

Screening strategies to identify sepsis in the prehospital setting: a validation study

Daniel J. Lane PhD, Hannah Wunsch MD MSc, Refik Saskin MS, Sheldon Cheskes MD, Steve Lin MDCM MSc, Laurie J. Morrison MD MSc, Damon C. Scales MD PhD

■ Cite as: *CMAJ* 2020 March 9;192:E230-9. doi: 10.1503/cmaj.190966

ABSTRACT

BACKGROUND: In the prehospital setting, differentiating patients who have sepsis from those who have infection but no organ dysfunction is important to initiate sepsis treatments appropriately. We aimed to identify which published screening strategies for paramedics to use in identifying patients with sepsis provide the most certainty for prehospital diagnosis.

METHODS: We identified published strategies for screening by paramedics through a literature search. We then conducted a validation study in Alberta, Canada, from April 2015 to March 2016. For adult patients (≥ 18 yr) who were transferred by ambulance, we linked records to an administrative database

and then restricted the search to patients with infection diagnosed in the emergency department. For each patient, the classification from each strategy was determined and compared with the diagnosis recorded in the emergency department. For all strategies that generated numeric scores, we constructed diagnostic prediction models to estimate the probability of sepsis being diagnosed in the emergency department.

RESULTS: We identified 21 unique prehospital screening strategies, 14 of which had numeric scores. We linked a total of 131 745 eligible patients to hospital databases. No single strategy had both high sensitivity (overall range 0.02–

0.85) and high specificity (overall range 0.38–0.99) for classifying sepsis. However, the Critical Illness Prediction (CIP) score, the National Early Warning Score (NEWS) and the Quick Sepsis-Related Organ Failure Assessment (qSOFA) score predicted a low to high probability of a sepsis diagnosis at different scores. The qSOFA identified patients with a 7% (lowest score) to 87% (highest score) probability of sepsis diagnosis.

INTERPRETATION: The CIP, NEWS and qSOFA scores are tools with good predictive ability for sepsis diagnosis in the prehospital setting. The qSOFA score is simple to calculate and may be useful to paramedics in screening patients with possible sepsis.

Sepsis is a syndrome characterized by life-threatening organ dysfunction due to a dysregulated host response to infection.¹ Early identification and prompt intervention are critical to improving outcomes for patients with sepsis.^{2,3} Paramedics are the first to evaluate and manage most patients with sepsis,⁴ often for an extended period before arrival in the emergency department. However, in the prehospital setting, without access to laboratory results, it can be challenging to differentiate patients who have sepsis from those who have infection without organ dysfunction.

The potential for paramedics to contribute to the early identification of sepsis using clinical signs and symptoms has been discussed but seldom rigorously studied.^{5–8} Studies that propose screening strategies for identification of sepsis by paramedics are frequently limited by incomplete prehospital measurements, small sample size or the use of convenience

samples comprising only patients with a diagnosis of sepsis made in the hospital.⁵ Furthermore, these studies have often relied solely on measures of test accuracy that depend on the known diagnosis of sepsis (sensitivity and specificity), which are sensitive to spectrum bias due to underlying disease severity.⁹ They also require that a threshold be established to define a positive versus negative test result, which may conceal the diagnostic information in individual test results that is more relevant to clinical decision-making for individual patients in different settings.¹⁰

To determine which approach to screening for sepsis is optimal in the prehospital setting, we completed a validation of the accuracy and predictive ability of published approaches for identification of patients with sepsis within a large cohort of patients with suspected infection who were transported by emergency medical services.

Methods

Identification of screening strategies

We re-ran the search strategy from our previously published systematic review⁵ to find additional screening strategies for identification of infection or sepsis in the prehospital setting (search dates Oct. 1, 2015, to July 22, 2019). The search strategy and methods are described in Appendix 1, Part A1-1 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190966/-/DC1).

Study design

In this 1-year validation study, we compared the classification and predictive ability of published prehospital sepsis screening strategies applied to a cohort of patients with a single reference standard for diagnosis of sepsis. We used the STROBE¹¹ and RECORD¹² statements to guide reporting.

