

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/107021/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Szakmany, Tamas , Pugh, R., Kopczynska, Maja, Lundin, Robert M., Sharif, Ben, Morgan, P., Ellis, G., Abreu, Julimar, Kulikouskaya, Svetlana, Bashir, K., Galloway, Luke, Al-Hassan, H., Grother, Thomas, McNulty, Paul, Seal, Steffan T., Cains, Alice, Vreugdenhil, M., Abdimalik, Mustafa, Dennehey, Naomi, Evans, G., Whitaker, J., Beasant, Elizabeth, Hall, C., Lazarou, Maria, Vanderpump, Chloe V., Harding, K., Duffy, Leo, Guerrier Sadler, Abigail, Keeling, Rachel, Banks, Charldre, Ng, S. W. Y., Heng, Sieh Yen, Thomas, D., Puw, Elen W., Otahal, Igor, Battle, C., Minik, O., Lyons, R. A. and Hall, Judith Elizabeth 2018. Defining sepsis on the wards: results of a multi-centre point-prevalence study comparing two sepsis definitions. *Anaesthesia* 73 (2) , pp. 195-204. 10.1111/anae.14062

Publishers page: <http://dx.doi.org/10.1111/anae.14062>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Original Article

Defining sepsis on the wards: results of a multi-centre point-prevalence study comparing two sepsis definitions

T. Szakmany,^{1,2} R. Pugh,³ M. Kopczynska,⁴ R. M. Lundin,⁴ B. Sharif,⁵ P. Morgan,⁶ G. Ellis,^{1,7} J. Abreu,⁴ S. Kulikouskaya,⁴ K. Bashir,⁴ L. Galloway,⁴ H. Al-Hassan,⁴ T. Grother,⁴ P. McNulty,⁴ S. T. Seal,⁴ A. Cains,⁴ M. Vreugdenhil,⁴ M. Abdimalik,⁴ N. Dennehey,⁴ G. Evans,⁴ J. Whitaker,⁴ E. Beasant,⁴ C. Hall,⁴ M. Lazarou,⁴ C. V. Vanderpump,⁴ K. Harding,⁴ L. Duffy,⁴ A. Guerrier Sadler,⁴ R. Keeling,⁴ C. Banks,⁴ S. W. Y. Ng,⁴ S. Y. Heng,⁴ D. Thomas,⁴ E. W. Puw,⁴ I. Otahal,⁸ C. Battle,⁹ O. Minik,¹⁰ R. A. Lyons,¹¹ J. E. Hall,¹² on behalf of the Welsh Digital Data Collection Platform collaborators*

1 Senior Lecturer, 4 Medical Student, 5 Core Trainee, 12 Professor, Department of Anaesthesia, Intensive Care and Pain Medicine, Division of Population Medicine, Cardiff University, UK

2 Consultant, Anaesthetic Directorate, Aneurin Bevan University Health Board, Royal Gwent Hospital, Newport, Gwent, UK

3 Consultant, Anaesthetic Department, Glan Clywdd Hospital, Betsi Cadwaladar University Health Board, Bodelwyddan, Rhyl, UK

6 Consultant, 7 Nurse Consultant, Critical Care Directorate, University Hospital of Wales, Cardiff and Vale University Health Board, Cardiff, UK

8 Consultant, Anaesthetist, Anaesthetic Department, Glangwili General Hospital, Hywel Dda University Health Board, Carmarthen, UK

9 Consultant Physiotherapist, Critical Care Directorate, Morriston Hospital, Abertawe Bro Morgannwg University Health Board, Heol Maes Eglwys, Swansea, UK

10 Specialty Doctor, ACT Directorate, Royal Glamorgan Hospital, Cwm Taf University Health Board, Ynysmaerdy, Llantrisant, UK

11 Professor, Farr Institute, Data Science Building, Swansea University Medical School, Swansea, UK

Summary

Our aim was to prospectively determine the predictive capabilities of SEPSIS-1 and SEPSIS-3 definitions in the emergency departments and general wards. Patients with National Early Warning Score (NEWS) of 3 or above and suspected or proven infection were enrolled over a 24-h period in 13 Welsh hospitals. The primary outcome measure was mortality within 30 days. Out of the 5422 patients screened, 431 fulfilled inclusion criteria and 380 (88%) were recruited. Using the SEPSIS-1 definition, 212 patients had sepsis. When using the SEPSIS-3 definitions with Sequential Organ Failure Assessment (SOFA) score ≥ 2 , there were 272 septic patients, whereas with quickSOFA score ≥ 2 , 50 patients were identified. For the prediction of primary outcome, SEPSIS-1 criteria had a sensitivity (95%CI) of 65% (54–75%) and specificity of 47% (41–53%); SEPSIS-3 criteria had a sensitivity of 86% (76–92%) and specificity of 32% (27–38%). SEPSIS-3 and SEPSIS-1 definitions were associated with a hazard ratio (95%CI) 2.7 (1.5–5.6) and 1.6 (1.3–2.5), respectively. Scoring system discrimination evaluated by receiver operating characteristic curves was highest for Sequential Organ Failure Assessment score (0.69 (95%CI 0.63–0.76)), followed by NEWS (0.58 (0.51–0.66)) ($p < 0.001$). Systemic inflammatory response syndrome criteria (0.55 (0.49–0.61)) and quickSOFA score (0.56 (0.49–0.64)) could not predict outcome. The SEPSIS-3 definition identified patients with the highest risk. Sequential Organ Failure Assessment score and NEWS were better predictors of poor outcome. The Sequential Organ Failure Assessment score appeared to be the best tool for identifying patients with high risk of death and sepsis-induced organ dysfunction.

