generation sequencing for pathogens were also performed on blood samples in the Cambodia cohort as previously described.[16] Participants were followed throughout their hospitalisation and a record review performed at discharge.

An interview was performed, and blood samples were collected at a 28-day follow-up visit across cohorts. When patients could not return in person, study team members attempted to conduct an interview with patients or a legally authorised representative by telephone. Fatal outcomes among each discharged participant were also determined.



No. at Risk:

Cambodia	200	(13)	171	(5)	164	(4)	148
Ghana	187	(43)	137	(12)	124	(3)	118
USA	180	(6)	173	(5)	168	(2)	166

Figure 2 Kaplan-Meier survival plot of 28-day mortality stratified by site.

Using clinical data from case report forms and microbiology diagnostic information, clinical adjudication was performed by three physician reviewers (internal medicine or infectious diseases) to determine the source of infection by anatomic location and pathogen class (ie, bacterial, parasitic, viral or fungal). This was graded on a low, moderate and high level of confidence by two independent reviewers and a third reviewer served as a tiebreaker for discordant conclusions. If the third reviewer did not agree with either adjudicator, then the decision was determined by committee. Microbiological results presented include those adjudicated to be clinically relevant to participant's acute illness.

Patient and public involvement

Patients were not involved in recruitment, design, conduct or dissemination plans of our research. Results of this study were disseminated to hospital and clinical leadership at Takeo Provincial Hospital and Komfo Anokye Teaching Hospital.

Statistical analysis

Summary statistics were calculated for the cohorts individually and pooled, comparing baseline demographics (eg, gender, age, ethnicity, selected medical comorbidities), baseline screening tool scores, physiological parameters, baseline clinical laboratory values using either χ^2 (categorical values), Fishers exact (categorical values) or

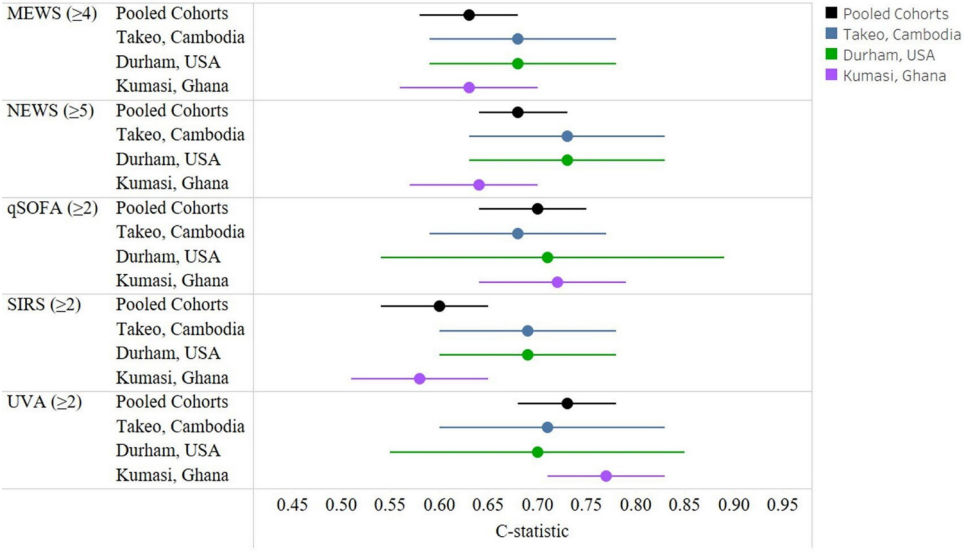


Figure 3 The C-statistic by score overall and by cohort (adjusted for age and sex). MEWS, Modified Early Warning Score; NEWS, National Early Warning Score; qSOFA, quick Sequential Organ Failure Assessment; SIRS, Systemic Inflammatory Response Syndrome; UVA, Universal Vital Assessment.

Kruskal-Wallis (continuous values) tests. Prevalence of diagnoses were described for each cohort by organ system and pathogen type and by anatomic site.