Study population and setting

Records for all adult patients (age \geq 18 yr) transported between Apr. 1, 2015, and Mar. 31, 2016, by a large provincial emergency medical service in Alberta were deterministically linked to a population-based emergency administrative database (National Ambulatory Care Reporting System) and an inpatient database (Discharge Abstract Database) by analysts in the emergency medical service using each patient's unique health number, birth date, time and initial destination for patient transport.

Sepsis should be considered when clinicians have a strong suspicion of infection;¹ therefore, we assembled a cohort consisting only of patients who had an infection diagnosed in the emergency department. We identified these patients using previously validated diagnosis codes for use in the emergency department consistent with a bacterial or fungal infection¹³ (as listed in Appendix 2, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190966/-/DC1). We also assembled a subcohort of patients for whom paramedics documented a suspected infection (i.e., real-world application). Paramedic-suspected infection was determined by examining several fields, including those for the chief complaint (patient perspective) and the provider impression (paramedic perspective), for terms consistent with infection (e.g., cellulitis, urinary tract infection, pneumonia), sepsis, cough or flu-like symptoms. We also included patients for whom paramedics had selected the sepsis treatment protocol.

For patients with multiple transports or emergency admissions in a single day, we retained only the initial emergency medical service record and the final emergency department record, because these were the most complete records. Patients who were admitted to the initial destination hospital were also linked to an inpatient administrative database (the Discharge Abstract Database of the Canadian Institute for Health Information) to determine in-hospital disposition (admission to the intensive care unit, mechanical ventilation, length of stay and death).

Variables

We extracted from the emergency medical service records all measured patient characteristics (age, weight, vital signs), the documented physical examination findings (including Glasgow

Coma Scale score¹⁴) and operational characteristics (date and time stamps). For each characteristic, the first available measure was used for evaluation of all screening strategies, as we hypothesized that initial measurements were least likely to be influenced by medical intervention and most likely to inform subsequent care by paramedics. These measures are entered directly into the patient care record during the patient encounter or are imported from a monitoring device and verified by the paramedics before they leave the hospital.

Application of screening strategies

We applied the prehospital screening strategies identified in our updated search to our cohort (all patients with confirmed infection) using the recommended measures, which resulted in a “positive” or “negative” screening result for each strategy for each patient. For strategies with a numeric score, we determined both the numeric score and the screening result based on the recommended threshold. We conducted sensitivity analyses to evaluate the ability of each approach to identify sepsis in the entire population of transported patients and also in the subcohort of patients for whom paramedics documented a suspected infection.

Outcome measure

The primary goal of this study was to compare the ability of the screening strategies to identify the outcome of sepsis. We identified cases of sepsis using a strategy based on the Canadian version of the *International Statistical Classification of Diseases and Related Health Problems, 10th revision*, validated for the 2012 Surviving Sepsis definition,¹⁵ modified to align with the Sepsis-3 definition. This strategy identified patients as having sepsis if they were diagnosed with infection in the emergency department and were found to have organ dysfunction characteristic of sepsis. Organ dysfunction was identified from diagnostic codes or altered vital signs consistent with organ dysfunction (identified on the basis of abnormalities in documented pulse oximetry, mean arterial pressure or Glasgow Coma Scale score that would be consistent with a sequential organ failure assessment score of 2 or greater¹). We excluded patients who were discharged from the emergency department. This approach was found to be reliable, and it had good criterion and construct validity for identifying patients with sepsis in the emergency department.¹³

Statistical analysis

Within the cohort of patients with infection diagnosed in the emergency department, we assessed diagnostic accuracy by calculating sensitivity, specificity and the corresponding positive and negative predictive values according to the result from each screening strategy. We assessed the predictive ability of strategies with a numeric score using diagnostic prediction models, with diagnosis of sepsis in the emergency department as the binary outcome.¹⁶ Each patient's score was calculated and included in the model, along with their age and sex (if not already a component of the strategy) to adjust for nonrandom differences in these variables within our population. We addressed the possibility of dependence between observations

due to clustering by destination hospital using the Huber-White robust covariance matrix estimates.¹⁷ We created a visual comparison of the predicted probabilities from each score, representing a patient's probability of having sepsis from the minimum to maximum level of each score. We assessed discrimination with the C statistic, and we assessed calibration visually with calibration plots (Appendix 3, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190966/-/DC1).^{18,19} Probability estimates for sepsis diagnosis with 95% confidence intervals are reported here. We used normal-value imputation of missing values for both analyses. We also completed a sensitivity analysis that excluded patients for whom any measure was missing.