Correspondence to: T. Szakmany

Email: szakmany1@cardiff.ac.uk

Accepted: 14 August 2017

Keywords: mortality; NEWS; qSOFA; sepsis; SIRS; SOFA

*Membership of the Welsh Digital Data Collection Platform Collaborators is provided in the Appendix

Introduction

Sepsis is defined as a dysregulated host response to infection, resulting in acute organ dysfunction [1]. Although the condition has been thoroughly studied in the intensive care unit (ICU), accurate data collection outside of this setting is less well-developed. It is thought, however, that the number of cases in the wider hospital is far higher [2–4]. In the UK, anaesthetists and critical care practitioners have been at the forefront of developing effective systems to identify and treat patients with sepsis outside critical care areas. They have identified a clear need to understand the significance of the condition in the pre-ICU environment and the tools we might use to identify and treat those most at risk [5].

We previously reported the results of a point-prevalence feasibility study and subsequent study of all Welsh centres using the 1992 International Consensus Criteria for sepsis (SEPSIS-1), using electronic data collection and real-time data monitoring [6–8]. We found that 4% of hospitalised patients had sepsis, half of whom had significant organ dysfunction (severe sepsis). Strikingly, the 90-day mortality among the whole hospital cohort was in excess of 30% for sepsis, and almost 40% for severe sepsis [7].

Concurrently, the validity and clinical utility of the existing sepsis definitions, which had previously been based on the concept of the systemic inflammatory response syndrome (SIRS), were questioned [9]. The Third International Consensus Definitions for sepsis (SEPSIS-3) have recently been published, with significantly revised clinical criteria, including the use of Sequential Organ Failure Assessment (SOFA) scores and the quick SOFA (qSOFA) screening tool for non-ICU settings [1, 10]. During the development phase, most of the datasets used were from North America, and included variable proportions of non-ICU patients

[10]. It is not known how the new SEPSIS-3 definitions would perform compared with SEPSIS-1 definitions in identifying patients at risk with sepsis in a UK ward setting, and furthermore how they might perform compared with a well-established track and trigger tool, the National Early Warning Score (NEWS) [11, 12].

Our objectives were to determine the ability of the SEPSIS-1 definition using the SIRS criteria, the SEPSIS-3 definition using SOFA and qSOFA scores and the NEWS track and trigger tool to predict outcome outside of the ICU.

Methods

This multi-centre, prospective, observational study of patients with suspected sepsis in 13 hospitals in Wales was approved by the South Wales Regional Ethics Committee, and patients or their proxy, in cases of patients lacking capacity, gave written informed consent. We enrolled consecutive patients presenting to hospitals in Wales with 24/7 consultant-level emergency department (ED) supervision, and the facility to admit and treat any acutely unwell patient. We screened patients in the ED or in an acute in-patient ward setting with suspected or proven infection on 19 October 2016, Wednesday (08:00 h to 07:59 h the following day). This date represented a typically ‘average’ day in the national health service [13, 14]. We approached all patients with NEWS ≥ 3 in whom the treating clinical teams had a high degree of clinical suspicion of an infection (documented as such in the medical or nursing notes), and following consent, we screened for the presence of sepsis using either SEPSIS-1 or SEPSIS-3 definitions.

We did not study patients if they were less than 18 years of age, or if they were already in intensive care or high dependency units. We referred patients to the clinical teams if the medical student data collectors felt

they needed urgent medical attention due to their condition, in line with the requirements of the ethics approval. To facilitate linkage to national databases for the collection of follow-up data, we collected patient-identifiable data and entered it on to the secure data collection tool.

We defined sepsis as the presence or strong suspicion of infection, together with two or more SIRS criteria according to the SEPSIS-1 definition; or as the presence or strong suspicion of infection together with SOFA score 2 or above, or qSOFA score 2 or above, according to the SEPSIS-3 definition. We used the SIRS criteria: respiratory rate greater than 20 breaths.min⁻¹; temperature greater than 38°C or less than 36°C; heart rate greater than 90 beats.min⁻¹; and white blood cell count greater than 12,000 mm⁻³, less than 4000 mm⁻³ or greater than 10% bands [6]. We defined qSOFA scores as systolic blood pressure ≤ 100 mmHg, respiratory rate ≥ 22 breaths.min⁻¹ and altered mental status (defined as either a Glasgow Coma Scale score ≤ 13 or an Alert Voice Pain Unresponsive scale (AVPU) other than 'Alert') [10]. We calculated SOFA and NEWS scores based on previously published tables [11, 15].

To calculate SOFA scores and determine organ dysfunction according to the SEPSIS-1 definition, we used laboratory values within 24 h of study enrolment, and if no prior values were available, a median (normal) value was imputed, as in previous studies [3, 10, 16]. Most patients did not have an arterial blood gas available at time of observation, so to calculate the respiratory component of the SOFA score, we followed the algorithm developed and validated by Pandharipande et al. [16]. We defined infection-related acute organ dysfunction according to the SEPSIS-1 criteria as any of the following present: systolic BP < 90 mmHg or mean arterial pressure < 65 mmHg or lactate > 2.0 mmol.l⁻¹ (after initial fluid challenge), international normalised ratio > 1.5 or activated partial thromboplastin time > 60 s, bilirubin > 34 µmol.l⁻¹, urine output < 0.5 ml.kg.h⁻¹ for 2 h, creatinine > 177 µmol.l⁻¹, platelets < 100 × 10⁹ l⁻¹, PaO₂/FiO₂ ratio below 250, or as SOFA score two or above according to the SEPSIS-3 definition [1, 6]. We recorded the NEWS score on study entry, and we noted if this was the worst value in the preceding 24-h period. [10, 15]

Data collectors, working in pairs to ensure data validity and appropriate clinical knowledge, were

supported by continuous online web chat. This ensured that senior clinicians identified through the Welsh Intensive Care Society Audit and Research Group and three study coordinators were available throughout the trial period. We provided key study information through emails, face-to-face training and online video tutorials, which included the protocol, answers to key questions and description of the electronic case report form (eCRF) on the electronic tablets. We previously published the details of the digital data collection platform developed for this study [8].