After checking the proportional hazards assumption, Cox regression was performed with bivariate models to evaluate increased risk of death in each cohort by baseline demographics, comorbid conditions, physiological parameters and clinical laboratory parameters. Physiological parameters and clinical laboratory parameters were modelled as dichotomous or ordinal parameters at clinically relevant abnormal range cut-offs (eg, blood urea nitrogen ≥20 mg/dL) to explore associations with increased risk of death and for clinical inference. Screening tools were dichotomised according to current usage, including qSOFA score ≥2 (range, 0 (best) to 3 (worst) points), SIRS score ≥2 (range, 0 (best) to 4 (worst) points), MEWS ≥5 (range, 0 (best) to 13 (worst) points), NEWS ≥5 (range, 0 (best) to 20 (worst) points) and UVA ≥2 (range, 0 (best) to 13 (worst))[12] and were evaluated in Cox regression models unadjusted and adjusted for age and sex for risk of death.[8] Glasgow Coma Scale score (GCS; range, 3 (worst) to 15 (best) points) of less than 15 was used for estimation of the qSOFA score, GCS of ≤3 for unresponsiveness for NEWS and GCS score 3–15 for the 'alert, verbal, pain, unresponsive' scale (alert: GCS 13–15, voice: GCS 9–12; pain: GCS 4–8; unresponsiveness: GCS ≤3) score approximation for MEWS.[17 18] Data were administratively right censored past 28 days. The Harrell's C-statistic was calculated for each screening tool for each cohort, the cohorts combined and Cambodia and Ghana cohorts pooled.[19] This statistic is a performance analogous to area under the receiver operating characteristic curve (AUROC) but accounts for differences over time with survival outcomes. C-statistic confidence interval (CI) estimates were determined.[20] The Cox regression Wald test p values were calculated for each score covariate adjusting for baseline risk estimated by age and sex to determine if scores improved model accuracy above baseline risk.[8 21] Adjustment was limited to age and sex covariates to avoid introducing confounding (eg, ascertainment bias from medical history), type I error from multiple comparisons or overfitting. P values <0.002 were considered significant using a Bonferroni correction for multiple comparisons. Cohort sample sizes were determined a priori through Monte Carlo simulation modelling for prognostic biomarker identification. All statistical analyses were performed in SAS (V.9.4), R V.4.0.2[22] or Stata (V.15.0; StataCorp LLC, College Station, TX, USA).[23]

RESULTS

Summary demographics, sepsis severity and laboratory findings

There were 567 participants across the cohorts including 187 from Kumasi, Ghana, 200 from Takeo, Cambodia and 180 from Durham, North Carolina, USA (figure 1). The study population was predominantly male (57.1% men), with more male participants enrolled in Cambodia than at other sites (68.0% vs 55.0% in the USA and 52.4% in Ghana). The overall median age was 50 years (IQR (IQR), 36 to 63), which was similar across cohorts (table 1). Previously diagnosed comorbid conditions were most common at the US site including a history of cardiovascular (65.6%; N=118), respiratory (42.2%; N=76) or gastrointestinal (36.7%; N=66) conditions.

Clinical physiological and laboratory value abnormalities at enrolment were common with median respiratory rate at 24 (IQR: 20–30), the median white blood count elevated at 12.05×10^9 cells/L (IQR: 8.13 to 16.6×10^9 cells/L) and median lactate elevated at 2.27 mmol/L (IQR: 1.66–3.09 mmol/L) (online supplemental table S1). At enrolment, the proportion of an elevated qSOFA (≥2) at baseline was highest at the Ghana site with 44.4% (N=83) of participants compared with 26.0% (N=52) in Cambodia and 22.2% (N=40) in the USA. The SIRS, MEWS, NEWS and UVA screening tools were similarly higher in the Ghana cohort.

Pathogens detected

The most common positive microbiological results overall included bacteraemia (N=83), respiratory culture growth (N=19), serum hepatitis B surface antigen (N=15) and malaria rapid diagnostic tests (N=11). A minority (121 of 567, 21.3%) of subjects had confirmed infections with complete adjudicator agreement using all available sources of clinical microbiological results (with the notable addition of RNA sequencing of samples from Cambodia[16]) including 90 (15.9%) bacterial, 17 viral (3.0%), 20 malarial (3.5%) and 2 (0.3%) fungal infections identified across all cohorts (online supplemental figure S1). These infection classes were different among sites (χ^2 test p<0.001).

