Quick sequential organ failure assessment versus systemic inflammatory response syndrome criteria for emergency department patients with suspected infection

Atsushi Shiraishi1,2*, Satoshi Gando1,3, Toshikazu Abe4,5, Shigeki Kushimoto6, Toshihiko Mayumi7, Seitaro Fujishima8, Akiyoshi Hagiwara9,10, Yasuazu Shino11, Shin-ichiro Shiraishi12, Toru Hifumi13, Yasuhiro Otomo14, Kohji Okamoto15, Junichi Sasaki16, Kiyotsugu Takuma17, Kazuma Yamakawa18, Yoshihiro Hanaki19, Masahiro Harada20 & Kazuma Morino21

Previous studies have shown inconsistent prognostic accuracy for mortality with both quick sequential organ failure assessment (qSOFA) and the systemic inflammatory response syndrome (SIRS) criteria. We aimed to validate the accuracy of qSOFA and the SIRS criteria for predicting in-hospital mortality in patients with suspected infection in the emergency department. A prospective study was conducted including participants with suspected infection who were hospitalised or died in 34 emergency departments in Japan. Prognostic accuracy of qSOFA and SIRS criteria for in-hospital mortality was assessed by the area under the receiver operating characteristic (AUROC) curve. Of the 1060 participants, 402 (37.9%) and 915 (86.3%) had qSOFA ≥ 2 and SIRS criteria ≥ 2 (given thresholds), respectively, and there were 157 (14.8%) in-hospital deaths. Greater accuracy for in-hospital mortality was shown with qSOFA than with the SIRS criteria (AUROC: 0.64 versus 0.52, difference + 0.13, 95% CI [+ 0.07, + 0.18]). Sensitivity and specificity for predicting in-hospital mortality at the given thresholds were 0.55 and 0.65 based on qSOFA and 0.88 and 0.14 based on SIRS criteria, respectively. To predict in-hospital mortality in patients visiting to the emergency department with suspected infection,
qSOFA was demonstrated to be modestly more accurate than the SIRS criteria albeit insufficiently sensitive.

Clinical Trial Registration: The study was pre-registered in the University Hospital Medical Information Network Clinical Trials Registry (UMIN000027258).

Sepsis is a health burden in various healthcare settings1, especially in emergency departments (EDs). Quick sequential (sepsis-related) organ failure assessment (qSOFA) is a prediction model for mortality following sepsis in patients suspected of having an infection outside the intensive care unit2,3. qSOFA was developed and validated in 2016 to replace the systemic inflammatory response syndrome (SIRS) criteria, that were originally designed to determine a systemic inflammatory syndrome. Earlier, SIRS criteria were prerequisite to determine whether the patient had sepsis based on the previous definitions of sepsis4,5; however, the performance of these criteria was reportedly poor for positive prediction, and insufficient for negative prediction6,7.

Of several studies involving patients with suspected infection in EDs to externally validate the prognostic accuracy of qSOFA compared with SIRS criteria in predicting mortality8–25 and systematic reviews26–28, inconsistency occurred in subsequent external validation results, with better prognostic accuracy of qSOFA for in-hospital mortality than for SIRS criteria. However, most of these studies were retrospective in nature or based on retrospective analyses of prospectively collected data8–10,22. These designs had multiple flaws, such as the use of the worst values of predictor variables during ED stay4–10 and the use of complete case analysis (excluding subjects with missing variables) for the index tests9,10,13,14,16–21,23,24.

Our multicentre prognostic study aimed to prospectively test the hypothesis that qSOFA could predict in-hospital mortality in patients with suspected infection with more accuracy than SIRS criteria using variables obtained at the time when a patient was first suspected of having an infection in the ED.

Results

Characteristics of study subjects. This study was discontinued in February 2018 after the number of participants reached 1060 following the recaluation of the required sample size at the interim analysis, which indicated a size of 439 participants. The study participants were mostly older individuals (median age 78, interquartile range [IQR] 68–85 years) with physical impairments (median clinical frailty scale score 4; IQR 3–6) and comorbidities (median Charlson comorbidity index 2; IQR 0–3) (Table 1). Distribution of the site of infection at the ED demonstrated that the most frequent site of infection was the respiratory tract (47.1%), followed by the abdomen (18.7%) and the urinary tract (14.2%).