We collected data from medical and nursing records, including baseline characteristics, baseline comorbidity and frailty (according to the Dalhousie Clinical Frailty Scale), physiological and laboratory values and process measures (such as critical care involvement and completion of sepsis care bundles) [17]. We followed up patients until 30 days after study enrolment.

The primary outcome measure was mortality within 30 days of recruitment. Secondary outcomes were the presence of organ dysfunction defined by SOFA score > 2 or the presence of 'severe sepsis' according to the SEPSIS-1 definition [10, 18].

Categorical variables are described as proportions and were compared using Chi-square or Fisher's exact test. We performed comparisons of continuous variables using one-way ANOVA or Mann-Whitney U-test as appropriate.

To assess the performances of the SEPSIS-1 and SEPSIS-3 definitions to predict the primary end-point, we calculated diagnostic performances (sensitivity, specificity, negative and positive predictive values). We constructed a receiver operating characteristic (ROC) curve and calculated the corresponding area under the ROC curve (AUROC). We plotted Kaplan-Meier survival curves and compared time-to-event data using log-rank testing. We estimated the respective hazard ratios (HRs) for the primary outcome within 30 days of SEPSIS-1 and SEPSIS-3 definitions with a Cox proportional hazards model after adjustment for measured confounders. The model fit was assessed by the -2 log likelihood statistics and Chi-square test. All statistical tests were calculated using SPSS 20.0 (SPSS Inc., Chicago, IL, USA). A two-tailed $p < 0.05$ was considered statistically significant.

Results

There were 5422 inpatients in the 24-h study period in the 13 participating hospitals (Fig. 1). Four hundred and thirty-one (7.9%) patients had NEWS ≥ 3 and documented clinical suspicion of infection, and all were approached for recruitment. Sixty-four patients (16.8%) were recruited in the ED, and the others from a variety of ward-based environments. Baseline characteristics are summarised in Table 1.

We identified 212 patients as having sepsis using the SEPSIS-1 definition, and 272 patients using the SEPSIS-3 definition with SOFA ≥ 2 (Fig. 2); using the qSOFA, 50 fulfilled the definition criteria (Fig. 2). Out of the cohort of 380 patients, 44 fulfilled neither the SEPSIS-1 nor the SEPSIS-3 criteria (Fig. 2). The characteristics of these groups and secondary outcomes are shown in the Supporting Information Table S1. Sepsis-related organ dysfunction ('severe sepsis') was present in 124 out of 212 patients (58.5%) according to SEPSIS-1 criteria. Ninety-nine out of 124 (79.8%) patients had SOFA ≥ 2 and 24 out of 124 (19.4%) had qSOFA ≥ 2 .

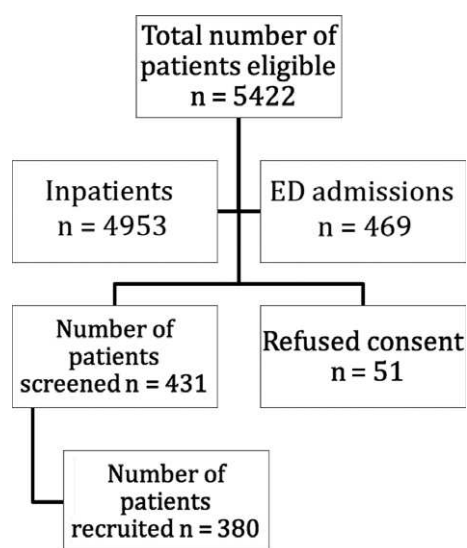


Figure 1 Organisational flow chart of the study. Two hundred and seven patients gave consent on the day, 66 patient representatives gave assent and 107 patients were entered following professional assent. Fifty-one patients (46 patients and 5 patient representatives) refused participation and no data were collected. ED, emergency department.

Out of the 272 patients with sepsis using the SEPSIS-3 definition, 183 (67.3%) fulfilled 'severe sepsis' criteria. Two hundred and thirty-two out of 272 (85.3%) patients had SOFA ≥ 2 using only basic physiological (respiratory, cardiovascular and neurological) parameters.

SEPSIS-1 and SEPSIS-3 definitions identified various proportions of 78 out of 380 (20.5%) patients who died within 30 days (Fig. 3). We found a statistically significant difference in the survival of patients described by the SEPSIS-1 and SEPSIS-3 definitions, or meeting both criteria (Fig. 4).

We report the predictive performances of SEPSIS-1 and SEPSIS-3 definitions in Table 2 and Supporting Information Fig. S1.

After adjustment for age and the presence of heart failure and using a Cox model, we found that the SEPSIS-3 definition was associated with death with HR (95%CI) 2.7 (1.5–5.6). The previous SEPSIS-1 definition had a HR of 1.6 (1.03–2.5).

Scoring system discrimination for the primary outcome was highest for SOFA (AUROC (95%CI) 0.70 (0.63–0.77), $p < 0.001$), followed by NEWS (0.59 (0.51–0.66), $p = 0.02$). The positive likelihood ratio (95%CI) was 1.27 (1.13–1.43) for SOFA, and 1.48 (1.02–2.16) for NEWS. The negative predictive value for SOFA was 89% (81–94%) and for NEWS 73% (67–77%). The SIRS (0.55 (0.48–0.62)) and qSOFA score (0.57 (0.49–0.64)) could not statistically predict outcome in this patient population ($p = 0.21$ and 0.07 for SIRS and qSOFA, respectively). We report the predictive capabilities of qSOFA ≥ 2 , severe sepsis criteria defined by the SEPSIS-1 definition and NEWS ≥ 6 in Table 2.