In Cambodia, the most common bacterial infections with complete adjudicator agreement were *B. pseudomallei* (N=10, with blood or respiratory culture growth), presumptive *M. tuberculosis* (N=5, with acid fast positive smears), polymicrobial (N=5) and *O. tsutsugamushi* (N=4, determined by sequencing). The most common causes of bacteraemia (17 total of 200 participants) were *B. pseudomallei* (N=8), *E. coli* (N=3) and polymicrobial infections (N=3). Three participants had a positive malaria RDT. Fungal infections were uncommon with one participant with non-albicans Candidemia and one with cryptococcal meningitis. Two individuals had dengue fever (one PCR positive and one adjudicated IgM positive).

In Ghana, the most common causes of bacteraemia (culture growth from 28 of 187 participants) were *E. coli* (N=6), *S. aureus* (N=6), *Salmonella* spp (N=5) and *S. pneumoniae* (N=3). Nine participants had a positive malaria RDT and 15 had a positive hepatitis B surface antigen.

In the USA, the most common causes of bacteraemia (culture growth from 19 of 180 participants) were *E. coli* (N=5), *K. pneumoniae* (N=3), polymicrobial (N=2), *Pseudomonas* spp (N=2) or *S. aureus* (N=2). Viral infections detected by PCR included rhinovirus (N=5), influenza A

Table 2 Performance characteristics of sepsis score across cohorts for predicting 28-day mortality

Score	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Unadjusted bivariate Cox model C-statistic (95% CI)	Adjusted* Cox model C-statistic (95% CI)	P value (Wald test)
Age and sex	–	–	–	–	–	0.59 (0.53 to 0.64)	
MEWS ≥4*	0.73 (0.63 to 0.82)	0.45 (0.40 to 0.49)	0.21 (0.16 to 0.26)	0.89 (0.85 to 0.93)	0.59 (0.54 to 0.63)	0.63 (0.58 to 0.68)	**0.002†**
NEWS ≥5*	0.86 (0.77 to 0.93)	0.43 (0.38, to 0.48)	0.25 (0.23 to 0.28)	0.93 (0.89 to 0.95)	0.65 (0.64 to 0.67)	0.68 (0.64 to 0.73)	**<0.001†**
qSOFA ≥2*	0.54 (0.44 to 0.65)	0.80 (0.76 to 0.84)	0.35 (0.27 to 0.44)	0.90 (0.87 to 0.93)	0.66 (0.61 to 0.71)	0.70 (0.64 to 0.75)	**<0.001†**
SIRS ≥2*	0.89 (0.80 to 0.94)	0.19 (0.16 to 0.23)	0.17 (0.14 to 0.21)	0.90 (0.82 to 0.95)	0.53 (0.50 to 0.57)	0.60 (0.54 to 0.65)	0.134
UVA≥2*	0.74 (0.64 to 0.83)	0.70 (0.65 to 0.74)	0.33 (0.27 to 0.40)	0.93 (0.90 to 0.95)	0.70 (0.65 to 0.74)	0.73 (0.68 to 0.78)	**<0.001†**

P values are from Wald test of the adjusted Cox regression model.
*Adjusted Cox model C-statistic is adjusted for age and gender.
†Significant at p<0.002.
MEWS, Modified Early Warning Score; NEWS, National Early Warning Score; NPV, Negative predictive value; PPV, Positive predictive value; qSOFA, quick Sequential Organ Failure Assessment; SIRS, Systemic Inflammatory Response Syndrome; UVA, Universal Vital Assessment.

(N=4), respiratory syncytial virus (N=4), human immunodeficiency virus (N=3) and human metapneumovirus (N=3). There was one participant with *Aspergillus fumigatus* fungal pneumonia.