Main results. Missing data for qSOFA, SIRS criteria, and both were found in 1 (0.1%), 408 (38.5%), and 7 (0.7%) participants, respectively (Fig. 1). No missing in-hospital mortality data were reported. In the multiply imputed population, 402 (37.9%) and 915 (86.3%) participants met the thresholds for qSOFA ≥ 2 and SIRS criteria ≥ 2, respectively. A total of 157 (14.8%) participants died in the participating hospitals. The primary analysis demonstrated greater diagnostic accuracy for in-hospital mortality with qSOFA than with SIRS criteria (area under the receiver operating characteristic [AUROC] curve 0.64 versus 0.52, difference + 0.13 95% confidence interval [CI] [+ 0.07, + 0.18]) (Table 2, Fig. 2). Sensitivity and specificity to predict in-hospital mortality at the given thresholds (qSOFA ≥ 2 and SIRS criteria ≥ 2) were 0.55 and 0.65 with qSOFA and 0.88 and 0.14 with SIRS criteria, respectively. The secondary analysis also demonstrated a positive net reclassification improvement (NRI) between qSOFA and SIRS criteria (+ 0.39 95% CI [+ 0.15, + 0.57]).

Sensitivity analysis of the naive dataset similarly demonstrated greater diagnostic accuracy for in-hospital mortality with qSOFA than with SIRS criteria (AUROC 0.64 versus 0.54, difference + 0.10 95% CI [+ 0.03, + 0.17]), where the sensitivity and specificity to predict in-hospital mortality at the given thresholds were 0.58 and 0.62 with qSOFA and 0.94 and 0.10 with SIRS criteria, respectively. Subgroup analyses did not demonstrate significant interactions with age, sex, comorbidities, and frailty (Fig. 3).

Discussion

A clinical prediction score is a mathematical model used to estimate the probability of a future event in patients with specific medical conditions. It is essential for a clinical prediction score in the ED to not only deliver a good prediction of the outcome but also provide fewer instances of missing data. The current study demonstrated that the simple qSOFA could provide a better prognostic value as well as better reclassification for in-hospital mortality in patients with suspected infection in the ED, with less frequent cases of missing data than with the SIRS criteria. However, prognostic accuracy of qSOFA (AUROC of 0.64) for mortality is nevertheless insufficient, especially in terms of sensitivity (0.55). Heterogeneity of patients’ baseline characteristics, including age, gender, comorbidity index, and frailty score, did not significantly interact with the prognostic ability of both qSOFA and SIRS criteria.

Studies comparing the prognostic accuracy of qSOFA and SIRS criteria in ED settings, except for a single prospective study3, have mostly been either purely retrospective in nature4,12,13,15–21,23–25 or have been retrospective studies based on prospectively collected data9,11,14,22. These studies had several flaws, including acquisition of predictor variables in wide time windows4–10 and/or the lack of assessment of computability of missing values9,10,13,14,16–21,23,24.

Practically, the qSOFA and SIRS criteria may be applied at the time of the initial evaluation after the patient’s visit to the ED. However, that several studies acquired the predictor variables within a wide time window meant that they could have obtained the worst values during the ED visits, which could bias the prognostic analyses through multiple measurements4–10. Use of the worst predicted values obtained with multiple measurements...
over time could lead to increased and decreased number of participants with positive and negative test results, respectively. Such results might improve the sensitivity and negative predictive value (fewer false negative) or worsen the specificity and positive predictive value (fewer true positive)\textsuperscript{15,29}. Furthermore, wide time windows may narrow the time between the prediction and outcome, which can lead to an apparent improvement in the positive prediction result\textsuperscript{30}. Finally, prolonged delay in the estimation of the score in the ED is inappropriate in the clinical settings. To avoid such biases and any others arising from multiple measurements, the current study used single baseline values obtained at the time the infection was first suspected.