Prognostic performances of SOFA, qSOFA, SIRS and NEWS to predict acute organ dysfunction are reported in Supporting Information Table S2 and Fig. S2. Sequential Organ Failure Assessment was the best predictive model (AUROC 0.950 (0.930–0.971), $p < 0.001$), followed by NEWS (0.694 (0.634–0.754), $p < 0.001$), qSOFA (0.668 (0.606–0.730), $p < 0.001$) and SIRS (0.580 (0.514–0.647), $p = 0.029$).

Fifty-nine patients (15.5%) were screened for sepsis using the official All Wales sepsis screening tool. The 'Sepsis 6' bundle was completed on 44 occasions (11.6%), and critical care outreach was involved in 33

Table 1 Baseline characteristics of the patients in the whole cohort and comparing the survivors with those who died within 30 days. Values are number (proportion) or median (IQR [range]).

	All patients n = 380	Patients who died n = 78	Survivors n = 302
Sex; men	180 (47%)	46 (59%)	134 (44%)
Age; y	74 (61–83 [18–100])	77 (72–87 [22–100])	73 (58–81 [18–96])
Systolic blood pressure; mmHg	113 (99–136 [74–243])	113 (101–137 [80–243])	112 (99–136 [74–220])
Respiratory rate; breaths.min ⁻¹	20 (18–22 [13–40])	20 (18–23 [13–32])	20 (18–22 [13–40])
Heart rate; beats.min ⁻¹	94 (80–105 [38–198])	95 (83–108 [47–179])	93 (80–104 [38–198])
Glasgow Coma Scale < 15	51 (13%)	21 (25%)	30 (10%)
Temperature; Celsius	36.6 (36.1–37.3 [34.2–39.7])	36.5 (35.9–37.2 [34.3–38.7])	36.8 (36.1–37.4 [34.4–39.7])
AVPU < Alert	20 (5%)	10 (12%)	10 (3%)
Clinical signs of infection			
Cough	176 (46%)	42 (53%)	134 (44%)
Dysuria	39 (10%)	10 (13%)	29 (10%)
Abdominal pain	68 (18%)	17 (22%)	51 (17%)
Headache	20 (5%)	1 (1%)	19 (6%)
Other	162 (55%)	34 (44%)	128 (42%)
Laboratory results			
White blood cell count; cells.ml ⁻¹	11,250 (8375–15,225 [0–56,000])	11,700 (8900–17,250 [1300–45,000])	11,200 (8200–14,625 [0–56,000])
Platelet count; 1000.ml ⁻¹	269 (187–348 [0–920])*	223 (147–376 [10–876])	274 (197–347 [0–920])
Creatinine; µmol.l ⁻¹	76 (61–107 [2–671])*	89 (63–137 [21–671])	75 (60–99 [2–588])
Bilirubin; µmol.l ⁻¹	10 (6–16 [0–570])*	14 (7–28 [3–341])	10 (6–15 [2–570])
Lactate; mmol.l ⁻¹	1.5 (1.1–2.6 [0.0–14.7])*	1.8 (0.9–3.0 [0–7.4])	1.5 (1.1–2.45 [0–14.7])
Clinical frailty score	5 (3–6 [1–9])	6 (5–7 [2–9])	4 (3–6 [1–9])
SIRS			
0	20 (5%)	4 (5%)	16 (5%)
1	116 (31%)	18 (23%)	98 (32%)
2	127 (33%)	28 (36%)	99 (33%)
3	92 (24%)	24 (31%)	68 (22%)
4	25 (7%)	4 (5%)	21 (7%)
qSOFA			
0	152 (40%)	27 (34%)	125 (41%)
1	177 (47%)	33 (42%)	144 (48%)
2	43 (11%)	15 (19%)	29 (10%)
3	7 (2%)	3 (4%)	4 (1%)
SOFA	2 (1–4 [1–11])	4 (2–5 [1–11])	2 (1–3 [1–10])
SOFA ≥ 2			
No	108 (28%)	10 (13%)	98 (32%)
Yes	272 (72%)	67 (86%)	205 (67%)
NEWS	4 (3–6 [3–15])	5 (3–7 [3–15])	4 (3–6 [3–12])
NEWS ≥ 6			
No	265 (70%)	46 (59%)	219 (73%)
Yes	115 (30%)	32 (41%)	83 (27%)

*Data unavailable in 36.3% (serum bilirubin), 7.9% (serum creatinine), 6.3% (platelet count) and 66.3% (serum lactate) cases.

AVPU, Alert/Verbal response/Response to pain/Unresponsive; SIRS, systemic inflammatory response syndrome; qSOFA, quick Sequential Organ Failure Assessment score; SOFA, Sequential Organ Failure Assessment score; NEWS, National Early Warning Score.

cases (8.7%). Intravenous antibiotics were administered either as a mono- or a combination therapy to 220 (57.9%) patients.