All statistical analyses were completed in R statistical and computing software. We used the “tableone” package for descriptive statistics, the Quan method (using the “icd” package)

for calculating Charlson scores²⁰ and the “rms” package for constructing prediction models.^{17,21,22}

Ethics approval

This study was reviewed and approved under a waiver of informed consent by the University of Calgary Conjoint Health Research Ethics Board and the University of Toronto Health Science Research Ethics Boards.

Results

Screening strategies to identify sepsis

We identified 32 studies,^{23–54} 1 abstract⁵⁵ and 3 ongoing registered studies^{56–58} that described 21 unique screening strategies for the identification of sepsis in the prehospital setting (Table 1). Among these strategies, 14 had a numeric score. All of the

Table 1: Criteria included in screening strategies for identification of sepsis by paramedics

Strategy*	Presence/absence in strategy						
	Temperature	Heart rate	Systolic blood pressure	Respiratory rate	Oxygen saturation	Glasgow Coma Scale	Other criteria
BAS 90–30–90 ^{53†}			X	X	X		
Borrelli et al. ^{29†}	X	X	X	X	X	X	
CIP ^{40,43†}		X	X	X	X	X	Age
HEWS ^{47†}	X	X	X	X	X	X	
MBIS ⁵⁵	X	X	X			X	
MEWS ^{37,42†}	X	X	X	X		X	
NEWS ^{45†}	X	X	X	X	X	X	
PHANTASi ²⁴	X	X		X			
PITSTOP ⁵⁶	X		X				
PreSAT ^{52†}	X	X	X	X			
PRESEP ^{28†}	X	X	X	X	X		
PRESS ^{41†}	X		X		X		Age, dispatch card‡
PSP ^{26†}	X	X	X	X			
qSOFA ^{27,30,33,36,38–40,44,46,51†}			X	X		X	
qSOFA + end-tidal carbon dioxide ³⁵			X	X		X	End-tidal carbon dioxide
Robson score ^{37,53†}	X	X		X		X	Blood glucose level
SEPSIS ^{48†}	X	X	X	X	X	X	Age, appearance of skin§
Sepsis Alert ³²	X	X	X	X			Mean arterial pressure
SIRS ^{23,25,27,33,39,42,50,54†}	X	X		X			
SIRS + end-tidal carbon dioxide ³⁴	X	X		X			End-tidal carbon dioxide
Suffoletto et al. ⁴⁹	X		X				

Note: BAS = Blood Pressure *Andningsfrekvens* (meaning “respiratory rate” in Swedish) Saturation, CIP = Critical Illness Prediction score, HEWS = Hamilton Early Warning Score, MBIS = Mecklenburg Bacterial Infection Scale, MEWS = Modified Early Warning Score, NEWS = National Early Warning Score, PHANTASi = Prehospital Antibiotics Against Sepsis, PITSTOP = Paramedic Initiated Treatment of Sepsis Targeting Out-of-Hospital Patients clinical trial, PreSAT = Prehospital Sepsis Assessment Tool, PRESEP = Prehospital Early Sepsis Detection, PRESS = Prehospital Severe Sepsis, PSP = Prehospital Sepsis Project, qSOFA = Quick Sepsis-Related Organ Failure Assessment, SEPSIS = Screening to Enhance Prehospital Identification of Sepsis, SIRS = Systemic Inflammatory Response Syndrome.

*For more detailed information about the components of each strategy, see Appendix 1, Table A1-1 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190966/-/DC1).

†Strategies that produce a numeric score.