Computation of qSOFA requires three variables that are readily available at the bedside (respiratory rate, systolic blood pressure, and the Glasgow coma scale [GCS] score). Thus, the score can be rapidly calculated, and the frequency of missing values is low. In contrast, computation of SIRS criteria is complicated and time consuming, and there is a greater likelihood of missing data as six variables consisting of bedside information (respiratory rate, heart rate, and body temperature) and laboratory data (partial carbon dioxide pressure, white blood cell count, and stab cell percentage) are required. In the current study, stab cell count data were missing for almost half of the study population, which led to an increase in the number of incomputable points (39%) and cases with undetermined positivity (12%) of SIRS criteria. However, for qSOFA, these percentages were 1% and 0%, respectively. Furthermore, the large number of missing stab cell count data also led to selection bias if patients with a score of 1 point for normal white blood cell count who lacked stab cell count data were excluded.

The current study estimated the prevalence of missing scores and compared the score performance based on both multiply imputed datasets and naive data, unlike previous studies, which did not sufficiently assess the missing scores\textsuperscript{9,10,13,14,16–21,23,24}.

To predict in-hospital mortality in ED patients with suspected infection, qSOFA was generally better and more specific, whereas SIRS criteria were generally worse and more sensitive based on the systematic reviews\textsuperscript{26–28}. However, a prominent inconsistency in the estimation of these diagnostic indices has been observed in previous ED studies\textsuperscript{4–25}. As discussed, biases in the acquisition of multiple predictor variables and statistical approaches,

| Variables | Before multiple imputation | After multiple imputation |
|-----------|---------------------------|--------------------------|
|           | Value                     | Missing, %               | Value                     | Missing, %               |
| N         | 1060                      | 0.0                      | 1060                      | 0.0                      |
| Age\textsuperscript{a}, years | 78 (68, 85)               | 0.0                      | 78 (68, 85)               | 0.0                      |
| Male sex\textsuperscript{b} | 633 (59.7)                | 0.0                      | 633 (59.7)                | 0.0                      |
| Clinical frailty score\textsuperscript{a} | 4 (3, 6)                  | 0.4                      | 4 (3, 6)                  | 0.0                      |
| Charlson comorbidity index\textsuperscript{a} | 2 (0, 3)                  | 0.0                      | 2 (0, 3)                  | 0.0                      |
| Site of infection\textsuperscript{a} | 0.0                      | 0.0                      |
| Respiratory tract | 499 (47.1)                | 0.0                      | 499 (47.1)                | 0.0                      |
| Urinary tract   | 151 (14.2)                | 0.0                      | 151 (14.2)                | 0.0                      |
| Abdomen        | 198 (18.7)                | 0.0                      | 198 (18.7)                | 0.0                      |
| Central nervous system | 14 (1.3)           | 0.0                      | 14 (1.3)                  | 0.0                      |
| Skin and soft tissue | 48 (4.5)                  | 0.0                      | 48 (4.5)                  | 0.0                      |
| Bone and joint | 9 (0.8)                   | 0.0                      | 9 (0.8)                   | 0.0                      |
| Wounds         | 4 (0.4)                   | 0.0                      | 4 (0.4)                   | 0.0                      |
| Catheter       | 3 (0.3)                   | 0.0                      | 3 (0.3)                   | 0.0                      |
| Endocardium    | 6 (0.6)                   | 0.0                      | 6 (0.6)                   | 0.0                      |
| Implant        | 2 (0.2)                   | 0.0                      | 2 (0.2)                   | 0.0                      |
| Other          | 57 (5.4)                  | 0.0                      | 57 (5.4)                  | 0.0                      |
| Unknown        | 69 (6.5)                  | 0.0                      | 69 (6.5)                  | 0.0                      |
| Respiratory rate, 1/min\textsuperscript{a} | 22 (18, 28)               | 0.8                      | 22 (18, 28)               | 0.0                      |
| Systolic blood pressure, mmHg\textsuperscript{a} | 126 (105, 149)             | 0.0                      | 126 (105, 149)            | 0.0                      |
| Heart rate, 1/min\textsuperscript{a} | 98 (84, 113)               | 0.0                      | 98 (84, 113)              | 0.0                      |
| Glasgow coma scale score\textsuperscript{a} | 14 (13, 15)               | 0.0                      | 14 (13, 15)               | 0.0                      |
| Body temperature, °C\textsuperscript{a} | 37.5 (36.7, 38.5)         | 0.2                      | 37.5 (36.7, 38.5)         | 0.0                      |
| White blood cell count, /µL\textsuperscript{a} | 10,950 (7500, 14,920)     | 0.0                      | 10,950 (7500, 14,920)     | 0.0                      |
| Stab cell count, %\textsuperscript{a} | 69 (11, 88)               | 66.8                     | 69 (11, 88)               | 66.8                     |
| Blood gas analysis\textsuperscript{a} | 1.9 (1.3, 3.2)             | 9.3                      | 1.9 (1.3, 3.1)            | 0.0                      |
| Lactate level, mmol/L | 37.2 (31.1, 44.1)        | 9.6                      | 37.2 (31.1, 44.1)         | 9.6                      |
| Carbon dioxide level, mmHg\textsuperscript{a} | 37.0 (31.1, 44.0)       | 0.0                      | 37.0 (31.1, 44.0)         | 0.0                      |