Discussion

This is, to our knowledge, the first prospective evaluation of the diagnostic and predictive capability of

SEPSIS-1 vs. SEPSIS-3 criteria in the UK. There was considerable overlap between them, with SEPSIS-3 identifying a larger proportion of patients at risk. However, of 63 (16.6%) patients, 12 of them falling into the previous 'severe sepsis' category would have been missed by applying only the new SEPSIS-3 definitions. On the other hand, application of SIRS-based criteria

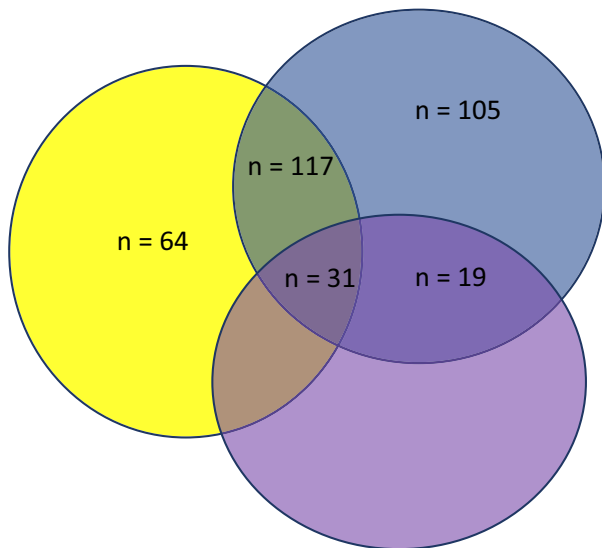


Figure 2 Patients identified having sepsis using the SEPSIS-1 and SEPSIS-3 clinical criteria. ● SOFA, Sequential Organ Failure Assessment score; ● SIRS, systemic inflammatory response syndrome criteria; ● qSOFA, quick Sequential Organ Failure Assessment score; SEPSIS-1 is defined by $\text{SIRS} \geq 2$. SEPSIS-3 is defined by $\text{SOFA} \geq 2$ and/or $\text{qSOFA} \geq 2$. Forty-four did not fulfil either SEPSIS-1 or SEPSIS-3 criteria.

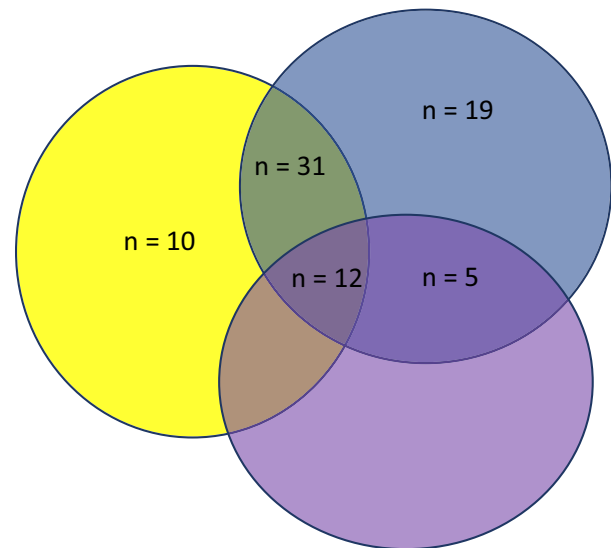


Figure 3 Distribution of SEPSIS-1 and SEPSIS-3 clinical criteria in patients who died within 30 days ($n = 78$). ● SOFA, Sequential Organ Failure Assessment score; ● SIRS, systemic inflammatory response syndrome criteria; ● qSOFA, quick Sequential Organ Failure Assessment score; SEPSIS-1 is defined by $\text{SIRS} \geq 2$. SEPSIS-3 is defined by $\text{SOFA} \geq 2$ and/or $\text{qSOFA} \geq 2$. One patient did not fulfil either SEPSIS-1 or SEPSIS-3 criteria.

(SEPSIS-1) did not identify 105 (27.3%) patients, all of whom had evidence of acute organ dysfunction.

Our results add further to the debate about the clinical usefulness of the qSOFA score, which was developed as an easy-to-use prediction tool for identifying patients at risk in the sepsis population [10]. There is ongoing controversy surrounding the utility and efficacy of qSOFA in the pre-hospital, ED and general ward setting [19–23]. We found that using only the qSOFA score, 50 (13.2%) patients would have been diagnosed with sepsis, missing 116 (30.5%) patients with organ dysfunction. In contrast to the results of Seymour et al., qSOFA also failed to predict outcome at 30 days, and did not offer any predictive value over the SOFA and NEWS scores for ability to predict infection-induced acute organ failure in this patient population [10]. We found a striking disconnect between SOFA and qSOFA scores. Although our sample size is too small to draw firm conclusions, we have seen that the biggest discrepancy was between the respiratory element of SOFA and qSOFA scores (data

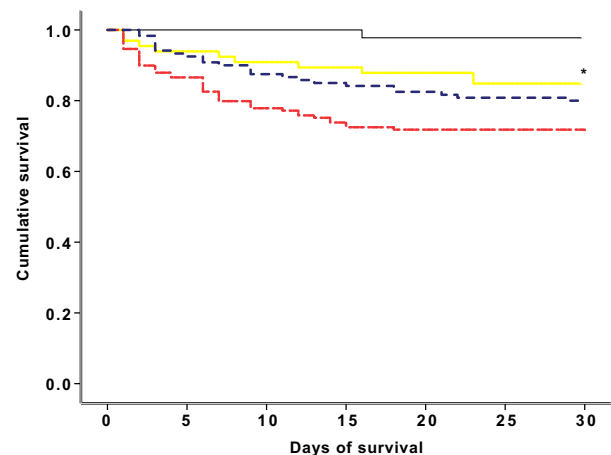


Figure 4 Survival difference of patients with different definitions of sepsis. Not meeting any sepsis criteria (black solid line), SEPSIS-1 definition (yellow dotted line), SEPSIS-3 definition (blue dashed line), both SEPSIS-1 and SEPSIS-3 definition (red dashed line with dots), $*p = 0.015$.

not shown). It is possible that the $\text{SpO}_2/\text{FiO}_2$ ratio used in our study is a much more sensitive parameter to indicate respiratory compromise, than the high

Table 2 Diagnostic performances of different sepsis definitions and clinical tools for the prediction of mortality at 30 days.