‡Dispatch category specified as “sick person.”

§Mottled or ashen appearance or cyanosis of the skin, lips or tongue; nonblanching rash of the skin.

strategies included more than 1 patient measure, whereas 3 of the strategies (Critical Illness Prediction [CIP], Screening to Enhance Prehospital Identification of Sepsis [SEPSIS] and Prehospital Severe Sepsis [PRESS]) also included patient characteristics other than measured patient values. A detailed description of the components of each screening strategy is available in Appendix 1, Table A1-1.

Patients

Of 146 626 adult patients transported during the study period, 131 745 (90%) were successfully linked to hospital databases. The most common reasons for linkage failure were lack of a unique health number documented on the emergency medical service report (36%) and inability to match emergency medical service and hospital records (38%). Among the patients with successful linkage, 12 740 had infection and therefore constituted our primary cohort; for 2740 (22%) of these, sepsis was diagnosed in the emergency department (Figure 1). The proportion of patients with missing values was low (range 0%–4%) for most vital signs, but was high for blood glucose level (24%) and end-tidal carbon dioxide (86%). Characteristics of the patients are presented in Table 2.

Accuracy of classification

We observed considerable variation in classification accuracy among the screening strategies. None of the screening strategies was both highly sensitive and highly specific for the diagnosis of sepsis. Positive predictive values ranged from 0.23 to 0.68, with only 6 strategies having positive predictive values above 0.5 (Table 3). When we excluded patients with missing measures, strategies for which measures were missing for a large proportion of patients (e.g., end-tidal carbon dioxide) had increased sensitivity and decreased specificity (Appendix 4, Table A4-1, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190966/-/DC1). Classification accuracy for the entire cohort of patients for whom linkage was successful ($n = 131\,745$) is presented in Appendix 4, Table A4-2.

The findings of our sensitivity analysis to evaluate the ability of each approach to identify sepsis in the entire population of transported patients is presented in Appendix 4, Table A4-2 and Table A4-3. In the sensitivity analysis conducted in the subcohort of 4138 patients for whom paramedics documented a suspected infection, 420 (10%) had sepsis diagnosed in the emergency department. Relative to the primary analysis, the positive predictive values for all screening strategies decreased (Appendix 4, Table A4-4).