Diagnoses and treatments

Across cohorts, the most common organ system sites of infection were lower respiratory tract infection (28.7%; N=163), multifocal or generalised source of infection (including malaria) (13.6%; N=77) and gastrointestinal (including hepatic) (12.7%; N=72)online supplemental figure S1. The most common antibiotics administered in USA, Ghana and Cambodia were beta-lactam antibiotics (online supplemental figure S2), but antibiotic regimens varied widely among sites. The most common antibiotics classes used were other antibacterials (eg, glycopeptide antibiotics, 58.9%), beta-lactam antibacterials, penicillins (51.7%) and cephalosporin and carbapenem antibacterials (44.4%) in the USA, cephalosporins and carbapenems (64.2%), macrolides, lincosamides and streptogramins (37.4%) and other antibacterials (33.7%) in Ghana, and cephalosporins and carbapenems (73.0%), beta-lactam antibacterials, penicillins (46.5%) and aminoglycoside antibacterials (39.0%) in Cambodia.

Survival

Among all cohorts, 16.4% (N=93) of participants had died at 1 month, including 58 (31.0%) in Ghana, 22 (11.0%) in Cambodia and 13 (7.2%) in the USA (figure 1). Among those that died within 1 month, median time to death was 4 days (IQR: 1–11) in Ghana, 7 days (IQR: 3–16) in Cambodia, 10 (IQR: 5–19) in the USA and 5 days (IQR: 2–13) overall. Parameters to calculate the qSOFA score and 28-day mortality were available for 96.4% participants. Hypernatraemia (>145 mEq/L) had the highest unadjusted risk of death (HR 6.89, 95% CI 3.43 to 13.85) among parameters tested in bivariate models (online supplemental figure S3). All screening tools were associated with an increased risk of death (figure 2) with the largest increase among those with an elevated UVA score (online supplemental figure S3). For individuals with a UVA ≥2 there was a 5.45 times increased risk of death (95% CI 3.39 to 8.76; C-statistic: 0.70) and those with a qSOFA ≥2 had a 4.11 times increased risk of death (95% CI 2.71 to 6.22; C-statistic: 0.66). Those with an elevated SIRS had a 1.81 times increased risk of death (95% CI 0.94 to 3.50; C-statistic:0.53). Elevated NEWS (HR: 4.03; 95% CI 2.24 to 7.26; C-statistic: 0.66) and MEWS (HR: 2.03; 95% CI 1.28 to 3.23; C-statistic: 0.53) had similarly increased risks (figure 3).

Accuracy in an adjusted Cox model was highest for UVA (0.73; 95% CI 0.68 to 0.78) followed by qSOFA (C-statistic: 0.70; 95% CI 0.64 to 0.75) (table 2). The sensitivity for predicting death was highest with SIRS (89%; 95% CI 80% to 94%) but specificity was lowest (19%; 95% CI 16% to 26%). The UVA score had a sensitivity of 74% and specificity of 70%. The qSOFA score had the lowest sensitivity (54%; 95% CI 44% to 65%) but highest specificity

Table 3 Performance characteristics of sepsis score across cohorts for predicting 28-day mortality stratified by site

Model	Takeo, Cambodia C-statistic (95% CI)	P value	Durham, USA C-statistic (95% CI)	P value	Kumasi, Ghana C-statistic (95% CI)	P value
Age and Sex	0.68 (0.59 to 0.78)	–	0.68 (0.54 to 0.81)	–	0.57 (0.49 to 0.64)	–
MEWS	0.68 (0.59 to 0.78)	0.2102	0.68 (0.57 to 0.79)	0.4991	0.63 (0.56 to 0.70)	0.0097
NEWS	0.73 (0.63 to 0.83)	0.0106	0.71 (0.59 to 0.84)	0.2557	0.64 (0.57 to 0.70)	0.0022
qSOFA	0.68 (0.59 to 0.77)	0.5101	0.71 (0.54 to 0.89)	0.0365	0.72 (0.64 to 0.79)	**<0.001***
SIRS	0.69 (0.60 to 0.78)	0.5020	0.69 (0.55 to 0.83)	0.5831	0.58 (0.51 to 0.65)	0.1882
UVA	0.71 (0.60 to 0.83)	0.0109	0.70 (0.55 to 0.85)	0.4753	0.77 (0.71 to 0.83)	**<0.001***

P value are from Wald test of the adjusted Cox regression model. Each model is adjusted for age and sex.
*Significant at p<0.002.
MEWS, Modified Early Warning Score; NEWS, National Early Warning Score; qSOFA, quick Sequential Organ Failure Assessment; SIRS, Systemic Inflammatory Response Syndrome; UVA, Universal Vital Assessment.