	SEPSIS-1	SEPSIS-3	qSOFA ≥ 2	SEPSIS-1 severe sepsis	NEWS ≥ 6
Sensitivity; % (95%CI)	68 (56–78)	86 (76–92)	22 (14–33)	92 (83–97)	41 (30–53)
Specificity; % (95%CI)	47 (41–53)	32 (27–38)	89 (85–92)	24 (19–29)	73 (67–77)
Positive predictive value; % (95%CI)	25 (19–31)	25 (20–30)	34 (22–49)	24 (19–29)	30 (26–35)
Negative predictive value; % (95%CI)	85 (78–90)	90 (82–95)	82 (77–85)	92 (83–97)	70 (65–74)
Positive likelihood ratio (95%CI)	1.28 (1.06–1.54)	1.27 (1.13–1.43)	1.99 (1.17–3.39)	1.21 (1.11–1.33)	1.49 (1.08–2.06)
Negative likelihood ratio (95%CI)	0.68 (0.49–0.95)	0.43 (0.25–0.76)	0.88 (0.78–0.99)	0.32 (0.15–0.71)	0.81 (0.67–0.98)

SEPSIS-1, 1992 definition of sepsis criteria as defined by Bone et al. [6]; SEPSIS-3, Third International Consensus Definition of sepsis criteria [1]; qSOFA, quick Sequential Organ Failure Assessment score; SEPSIS-1 severe sepsis, sepsis with organ dysfunction as defined by Bone et al. [6]; NEWS, National Early Warning Score.

respiratory rate in the qSOFA system. Our results warrant careful interpretation, as our sample size was orders of magnitude smaller compared with the original qSOFA study [10].

In their large dataset of non-ICU patients, median (IQR) SOFA was 1 (2), with significantly lower hospital mortality of 3% indicating a population at lower risk compared with ours. Interestingly, in their ICU cohort with a mortality of 17%, the AUROC of qSOFA at 0.66 (95%CI 0.64–0.68) was only slightly better than we have observed. Similarly, Raith et al. could not confirm the findings of the original paper in a patient population where the baseline risk was significantly higher with a hospital mortality of 18.7%, with AUROC 0.607 (0.603–0.611) [24]. These data suggest that qSOFA might not be a valuable tool to predict outcome in populations where the baseline risk of death is higher than 15%.

Our findings could support the use of SOFA scores even in a resource-limited ward setting, although it is unclear how this might best be integrated into already established track and trigger systems [10, 19, 22, 25, 26]. Donnelly et al. were able to show in a population-based study that high admission SOFA was the best tool predict poor outcome in the hospital and within one year after discharge, with similar AUROC (0.765) and HR 2.43 (95%CI 1.84–3.21) to ours [25]. It is clear from these emerging data that a SIRS-based classification of sepsis is inferior to SOFA for delineating patient cohorts at the highest risk of poor outcome [1, 22, 25]. The exact cut-off for SOFA might need further recalibration; however, the current threshold of two or more could be used in the vast majority of patients, by

calculating the SOFA score from physiological parameters readily available at the bed-side.

The high specificity and positive predictive value of NEWS ≥ 6 for acute organ dysfunction and adverse outcome underlines the utility and importance of the current escalation protocol ('NEWS Six = Sick') in our healthcare system [27]. Similar to our data, NEWS ≥ 7 was found as the best cut-off for predicting poor outcome in a large retrospective cohort of patients with sepsis [23]. Recently, a multi-centre Scottish study also found that NEWS ≥ 6 carried an increased risk of death and ICU admission in patients admitted to the ED with sepsis [28].

Sepsis either defined by the SEPSIS-1 or SEPSIS-3 criteria had a high mortality; 22.9% of the patients died within 30 days, significantly higher than the 2.2% mortality observed in the group which did not fulfil either criteria for the diagnosis of sepsis. This was almost identical to the 22% 30-day mortality observed in our previous study, but significantly higher than the 6% and 8% mortality observed in the recent studies involving ED and ward patients [7, 19, 23]. This could probably be explained by methodological differences between studies. Churpek et al. used a retrospective dataset, with wider screening criteria more likely to capture patients with lower acuity [23]. In fact, only 28% of their 30677 patients met severe sepsis definition and the mortality of this subset was not reported [23]. With a focus on the ED, only 20% of patients in Freund's study met the SEPSIS-3 definition and the mortality rate was not available for this cohort [19]. In our dataset, mortality was highest when patients met both SEPSIS-1 and SEPSIS-3 definitions, highlighting

the high risk of death when infection causes end-organ dysfunction [1, 9]. On the other hand, the recent PROMISE trial in the UK, that recruited patients with severe sepsis and septic shock according to the SEPSIS-1 definition, reported 24.5% mortality at 28 days in the control arm, where patient characteristics were similar to our study [29]. We included all patients regardless of their 'do not attempt resuscitation' status, or with treatment limited to certain levels, and this could have affected mortality rates in our study.

Our results highlight the need for a simple, fast assessment tool to identify patients on the general wards with sepsis. In the UK, anaesthetists, will see many of these patients and evaluate them using more sophisticated clinical tools but enabling the ward staff to streamline these referrals is crucial to improve processes of care.

The strengths of our work include the use of robust, previously published data collection methodology tested over subsequent studies, and the wide participation of centres [7, 8]. We prospectively collected data on patients where the clinical teams suspected infection, hence we were able to test the real-life utility of the new sepsis definition and proposed clinical tools and compared its performance with the already implemented SEPSIS-1 criteria. Our study has high internal validity as in our subsequent trials using similar methodology we recruited similar number of patients with almost identical outcomes in the same hospitals [7, 30].