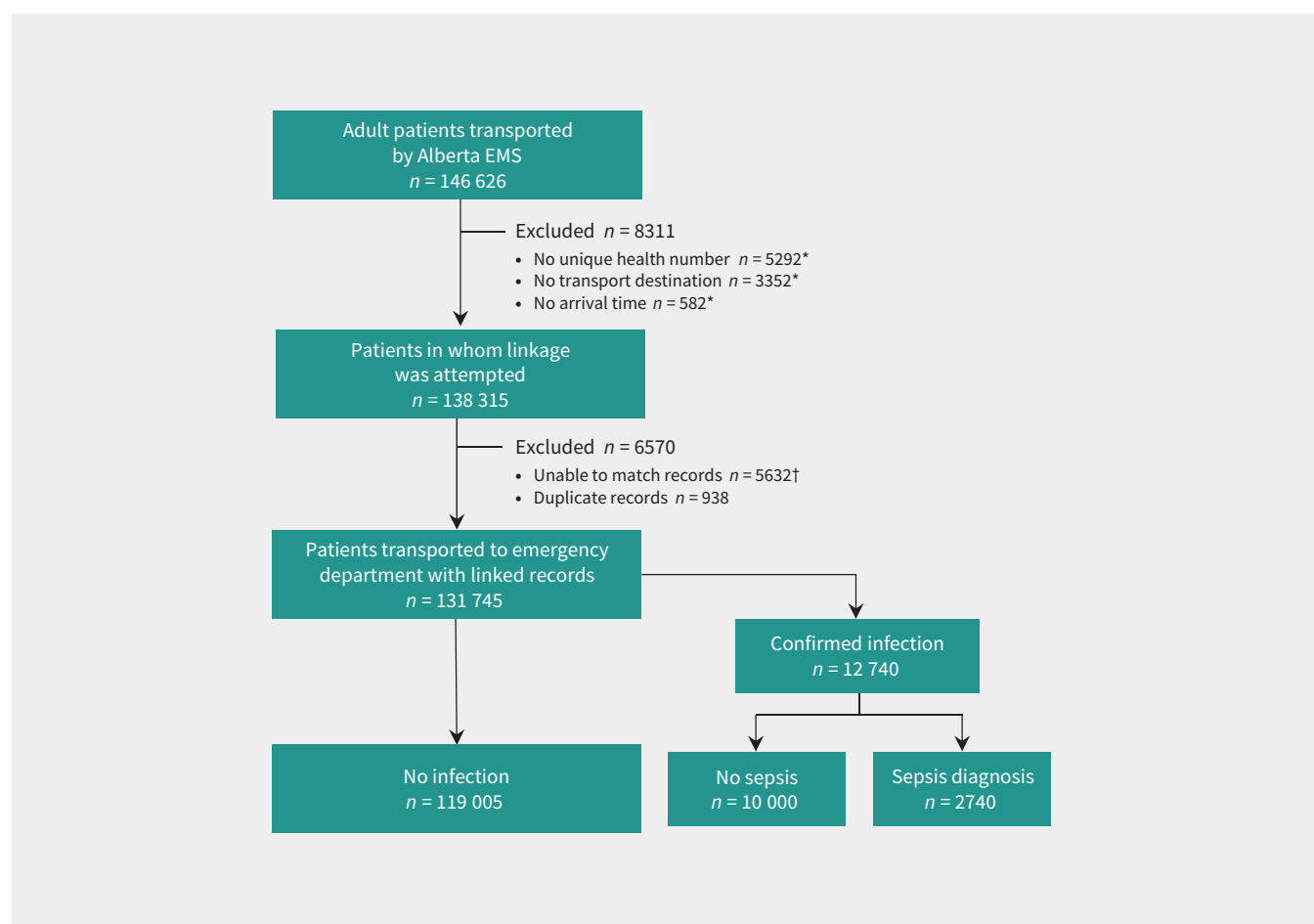


Figure 1: Study flow diagram. Note: EMS = emergency medical services. *These frequencies are not mutually exclusive. †For the 5632 patients with unmatched records, age and sex were as follows: 3667 patients were 18–65 years of age, and 1965 patients were older than 65 years; 3496 patients were women, and 2136 were men.

Table 2 (part 1 of 2): Characteristics of patients who were transported by emergency medical services