Table 1. Baseline characteristics of the cohort before and after multiple imputation. Data for all variables were obtained at the emergency department at the time when infection was suspected. \textsuperscript{a}Numeric variables are reported as median and 25th–75th percentile after pooling the values across the multiply imputed datasets. \textsuperscript{b}Nominal variables are displayed as number and percentages after averaging the counts across the multiply imputed datasets and rounding into integers.
which ignored missing values, might have led to differences in the estimation of specificity and sensitivity indices for these two methods. Furthermore, even in a prospective study, the authors retrospectively excluded patients with suspected infection at baseline who were later diagnosed with a non-infectious disease\(^9\). This retrospective exclusion might have led to the apparent improvement in diagnostic accuracy, similar to that in the per-protocol design\(^9\).

The strength of the current study is its prospective design, which eliminates possible biases in relation to multiple measurements, missing values, and per-protocol analysis. However, there are several limitations to the current study, which should be addressed in the future. First, this was a single national study in a developed country, which limits the generalisability of its results to other patients worldwide with suspected infection. In particular, the study participants were frail older patients with other comorbidities, which might make the results of our study inapplicable to younger populations. Second, even though this was a prospective study, a large percentage of missing stab cell result values prevented a realistic estimation for the SIRS criteria and required the use of multiple imputation. However, we believe the results could reflect those from a real-world setting.

**Conclusions**

This prospective, multi-centre study conducted for the external validation of qSOFA and SIRS criteria demonstrated that for patients in the ED who had suspected infection, qSOFA had modestly better prognostic accuracy in predicting in-hospital mortality albeit inadequate in sensitivity, and improved reclassification.

**Methods**

**Study design and setting.** In 2016, we designed the Sepsis Prognostication in Intensive Care Unit and Emergency Room (SPICE) study, a prospective observational study that consisted of two sub-studies—the SPICE-ER and SPICE-ICU—which were based in the ED and the intensive care unit, respectively. The current study is a primary study of the main SPICE-ER study, which involved a prospective prognostic analysis comparing qSOFA and SIRS criteria and externally validating the prognostic accuracy of these tools for in-hospital mortality among patients in the ED with suspected infection. The design and reporting of the study adheres to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD)
guidelines\textsuperscript{31}. The study included 34 EDs from 6 secondary and 28 tertiary emergency care centres and was conducted between December 2017 and February 2018. All procedures in these studies that involved human participants were performed in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. The study protocol for this observational study, including a waiver of the informed consent requirement, was first approved by the institutional review board of Hokkaido University (approval number 016-0385) and subsequently by the ethics committees of all participating hospitals.

Selection of participants. For this study, we included patients who visited the EDs of the participating hospitals; who were suspected of having an infection by the emergency physicians; who received any kind of antibiotics, underwent any fluid culture test, or underwent imaging for the detection of infection sites during ED stay; and who were hospitalised or died in the ED. Patients were excluded if they were transferred to another hospital without hospitalisation at the participating hospital.