(80%; 95% CI 76% to 84%). We observed that the qSOFA discrimination for mortality was moderate with a C-statistic of 0.70 adjusting for age and sex (figure 3). There was similar qSOFA accuracy in individual cohorts from the USA (C-statistic 0.71; 95% CI 0.54 to 0.89), Cambodia (C-statistic: 0.68; 95% CI 0.59 to 0.77) or Ghana (C-statistic: 0.72; 95% CI 0.64 to 0.79) (figure 3). Similarly, the UVA score had moderate accuracy with a C-statistics on 0.73 (95% CI 0.68 to 0.78). Other screening scores had similar moderate C-statistic values. The SIRS C-statistic was 0.60 (95% CI 0.54 to 0.65). Among participants with a NEWS score of ≥5 (62% of the pooled cohort), the C-statistic was 0.68 (95% CI 0.64 to 0.73) and among those with a MEWS score of ≥4 (58% of the pooled cohort), the C-statistic was 0.63 (95% CI 0.58 to 0.68) for death. The qSOFA and UVA scores were significantly greater than baseline risk in Ghana in contrast to other scores or cohorts (table 3). The qSOFA score increased prognostication accuracy in the US cohort with a p=0.02 but this was not significant after correcting for multiple comparisons. In Cambodia, while not significant after correction, NEWS (p=0.01) and UVA (p=0.01) scores increased accuracy greater than baseline risk. When pooling LMIC cohorts (ie, Ghana and Cambodia), after adjustment for age and sex, the qSOFA (C-statistic: 0.71; 95% CI 0.66 to 0.77) and UVA scores (C-statistic: 0.76; 95% CI 0.71 to 0.81) had higher accuracy compared with MEWS (C-statistic: 0.66; 95% CI 0.61 to 0.72), NEWS (C-statistic: 0.70; 95% CI 0.65 to 0.76) and SIRS (C-statistic: 0.61; 95% CI 0.55 to 0.67) (online supplemental table S2). In contrast, in the US cohort, NEWS, MEWS, SIRS, qSOFA and UVA scores after age and sex adjustment each had similar accuracy with C-statistics ranging from 0.66 to 0.71 (table 3 and online supplemental table S3).

DISCUSSION

In pooled prospective international cohorts in Cambodia, Ghana and the USA, the UVA score and Sepsis-3 (qSOFA) performed well with a C-statistic around 0.7 for predicting 28-day mortality. However, this improvement was largely identified in the cohort in Ghana and the accuracy was no different than baseline risk in the Cambodia cohort. There was a trend towards improving prognostication accuracy with the NEWS and UVA score in Cambodia and only with the qSOFA score in the USA. These results suggest that widely used sepsis screening tools may have varying performance for prognostication in diverse settings with different treatment regimens and aetiologies of sepsis. Therefore, screening tools should be selected after validation within populations prior to widespread adoption.

High sodium (hypernatraemia) was associated with the highest risk of 28-day death among individual clinical parameters. Hypernatraemia during critical illness has been previously associated with mortality in large observational studies from high-resource settings.[24 25] Hypernatraemia can occur in sepsis due to intravascular fluid loss due to breakdown of vascular cell junctions, insensible fluid losses or dehydration from the disease process.[26] There can also be an iatrogenic contribution from diuretics, sodium from intravenous fluids or with inadequate fluid resuscitation. Ultimately, there is no data available to precisely determine the causes of hypernatraemia among the participants in our cohorts. However, our results highlight the universal risk of death among those with hypernatraemia among those with sepsis and emphasise the need for close management of fluid and electrolytes across critical care settings.