Characteristic	Group; no (%) of patients*		Standardized difference	% with missing data
	No sepsis n = 10 000	Sepsis n = 2740		
Prehospital operational characteristics				
Dispatch priority			0.34	0.01
Omega/alpha (lowest priority)	3300 (33)	525 (19)		
Bravo/charlie	3998 (40)	1172 (43)		
Delta/echo (highest priority)	2695 (27)	1043 (38)		
Unit type			0.004	0
Advanced life support	8886 (89)	2438 (89)		
Basic life support	1114 (11)	302 (11)		
Emergent (i.e., lights and sirens) transport to hospital	369 (4)	460 (17)	0.44	0.3
Response time, min, median (IQR)	7.0 (5.0–10)	6.0 (4.0–9.3)	0.13	2.2
Prehospital time, min, median (IQR)	39 (31–50)	41 (34–52)	0.14	0
Transport distance, km, median (IQR)	14 (9.0–17)	12 (9.0–16)	0.05	0
Prehospital patient characteristics				
Sex			0.087	0
Male	4447 (44)	1337 (49)		
Female	5553 (56)	1403 (51)		
Age, yr, mean ± SD	71 ± 18	75 ± 15	0.23	0
Weight, kg, mean ± SD	79 ± 26	79 ± 27	0.02	38
Temperature, °C, mean ± SD	37.0 ± 0.93	37.1 ± 1.2	0.09	7.8
Heart rate, beats/min, mean ± SD	94 ± 23	100 ± 26	0.24	0.7
Respiratory rate, breaths/min, mean ± SD	21 ± 6.2	24 ± 10	0.41	2.1
Oxygen saturation, %, median (IQR)	95 (92–97)	91 (84–95)	0.75	1.2
End-tidal carbon dioxide, mm Hg, mean ± SD	31 ± 15	28 ± 13	0.26	86
Systolic blood pressure, mm Hg, mean ± SD	138 ± 27	126 ± 31	0.42	0.6
Diastolic blood pressure, mm Hg, mean ± SD	79 ± 18	72 ± 21	0.37	0.8
Mean arterial pressure, mm Hg, mean ± SD	99 ± 19	89 ± 22	0.43	15
Blood glucose level, mmol/L, mean ± SD	7.9 ± 3.69	8.4 ± 4.3	0.12	24
Glasgow Coma Scale score, ¹⁴ mean ± SD	15 ± 0.83	13 ± 2.9	0.92	3.7
CIP score, [†] median (IQR)	3 (1–5)	6 (4–9)	0.99	0
IV initiated by EMS personnel	3773 (38)	1467 (54)	0.43	0
Total volume of IV fluid administered, [‡] mL, median (IQR)	400 (238–500)	400 (250–500)	0.12	19
Weight-based volume of IV fluid administered (mL/kg), mean ± SD	5.85 ± 4.97	6.28 ± 5.23	0.08	19
Hospital characteristics				
CTAS score in emergency department			0.75	7.7
1 (most urgent)	137 (2)	366 (14)		
2	2999 (33)	1389 (52)		
3	4759 (52)	831 (31)		
4	1081 (12)	99 (4)		
5 (least urgent)	92 (1)	6 (< 1)		
CTAS score, median (IQR)	3 (2–3)	2 (2–3)	0.785	7.7
Time to triage, min, median (IQR)	11 (8.0–15)	10 (7.5–14)	0.12	9.8

Table 2 (part 2 of 2): Characteristics of patients who were transported by emergency medical services

Characteristic	Group; no (%) of patients*			
	No sepsis n = 10 000	Sepsis n = 2740	Standardized difference	% with missing data
Time to physician assessment, min, median (IQR)	77 (37–137)	44 (16–96)	0.38	11
Total time in emergency department, min, median (IQR)	555 (347–1035)	700 (445–1287)	0.19	0
Patient admitted to inpatient care	5775 (58)	2614 (95)	1.8	0
Patient transferred to another hospital	254 (2)	90 (3)	0.04	0
Charlson comorbidity index, median (IQR)	1 (0–2)	1 (0–3)	0.15	
Mechanical ventilation required	122 (1)	254 (10)	0.38	0
Patient admitted to ICU	266 (3)	356 (13)	0.39	0
Total ICU time,§ h, median (IQR)	116 (58–216)	117 (49–209)	0.18	0
Total length of stay, d, median (IQR)	7 (4–13)	7 (4–15)	0.04	0
In-hospital death	427 (5)	528 (28)	0.66	18

Note: CIP = Critical Illness Prediction score, CTAS = Canadian Triage Acuity Scale,⁵⁹ EMS = emergency medical services, ICU = intensive care unit, IQR = interquartile range, IV = intravenous, Prehospital time = time from EMS arrival on scene to arrival at hospital, Response time = time from call to EMS arrival on scene, SD = standard deviation.
*Except where indicated otherwise.
†This score was previously validated as a measure of illness severity in patients transported by paramedics.⁶⁰
‡Among patients who received fluid by IV administration.
§Among patients who were admitted to the ICU.