Measurements. Index tests (qSOFA and SIRS criteria) were assessed using the data collected on clinical variables at the time when the infection was first suspected. Both qSOFA and SIRS criteria tested positive at a score $\geq 2$ point based on the original definitions of sepsis\textsuperscript{2,4}. GCS scores $< 15$ were used to satisfy the altered mental status criteria in qSOFA\textsuperscript{2}. The reference standard for the study was in-hospital mortality. The application of multiple imputation on all the study variables enabled 100% calculation of the study index tests and 100% assessment of associations between the index tests and the study outcome.

Outcomes. The baseline characteristics examined included the patients’ statuses before index infection, i.e., age, sex, clinical frailty scale score\textsuperscript{32}, and Charlson comorbidity index\textsuperscript{33}, as well as the clinical data obtained at the ED, i.e., from the suspected site of infection and on the physiological status (respiratory rate, heart rate, systolic blood pressure, the GCS score, body temperature, white blood cell count, stab cell percentage, lactate level, and partial pressure of carbon dioxide in the blood gas analysis). The suspected sites of infection were classified into 12 regions: the respiratory tract, urinary tract, abdomen, central nervous system, skin and soft tissue, bones and joints, wounds, intravascular catheter, endocardium, any kind of implant aside intravascular catheter, other regions, and unknown origin. The study outcome was defined as in-hospital mortality during ED stay or hospitalisation. All study data from the participating hospitals were entered electronically into the data capture server provided by the University Hospital Medical Information Network Internet Data and Information Center for Medical Research.

Statistical analysis. The statistical parameters required for estimating the sample size in this study were not fully available from previous publications; therefore, this study employed an adaptive design for sample size estimation. Initial sample size estimation was done using the receiver operating characteristics (ROC)
curve power calculation method based on the hypothesis that SIRS criteria significantly predicted in-hospital mortality. Parameters needed for the initial sample size estimation included AUROC curve for in-hospital mortality with an SIRS criteria score of 0.64, probability of in-hospital mortality of 0.04, power of 0.8, and significance level of 0.05. The required initial sample size was estimated at 807 participants but was modified to 900 participants considering the decline in statistical power owing to missing values. Interim analysis was pre-planned to determine the final study sample size based on two ROC curve power calculations to detect differences in the AUROCs of tested scores when the number of the study participants exceeded the initial sample size. The study was to be discontinued after the number of participants exceeded the estimated sample size at the interim analysis. Additionally, the study was also to be discontinued if the estimated sample size exceeded the upper limit of 2000 participants.

To compensate for the missing values, mainly in laboratory variables, multiple imputation by chained equations, which generated 25 multiply imputed datasets with 20 iterations of calculations, was used. The statistical analysis plan consisted of the comparison of AUROC (primary analysis) and NRI analysis (secondary analysis). Integration of the point estimations with 95% confidence intervals across each analysis on the multiply imputed dataset was based on bootstrapping. It was repeated 800 times per multiply imputed dataset to a total of 20,000 times.

In consideration of possible inconsistencies in the results before and after multiple imputation, a sensitivity analysis was conducted to assess the robustness of the results from primary analysis, using the naive dataset prior to multiple imputation instead of the multiply imputed datasets.
It was assumed that heterogeneity in the baseline characteristics of the participants might interact with the predicted scores for in-hospital mortality. A post-hoc subgroup analysis, dichotomised by age, sex, the clinical frailty scale score\cite{2}, and Charlson comorbidity index \cite{3} was used to assess the association between positivity of the tested scores and in-hospital mortality, which was reported as an odds ratio.

All statistical analyses were performed using R version 3.5.2 statistical software (The R Foundation for Statistical Computing, Vienna, Austria).

Ethics approval and consent to participate. The study was conducted after obtaining approval from the ethics committees of all participating hospitals. All procedures in these studies that involved human participants were performed in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. The requirement for informed consent was waived by the ethics committees because of the observational nature of the study.