Current sepsis screening tools have had variable performance when applied for prognostication. SOFA or APACHE scores have been developed specifically for prognostication but required parameters including arterial blood oxygen saturation are often not available.[8] Performance of qSOFA and SIRS for mortality have performed poorly (SIRS, AUROC, 0.61; qSOFA: AUROC, 0.61) for prognostication in high-resource settings intensive care unit (ICU) settings[13] and in diverse LMICs (adjusted SIRS: AUROC, 0.59; adjusted qSOFA: AUROC, 0.70)[8] in prior studies for mortality prognostication. While qSOFA is generally more specific than other screening tools, it is less sensitive than SIRS, MEWS and NEWS, which is consistent with our data.[27] When applied to sepsis identification, Surviving Sepsis 2021 guidelines recommend against solely using qSOFA,[28] due to being a more specific rather

than sensitive screening test. Additionally, qSOFA has been found to be inferior to MEWS, and NEWS but more accurate and specific than SIRS for predicting in-hospital mortality and ICU transfer in a large retrospective cohort of over 30 000 patients in the USA (NEWS: AUROC, 0.77; MEWS: AUROC, 0.73; qSOFA: AUROC, 0.69; SIRS: AUROC, 0.65).[27] Different screening scores have been evaluated in prospective cohorts in sub-Saharan Africa (SSA) previously in Tanzania (qSOFA: AUROC, 0.57; MEWS: AUROC, 0.49)[29] and Rwanda[30] (MEWS: AUROC, 0.69; UVA: AUROC, 0.71; qSOFA: AUROC, 0.65) and in Gabon[31] (UVA: AUROC, 0.90; qSOFA: AUROC, 0.77; MEWS: AUROC, 0.72; SIRS: AUROC, 0.70). Given the performance variability that has been previously observed and was observed in this study, it is prudent to evaluate prediction scores within the populations they serve prior to widespread promotion.

The UVA score performed better than baseline risk in the Ghana cohort. Our results externally validated the UVA score for use prognostication of hospitalised patients with suspected sepsis in Kumasi, Ghana and potentially in the region when demographics are similar. The superiority of the UVA score in the Ghana cohort could be related to similarities in infectious causes of illness with other countries in SSA populations from which the UVA score was derived.[11] In contrast to the score derivation study,[11] UVA score performed similarly to qSOFA in Ghana. The accuracy of the UVA scores was not greater than baseline risk in the cohort in Cambodia after adjustment for multiple comparisons. While conclusions may be limited by sample size, sepsis scores derived from the regions of the world with more similar infectious aetiologies may perform better. Our results highlight the importance of validating scores in new patient populations prior to widespread use.

This study had multiple limitations. First, exclusion criteria of immunocompromising conditions except HIV may have led to a skewed populations from Ghana and Cambodia. These exclusion criteria were created to decrease the effect of comorbid conditions or medications on immune biomarkers. However, in Cambodia and Ghana, immunosuppressive medications or diagnoses of chronic liver or kidney disease may be less common in the general population due to limited access to specialists or specialised medications. Additionally, while there were differences in the baseline severity between cohorts, study processes including inclusion criteria were largely standardised across sites improving the comparability of the cohorts in diverse settings and baseline risk was adjusted in models using age and sex. Diagnostic testing differed at each site and mortality specifically due to sepsis could not be determined. Enrolment was by convenience sampling within the referral hospital catchment area and may not be representative of the general population within these countries. Approximation of the mental status for the MEWS scoring using GCS may not be generalisable to the use of GCS at other sites. However, similar MEWS and NEWS performance was observed across sites. Finally, due to the limited sample size in each of the cohorts, smaller improvements in accuracy may not have been identified in the Cambodia and US cohorts that had less deaths compared with the Ghana cohort.

Inexpensive and readily available tools are needed for triage in resource-limited areas in the world to help identify patients that need escalation and possible transfer to higher levels of care. Current widely used sepsis screening tools represent a clinical benchmark for the development of future triage tools. Research is ongoing to assess point-of-care diagnostics within our sepsis cohort research network. Assays with portable and low-cost inflammation biomarkers tests, molecular diagnostics or point-of-care ultrasound have the potential to augment the performance of clinical screening tools towards a more personalised approach to sepsis recognition and triage.

CONCLUSION

Sepsis screening tools that are widely used during clinical care had sub-optimal performance for risk stratification in three international cohorts with increased performance of the UVA and qSOFA scores in Ghana compared with baseline risk. There remains a need for reliable, low-cost and scalable prognostication methods that are validated in diverse settings.