Table 3: Diagnostic classification accuracy of screening strategies for prehospital identification of sepsis, based on published thresholds for patients with infection diagnosed in the emergency department

Strategy* (score threshold)	Sensitivity (95% CI)	Specificity (95% CI)	PPV	NPV
Sepsis Alert	0.07 (0.06–0.08)	0.99 (0.99–0.99)	0.68	0.79
qSOFA (2)	0.40 (0.38–0.42)	0.94 (0.94–0.95)	0.66	0.85
PITSTOP	0.02 (0.02–0.03)	1.00 (1.00–1.00)	0.65	0.79
qSOFA + end-tidal carbon dioxide	0.44 (0.42–0.46)	0.92 (0.91–0.92)	0.60	0.86
PRESS (2)†	0.11 (0.10–0.12)	0.98 (0.97–0.98)	0.56	0.80
SEPSIS (5)	0.26 (0.24–0.27)	0.94 (0.94–0.95)	0.55	0.82
BAS (1)	0.57 (0.55–0.59)	0.79 (0.78–0.80)	0.43	0.87
Borrelli strategy (3)	0.49 (0.47–0.51)	0.86 (0.86–0.87)	0.50	0.86
MEWS (4)	0.53 (0.51–0.55)	0.77 (0.77–0.78)	0.39	0.86
PRESEP (4)	0.49 (0.47–0.51)	0.76 (0.75–0.77)	0.36	0.84
MBIS	0.44 (0.42–0.46)	0.77 (0.76–0.77)	0.34	0.83
PSP (2)	0.42 (0.40–0.44)	0.77 (0.76–0.78)	0.33	0.83
PreSAT (2)	0.49 (0.47–0.50)	0.71 (0.70–0.72)	0.32	0.83
PHANTASi	0.20 (0.19–0.22)	0.88 (0.87–0.89)	0.32	0.80
Robson score (2)	0.75 (0.73–0.76)	0.54 (0.53–0.55)	0.31	0.89
SIRS (2)	0.45 (0.43–0.47)	0.72 (0.71–0.73)	0.31	0.83
HEWS (2)	0.85 (0.84–0.86)	0.41 (0.40–0.42)	0.28	0.91
SIRS + end-tidal carbon dioxide	0.74 (0.73–0.76)	0.40 (0.39–0.41)	0.25	0.85
Suffoletto strategy	0.70 (0.68–0.71)	0.38 (0.37–0.39)	0.23	0.82

Note: BAS = Blood Pressure *Andningsfrekvens* ("respiratory rate" in Swedish) Saturation, CI = confidence interval, HEWS = Hamilton Early Warning Score, MBIS = Mecklenburg Bacterial Infection Scale, MEWS = Modified Early Warning Score, NPV = negative predictive value, PHANTASi = Prehospital Antibiotics Against Sepsis, PITSTOP = Paramedic Initiated Treatment of Sepsis Targeting Out-of-Hospital Patients clinical trial, PPV = positive predictive value, PreSAT = Prehospital Sepsis Assessment Tool, PRESEP = Prehospital Early Sepsis Detection, PRESS = Prehospital Severe Sepsis, PSP = Prehospital Sepsis Project, qSOFA = Quick Sepsis-Related Organ Failure Assessment, SEPSIS = Screening to Enhance Prehospital Identification of Sepsis, SIRS = Systemic Inflammatory Response Syndrome.
*Strategies are ordered by decreasing PPV. No studies reporting thresholds for the CIP (Critical Illness Prediction) score or NEWS (National Early Warning Score) score for identification of sepsis were identified; as such these strategies are not included in this table.
†Authors included "nursing home transport" in calculation of their score, but this variable was not available in our data set.