Data availability
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 19 April 2020; Accepted: 17 February 2021
Published online: 05 March 2021

References
1. Fleischmann, C. et al. Assessment of global incidence and mortality of hospital-treated sepsis current estimates and limitations. Am. J. Respir. Crit. Care Med. 193, 259–272 (2016).
2. Seymour, C. W. et al. Assessment of clinical criteria for sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA 315, 762–774 (2016).
3. Singer, M. et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). JAMA 315, 801–810 (2016).
4. Bone, R. C. et al. American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference: Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. Crit. Care Med. 20, 864–874 (1992).
5. Levy, M. M. et al. 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. Crit. Care Med. 31, 1250–1256 (2003).
6. Churpek, M. M., Zadravec, F. J., Winslow, C., Howell, M. D. & Edelson, D. P. Incidence and prognostic value of the systemic inflammatory response syndrome and organ dysfunctions in ward patients. Am. J. Respir. Crit. Care Med. 192, 958–964 (2015).
7. Kaukonen, K. M., Bailey, M., Pilcher, D., Cooper, D. J. & Bellomo, R. Systemic inflammatory response syndrome criteria in defining severe sepsis. N. Engl. J. Med. 372, 1629–1638 (2015).
8. Williams, J. M. et al. Systemic inflammatory response syndrome, quick Sequential Organ Function Assessment, and organ dysfunction: Insights from a prospective database of ED patients with infection. Chest 151, 586–596 (2017).
9. Freund, Y. et al. Prognostic accuracy of sepsis-3 criteria for in-hospital mortality among patients with suspected infection presenting to the emergency department. JAMA 317, 301–308 (2017).
10. Henning, D. J. et al. An emergency department validation of the SEP-3 sepsis and septic shock definitions and comparison with 1992 consensus definitions. Ann. Emerg. Med. 70, 544–552.e5 (2017).
11. Askim, Å. et al. Poor performance of quick-SOFA (qSOFA) score in predicting severe sepsis and mortality—a prospective study of patients admitted with infection to the emergency department. Scand. J. Trauma Resusc. Emerg. Med. 25, 56 (2017).
12. Park, H. K., Kim, W. Y., Kim, M. C., Jung, W. & Ko, B. S. Quick sequential organ failure assessment compared to systemic inflammatory response syndrome for predicting sepsis in emergency department. J. Crit. Care 42, 12–17 (2017).
13. Haydar, S., Spaniel, M., Weems, P., Wood, S. & Stout, T. Comparison of QSOFA score and SIRS criteria as screening mechanisms for emergency department sepsis. Am. J. Emerg. Med. 35, 1730–1733 (2017).
14. González Del Castillo, J. et al. Prognostic accuracy of SIRS criteria, qSOFA score and GYM score for 30-day-mortality in older non-severely dependent infected patients attended in the emergency department. Eur. J. Clin. Microbiol. Infect. Dis. 36, 2361–2369 (2017).
15. Moskowitz, A. et al. Quick sequential organ failure assessment and systemic inflammatory response syndrome criteria as predictors of critical care intervention among patients with suspected infection. Crit. Care Med. 45, 1813–1819 (2017).
16. Tugsl, S., Carron, P. N., Yersin, B., Calandra, T. & Dami, F. Low sensitivity of qSOFA, SIRS criteria and sepsis definition to identify infected patients at risk of complication in the prehospital setting and at the emergency department triage. Scand. J. Trauma Resusc. Emerg. Med. 25, 108 (2017).
17. Ranzan, O. T. et al. New sepsis definition (Sepsis-3) and community-acquired pneumonia mortality. A validation and clinical decision-making study. Am. J. Respir. Crit. Care Med. 196, 1287–1297 (2017).
18. Müller, M. et al. Utility of quick sepsis-related organ failure assessment (qSOFA) to predict outcome in patients with pneumonia. PLoS One 12, e0188913 (2017).
19. Goudien, B. et al. qSOFA, SIRS and NEWS for predicting inhospital mortality and ICU admission in emergency admissions treated as sepsis. Emerg. Med. J. 35, 345–349 (2018).
20. van der Woude, S. W., van Doornmaal, F. F., Hutten, B. A., Nellen, F. J. & Hollemann, F. Classifying sepsis patients in the emergency department using SIRS, qSOFA or MEWS. Neth. J. Med. 76, 158–166 (2018).