Our study has some limitations. Firstly, our dataset was a compromise between being an exhaustive list of possible determinants of sepsis using different definitions, and being small enough to maintain data collector participation and data reliability. Secondly, we followed our patients up for only 30 days and did not collect data on cause of death. Long-term quality of life survey and healthcare utilisation will be taken forward as part of a longitudinal study. Thirdly, based on the findings of others, it is possible that we could also have missed patients with sepsis who had NEWS below 3 (e.g. patients with high temperature and white cell count, but normal respiratory rate and heart rate) [23, 31]. However, recent data suggest that $\text{NEWS} \geq 3$ may be the best trigger to screen patients for sepsis in the ED [32]. Fourthly, laboratory elements of the

SOFA score were missing in a number of patients; serum bilirubin in 36.3%, serum creatinine in 7.9% and platelet count in 6.3% of the cases. It is possible that, due to these omissions, the number of patients with sepsis according to the SEPSIS-3 criteria is under-represented in our sample. Similarly, only 33.7% of patients had their lactate measured, possibly resulting in misrepresentation of the severe sepsis category.

In conclusion, the SEPSIS-3 definition identified patients with the highest risk if the full SOFA score was used; however, there was considerable overlap between the patients identified by the two definitions. SOFA and NEWS were found to be better predictors of poor outcome than qSOFA or SIRS in our population. These findings will have important implications for clinicians at the bedside and for organisations trying to improve the quality of sepsis care. For health-care systems with established track and trigger mechanisms, the optimal approach to integrating the new sepsis screening criteria with pre-existing escalation tools has yet to be determined.

Acknowledgements

This work was supported by the Fiona Elizabeth Agnew Trust and the Welsh Intensive Care Society, and they had no access to the data and no role in study design, conduct, analysis or drafting this report. No competing interests declared.

References

1. 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). *Journal of the American Medical Association* 2016; **315**: 801–10.
2. SepNet Critical Care Trials Group. Incidence of severe sepsis and septic shock in German intensive care units: the prospective, multicentre INSEP study. *Intensive Care Medicine* 2016; **42**: 1980–9.
3. Vincent J-L, Marshall JC, Namendys-Silva SA, Francois B, Martin-Loeches I, Lipman J, et al. Assessment of the worldwide burden of critical illness: the Intensive Care Over Nations (ICON) audit. *Lancet Respiratory Medicine* 2014; **2**: 380–6.
4. Rohde JM, Odden AJ, Bonham C, Kuhn L, Malani PN, Chen LM, et al. The epidemiology of acute organ system dysfunction from severe sepsis outside of the intensive care unit. *Journal of Hospital Medicine* 2013; **8**: 243–7.
5. Angus DC, Seymour CW, Coopersmith CM, Deutschman CS, Klompas M, Levy MM, et al. A framework for the development and interpretation of different sepsis definitions and clinical criteria. *Critical Care Medicine* 2016; **44**: e113–21.
6. 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. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. *Chest* 1992; **101**: 1644–55.
7. Szakmany T, Lundin RM, Sharif B, Ellis G, Morgan P, Kopczynska M, et al. Sepsis prevalence and outcome on the general wards and emergency departments in Wales: results of a multi-centre, observational, point prevalence study. *PLoS ONE* 2016; **11**: e0167230.
 8. Sharif B, Lundin RM, Morgan P, Hall JE, Dhadda A, Mann C, et al. Developing a digital data collection platform to measure the prevalence of sepsis in Wales. *Journal of the American Medical Informatics Association* 2016; **23**: 1185–9.
 9. Shankar-Hari M, Deutschman CS, Singer M. Do we need a new definition of sepsis? *Intensive Care Medicine* 2015; **41**: 909–11.
 10. 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). *Journal of the American Medical Association* 2016; **315**: 762–74.
 11. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. 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.
 12. McGinley A, Pearse RM. A national early warning score for acutely ill patients. *British Medical Journal* 2012; **345**: e5310.
 13. Arulkumaran N, Harrison DA, Brett SJ. Association between day and time of admission to critical care and acute hospital outcome for unplanned admissions to adult general critical care units: cohort study exploring the 'weekend effect'. *British Journal of Anaesthesia* 2017; **118**: 112–22.
 14. Aldridge C, Bion J, Boyal A, Chen Y-F, Clancy M, Evans T, et al. Weekend specialist intensity and admission mortality in acute hospital trusts in England: a cross-sectional study. *Lancet* 2016; **388**: 178–86.
 15. 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 Medicine* 1996; **22**: 707–10.
 16. Pandharipande PP, Shintani AK, Hagerman HE, St Jacques PJ, Rice TW, Sanders NW, et al. Derivation and validation of SpO₂/FiO₂ ratio to impute for PaO₂/FiO₂ ratio in the respiratory component of the Sequential Organ Failure Assessment score. *Critical Care Medicine* 2009; **37**: 1317–21.
 17. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, et al. A global clinical measure of fitness and frailty in elderly people. *Canadian Medical Association Journal* 2005; **173**: 489–95.
 18. Levy MM, Fink MP, Marshall JC, Abraham E, Angus D, Cook D, et al. 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. *Intensive Care Medicine* 2003; **29**: 530–8.
 19. Freund Y, Lemachatti N, Krastinova E, Van Laer M, Claessens Y-E, Avondo A, et al. Prognostic accuracy of Sepsis-3 criteria for in-hospital mortality among patients with suspected infection presenting to the emergency department. *Journal of the American Medical Association* 2017; **317**: 301–8.
 20. Dorsett M, Kroll M, Smith CS, Asaro P, Liang SY, Moy HP. qSOFA has poor sensitivity for prehospital identification of severe sepsis and septic shock. *Prehospital Emergency Care* 2017; **15**: 1–9.
 21. Finkelsztajn EJ, Jones DS, Ma KC, Pabón MA, Delgado T, Nakahira K, et al. Comparison of qSOFA and SIRS for predicting adverse outcomes of patients with suspicion of sepsis outside the intensive care unit. *Critical Care* 2017; **21**: 73.
 22. Giamarellos-Bourboulis EJ, Tsaganos T, Tsangaris I, Lada M, Routsi C, Sinapidis D, et al. Validation of the new Sepsis-3 definitions: proposal for improvement in early risk identification. *Clinical Microbiology and Infection* 2017; **23**: 104–9.
 23. 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. *American Journal of Respiratory and Critical Care Medicine* 2017; **195**: 906–11.
 24. Raith EP, Udy AA, Bailey M, McGloughlin S, MacIsaac C, Bel-lomo R, et al. Prognostic accuracy of the SOFA score, SIRS criteria, and qSOFA score for in-hospital mortality among adults with suspected infection admitted to the intensive care unit. *Journal of the American Medical Association* 2017; **317**: 290–300.
 25. Donnelly JP, Safford MM, Shapiro NI, Baddley JW, Wang HE. Application of the third international consensus definitions for sepsis (Sepsis-3) classification: a retrospective population-based cohort study. *Lancet Infectious Diseases* 2017; **17**: 661–70.
 26. Besen BAMP, Romano TG, Nassar AP, Taniguchi LU, Azevedo LCP, Mendes PV, et al. Sepsis-3 definitions predict ICU mortality in a low-middle-income country. *Annals of Intensive Care* 2016; **6**: 107.
 27. Hancock C. A national quality improvement initiative for reducing harm and death from sepsis in Wales. *Intensive and Critical Care Nursing* 2015; **31**: 100–5.
 28. Corfield AR, Lees F, Zealley I, Houston G, Dickie S, Ward K, et al. Utility of a single early warning score in patients with sepsis in the emergency department. *Emergency Medicine Journal* 2014; **31**: 482–7.
 29. Mouncey PR, Osborn TM, Power GS, Harrison DA, Sadique MZ, Grieve RD, et al. Trial of early, goal-directed resuscitation for septic shock. *New England Journal of Medicine* 2015; **372**: 1301–11.
 30. Szakmany T, Ellis G, Lundin RM, Pignatelli I, Sharif B, Joshi S, et al. Sepsis in Wales on the general wards: results of a feasibility pilot. *British Journal of Anaesthesia* 2015; **114**: 1000–1.
 31. Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. *Critical Care Medicine* 2001; **29**: 1303–10.
 32. Keep JW, Messmer AS, Sladden R, Burrell N, Pinate R, Tunnicliffe M, et al. National early warning score at Emergency Department triage may allow earlier identification of patients with severe sepsis and septic shock: a retrospective observational study. *Emergency Medicine Journal* 2016; **33**: 37–41.