Author affiliations
[1]Austere environments Consortium for Enhanced Sepsis Outcomes (ACESO), Henry M Jackson Foundation for the Advancement of Military Medicine Inc, Bethesda, Maryland, USA
[2]Emergency Medicine, Komfo Anokye Teaching Hospital, Kumasi, Ghana
[3]Takeo Provincial Referral Hospital, Takeo, Cambodia
[4]Duke University School of Medicine, Durham, North Carolina, USA
[5]Danaher Diagnostics, Washington, D.C, USA
[6]Naval Medical Research Center Infectious Diseases Directorate, Bethesda, Maryland, USA
[7]Duke Global Health Institute, Durham, North Carolina, USA
[8]Henry M Jackson Foundation for the Advancement of Military Medicine Inc, Bethesda, Maryland, USA
[9]US Naval Medical Research Unit No 2, Phnom Penh, Cambodia
[10]Naval Medical Research Unit-3 Ghana Detachment, Accra, Ghana
[11]Global Center for Health Security, University of Nebraska Medical Center, Omaha, Nebraska, USA
[12]Department of Internal Medicine, University of Nebraska Medical Center, Omaha, Nebraska, USA
[13]Austere environments Consortium for Enhanced Sepsis Outcomes (ACESO), Biological Defense Research Directorate, Naval Medical Research Center-Frederick, Frederick, Maryland, USA
[14]Department of Clinical Microbiology, Komfo Anokye Teaching Hospital, Kumasi, Ghana
[15]Komfo Anokye Teaching Hospital, Kumasi, Ghana

Twitter Paul W Blair @PaulBlair9

Contributors PWB, RM, JB, LMK, DAS, REM performed data curation and analyses. PWB, KLS and DVC developed the manuscript concept. DVC and KLS provided resources for research development. AL, KLS, CKO, JC, ST, EK, ET, CB, CWW, AF, AL, DF, JVL, MP, MR, NA, CD, MGG, TV, AKO-O, DA and GO were involved in protocol development and data generation. LMK, MGS, WRH, SK were involved in research operations. AL, KLS, CWW, DF, ET, CB, DVC and PWB were involved in manuscript revisions. PWB as the guarantor accepts full responsibility of the finished work, had access to the data, and controlled the decision to publish. All authors reviewed and approved of this manuscript.

Funding Defense Threat Reduction Agency (JSTO-CBA) to Naval Medical Research Center (NMRC) (HDTRA1516108), Defense Health Bureau of Medicine & Surgery to NMRC for Combating Antibiotic Resistance Bacteria (FY1819 0130.1832), Naval Medical Logistics Command Cooperative Agreement (N626451920001).

Disclaimer KLS is an employee of the US government, and CB, NA, CD, MP and AL are military service members. This work was prepared as a part of official duties. Title 17 U.S.C. 105 provides that 'Copyright protection under this title is not available for any work of the United States Government'. Title 17 U.S.C. 101 defines a US Government work as a work prepared by a military service member or employee of the US Government as part of a person's official duties. The views expressed reflect the results of research conducted by the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the US Government.

Competing interests ET has held equity and consulted for Predigen and Biomeme, and he is an employee of Danaher Diagnostics. All other authors declared no competing interests.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and study protocols were approved by the Naval Medical Research Center (NMRC) Institutional Review Board (IRB) (Cambodia sepsis study # NMRC.2013.0019; Ghana sepsis study # NMRC.2016.0004-GHA; Duke sepsis study Duke#PRO00054849). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. De-identified data may be made available upon reasonable request to the corresponding author.