21. Martino, I. F. et al. The role of qSOFA compared to other prognostic scores in septic patients upon admission to the emergency department. Eur. J. Intern. Med. 53, e11–e13 (2018).
22. Boillat-Blanco, N. 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 13, e0197982 (2018).
23. Chen, F. C. et al. Quick sepsis-related organ failure assessment predicts 72-h mortality in patients with suspected infection. Eur. J. Emerg. Med. 26, 323–328 (2019).
24. Delahanty, R. J., Alvarez, J. A., Flynn, L. M., Sherwin, R. L. & Jones, S. S. Development and evaluation of a machine learning model for the early identification of patients at risk for sepsis. Ann. Emerg. Med. 73, 334–344 (2019).
25. Brink, A. et al. Predicting mortality in patients with suspected sepsis at the emergency department; a retrospective cohort study comparing qSOFA, SIRS and National Early Warning Score. PLoS One 14, e0211133 (2019).
26. Fernando, S. M. et al. Prognostic accuracy of the quick Sequential Organ Failure Assessment for mortality in patients with suspected infection: A systematic review and meta-analysis. Ann. Intern. Med. 168, 266–275 (2018).
Acknowledgements
The authors acknowledge all patients and hospitals who participated in the SPICE-ER study. Names of the hospitals in Japan that participated in the SPICE-ER study along with the number of patients who participated in each of the hospitals in parentheses are as follows: Public Toyooka Hospital, Tajima Emergency and Critical Care Medical Center (128), Kameda Medical Center (92), Juntendo University Urayasu Hospital (83), Juntendo University Nerima Hospital (83), Japanese Red Cross Nagoya Daichi Hospital (80), Center Hospital of the National Center for Global Health and Medicine (68), Aizu Chuo Hospital (66), Kawasaki Medical School Hospital (52), National Hospital Organization Kumamoto Medical Center (49), Kagoshima City Hospital (47), Kawasaki Municipal Kawasaki Hospital (44), Tokyo Medical and Dental University Hospital (38), Yamagata Prefectural Central Hospital (35), Kyoto Okamoto Memorial Hospital (23), Fukuoka University Hospital (17), Kagawa University Hospital (13), Tokyo Women's Medical University Medical Center East (13), Gifu University Hospital (13), Hokkaido University Graduate School of Medicine (12), Kitakyushu City Yahata Hospital (11), Ome Municipal General Hospital (10), Keio University Hospital (10), Saiseikai Kumamoto Hospital (10), Saku Central Hospital (10), Osaka General Medical Center (9), Saitama Red Cross Hospital (9), Fujieda Municipal General Hospital (9), Kyoto University Hospital (7), Kansai Medical University Hospital (7), St. Mary's Hospital (4), University of Occupational and Environmental Health Hospital (4), Fujisawa City Hospital (3), Kumamoto University Hospital (2), Osaka Mishima Emergency Critical Care Center (2) and Seirei Yokohama Hospital (1).

Author contributions
A.S. contributed to the conception and design of the work, literature research, data analysis and interpretation, drafting the article, and final approval of the article. S.G., S.K., and T.M. contributed to the conception and design of the work, data interpretation, critical revision of the article, and final approval of the article. T.A. contributed to the conception and design of the work, data interpretation, drafting the article, and final approval of the article. S.F., A.H., Y.S., T.H., Y.O., K.O., J.S., K.T., K.Y., Y.H., M.H., and K.M. contributed to the data interpretation, critical revision of the article, and final approval of the article.

Funding
The SPICE-ER study and the current study were funded by the Japanese Association for Acute Medicine (JAAM). The funders (except for JAAM itself) had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Competing interests
Atsushi Shiraishi has received personal fees from CSL Behring. Satoshi Gando has received personal fees from Asahi Kasei Pharma America Inc. and Asahi Kasei Pharma Japan Inc. The remaining authors have disclosed that they have no conflicts of interest.

Additional information
Correspondence and requests for materials should be addressed to A.S.

Reprints and permissions information is available at www.nature.com/reprints.
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021