Appendix

Membership of the Welsh Digital Data Collection Platform Collaborators is provided below: Steering committee: T. Szakmany (Chief Investigator); M. Kopczynska (National coordinator); R. Michael Lundin (National coordinator). B. Sharif (National coordinator).

Local coordinators: J. Abreu, S. Kulikouskaya, K. Bashir, L. Galloway, H. Al-Hassan, T. Grother,

P. McNulty, S. T. Seal, A. Cains, M. Vreugdenhil, M. Abdimalik, N. Dennehey, G. Evans, J. Whitaker, E. Beasant, C. Hall, M. Lazarou, C. V. Vanderpump, K. Harding, L. Duffy, A. G. Sadler, R. Keeling, C. Banks, S. W. Y. Ng, S. Y. Heng, D. Thomas, E. W. Puw, I. Otahal, C. Battle, O. Minik, R. Pugh, P. Morgan, G. Ellis, R. A. Lyons (Secure Anonymous Information Linkage (SAIL) representative), J. E. Hall, T. Canning (Independent patient representative).

Local data collectors: E. Walters, C. Durie, R. J. H. Sinnerton, B. Tanner, B. Cunningham-Walker, C. Spooner, A. Kiran, N. Nadeem, V. Unadkat, A. Poacher, S. Ganesananthan, J. N. C. Hui, E. Sparey, D. C. Li, J. Smith, I. Corrin, A. Kurani, H. Waring, A. Khan, C. Smith, N. Doyle, E. Baker, A. Shipley, M. Y. Amjad, M. Cynan, N. N. Azis, I. Hay, C. Russell, J. Davies, R. Parsonson, J. Joseph-Gubral, A. Arunthavarajah, J. Nicholas, A. Harris, J. Hale, H. Atkinson, J. Webster, T. Burnett, J. R. Gowar, S. DeFriend, D. Lawson, C. Maden, H. Jones, H. Preston, N. A. B. Maliki, M. Zimmerman, J. Webber, L. Jones, R. Phillips, L. McCarthy, L. Wirt, E. Hubbard, E. Evans, L. J. Davis, B. Atkins, L. W. Davies, L. Sanders-Crook, N. Vadivale, C. Lee, A. Dhadda, S. Cleaver, G. Logue,

J. Inns, I. Jones, R. Howcroft, C. Gilbert, M. Bradley, L. Pike, A. Urazbayeva, N. H. A. Hanif, Y. K. Ying, A. Coleclough, E. Higgins, L. Morgan, N. Spencer, T. G. Ng, S. Booth, N. Vignarajah, T. Chamberlain, D. Zhao, N. N. Tantirige, J. Watts, R. Walford, A. Prideaux, A. Tee, A. Hook, A. Mounce, E. Eccles, R. Edwards, K. Ramesh, L. Bausor, C. Lambrinudi, A. Sharma, A. Handley, A. Gordon, L. Allen, R. Paddock, H. Penney, L. Banerjee, L. D'Souza, K. Thomas, P. Havalda, C. Littler, N. West, C. Subbe, M. Hobrok, R. Self, V. Hamlyn

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Characteristics of patients identified by the different definitions.

Table S2. Diagnostic performances for the prediction of presence of organ dysfunction.

Figure S1. Ability of different clinical scoring systems to predict mortality.

Figure S2. Diagnostic capability of different clinical scoring systems to predict presence of organ dysfunction.