ORCID iDs
Paul W Blair http://orcid.org/0000-0002-3817-5987
Alex Kwame Owusu-Ofori http://orcid.org/0000-0002-2172-8645

REFERENCES

1 Reinhart K, Daniels R, Kissoon N, *et al*. Recognizing sepsis as a global health priority-a who resolution. *N Engl J Med* 2017;377:414–7.
2 Rudd KE, Johnson SC, Agesa KM, *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.
3 Maitland K, Kiguli S, Opoka RO, *et al*. Mortality after fluid bolus in African children with severe infection. *N Engl J Med* 2011;364:2483–95.
4 Andrews B, Semler MW, Muchemwa L, *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.
5 WHO. *Dengue: guidelines for diagnosis, treatment, prevention and control: new edition*. Geneva, 2009.
6 Singer M, Deutschman CS, Seymour CW, *et al*. The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* 2016;315:801–10.
7 Zimmerman JE, Kramer AA, McNair DS, *et al*. Acute physiology and chronic health evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med* 2006;34:1297–310.
8 Rudd KE, Seymour CW, Aluisio AR, *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;319:2202–11.
9 Subbe CP, Kruger M, Rutherford P, *et al*. Validation of a modified early warning score in medical admissions. *QJM* 2001;94:521–6.
10 Smith GB, Prytherch DR, Meredith P, *et al*. The ability of the national early warning score (news) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. *Resuscitation* 2013;84:465–70.
11 Moore CC, Hazard R, Saulters KJ, *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 Health* 2017;2:e000344.
12 Adegbite BR, Edoa JR, Ndzebe Ndoumba WF, *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.
13 Raith EP, Udy AA, Bailey M, *et al*. Prognostic accuracy of the SOFA score, SIRS criteria, and qsofa score for in-hospital mortality among adults with suspected infection admitted to the intensive care unit. *JAMA* 2017;317:290–300.
14 Krishnan S, Beckett C, Espinosa B, *et al*. Austere environments consortium for enhanced sepsis outcomes (ACESO). *Shock* 2020;53:377–8.
15 Schully KL, Berjohn CM, Prouty AM, *et al*. Melioidosis in lower provincial cambodia: a case series from a prospective study of sepsis in takeo province. *PLoS Negl Trop Dis* 2017;11:e0005923.
16 Rozo M, Schully KL, Philipson C, *et al*. An observational study of sepsis in takeo province cambodia: an in-depth examination of pathogens causing severe infections. *PLoS Negl Trop Dis* 2020;14:e0008381.
17 Kelly CA, Upex A, Bateman DN. Comparison of consciousness level assessment in the poisoned patient using the alert/verbal/painful/unresponsive scale and the Glasgow coma scale. *Ann Emerg Med* 2004;44:108–13.
18 Nedea D. AVPU scale. Available: https://www.mdapp.co/avpu-scale-calculator-163/ [Accessed 18 Jan 2022].
19 Harrell FE Jr, Califf RM, Pryor DB, *et al*. Evaluating the yield of medical tests. *JAMA* 1982;247:2543–6.
20 Han X, Zhang Y, Shao Y. On comparing 2 correlated C indices with censored survival data. *Stat Med* 2017;36:4041–9.
21 Pepe MS, Kerr KF, Longton G, *et al*. Testing for improvement in prediction model performance. *Stat Med* 2013;32:1467–82.
22 Team T R D C. R: A language environment for statistical computing. R foundation for statistical computing; 2020.
23 StataCorp L. *Stata statistical software: release 16 college station*. TX StataCorp LP, 2019.
24 Castello LM, Gavelli F, Baldrighi M, *et al*. Hypernatremia and moderate-to-severe hyponatremia are independent predictors of mortality in septic patients at emergency department presentation: a sub-group analysis of the need-speed trial. *Eur J Intern Med* 2021;83:21–7.
25 Olsen MH, Møller M, Romano S, *et al*. Association between ICU-acquired hypernatremia and in-hospital mortality: data from the medical information mart for intensive care III and the electronic ICU collaborative research database. *Crit Care Explor* 2020;2:e0304.
26 Lindner G, Funk GC. Hypernatremia in critically ill patients. *J Crit Care* 2013;28:216.
27 Churpek MM, Snyder A, Han X, *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.
28 Evans L, Rhodes A, Alhazzani W, *et al*. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Intensive Care Med* 2021;47:1181–247.
29 Carugati M, Zhang HL, Kilonzo KG, *et al*. Predicting mortality for adolescent and adult patients with fever in resource-limited settings. *Am J Trop Med Hyg* 2018;99:1246–54.
30 Klinger A, Mueller A, Sutherland T, *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.
31 Schmedding M, Adegbite BR, Gould S, *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.