Predictive ability

For the 14 screening strategies that generated a numeric score (Table 1), assessing predictive ability provides knowledge about the probability of sepsis diagnosis across each level of the numeric scores. We observed considerable variation in discrimination (C statistic range 0.61–0.79) and considerable change in probabilities for increasing scores among different strategies (Figure 2). Strategies using more measures and with a greater range of possible points generally identified patients with the highest probability of sepsis (e.g., CIP, National Early Warning Score [NEWS]; Appendix 3, Table A3-1). Strategies that included measures of the Glasgow Coma Scale and systolic blood pressure (e.g., CIP, NEWS, Quick Sepsis-Related Organ Failure Assessment [qSOFA]) generally identified patients with a higher probability of sepsis than strategies incorporating a similar number of predictors without either of these 2 measures (e.g., Prehospital Sepsis Assessment Tool, Systemic Inflammatory Response Syndrome [SIRS], Prehospital Sepsis Project [PSP]; Table 1). Two strategies with only 3 points on the scoring system, namely the qSOFA and the BAS (Blood Pressure *Andningsfrekvens* [“respiratory rate” in Swedish] Saturation), identified patients with a 20% to 30% increase in probability for each additional point (qSOFA range 0.07–0.87; BAS range 0.13–0.82); that is, qSOFA identified patients with a 7% (lowest score) to 87% (highest score) probability of sepsis diagnosis. However, another simple strategy (SIRS) had little change in probability of sepsis across the entire range of the score, identifying patients with only a 19% difference in probability between the minimum and maximum scores (Appendix 3, Table A3-1). Among patients with paramedic-suspected infection, the overall discrimination improved for all strategies (C statistic range 0.71–0.84), but the probabilities of sepsis diagnosis decreased for all strategies (Appendix 3, Table A3-2 and Figure A3-1). Calibration curves showed that the BAS and PSP strategies overestimated the probability of sepsis at high scores, whereas the PRESS score underestimated sepsis probability at high scores. Calibration for the Prehospital Early Sepsis Detection score was poor, whereas the remaining scores had consistent overlap of observed and predicted probabilities, which indicated good calibration (Appendix 3, Figures A3-2 to A3-15).

Interpretation

The accuracy of prehospital screening strategies for identification of sepsis by paramedics varied considerably, with no strategy having both high sensitivity and high specificity. However, in validating the predictive ability of strategies that used a numeric score, we found 3 strategies (CIP, NEWS and qSOFA) that had good discrimination and good calibration. With these strategies, low scores identify patients with low probability of sepsis, and high scores identify those with high probability of sepsis in the prehospital setting.

Sepsis is a syndrome rather than a disease.¹ Thus, a spectrum of severity of illness among patients is expected, and a gold standard test to accurately diagnose patients with sepsis is not available.⁶¹ To effectively navigate this uncertainty, clinicians need to know what information a screening strategy provides about an

individual patient’s risk of having sepsis. Accuracy of classification provides limited information about uncertainty, because it relies on a single result or threshold to identify the patient as having the disease or not.^{9,62–64} Our approach of evaluating the predictive ability of sepsis diagnosis across the entire range of scores helps to address this uncertainty by highlighting the change in risk at different scores. Screening strategies that can identify patients with low and high probability of sepsis may help clinicians determine which patients with suspected infection have low risk, and which patients are at high risk of having sepsis. Conversely, screening strategies with little change in probability from their lowest to their highest scores do not convey useful information to clinicians about an individual patient’s risk of sepsis.

Previous studies have compared the accuracy of classification of screening strategies for prehospital identification of sepsis in the same population^{30,53,55} or in systematic reviews.^{5,6,65} However, these comparisons were limited because they assessed only a few of the available screening strategies, they compared studies using different case definitions for sepsis, or they used diagnostic accuracy metrics. In contrast, in our study, we compared all published strategies within the same population using the same case definition for sepsis and using diagnostic prediction models. No screening strategy will be perfectly accurate for the diagnosis of sepsis, but our estimates of the diagnostic predictive value or the probability of diagnosis with each strategy provide clinicians with knowledge about how certain a diagnosis may be, given an individual patient’s presentation, thus allowing them to determine who might benefit from earlier intervention.

The CIP, NEWS and qSOFA scores had good predictive ability and the greatest range in the probability of sepsis diagnosis from their minimum to their maximum scores. Prehospital systems may consider integration of these screening strategies into paramedic treatment protocols, using a higher probability of sepsis (i.e., a higher score) to inform a stepwise approach to more aggressive intervention by paramedics during transport of these patients. For example, paramedics might consider notifying the emergency department in advance if screening reveals that a patient has moderate probability of sepsis (e.g., qSOFA score of 1 or 2), whereas they might initiate prehospital interventions and emergency transport and provide advance notification for a patient with a higher probability of sepsis (e.g., qSOFA score of 3). Screening strategies requiring only 3 measures, such as qSOFA, are simple and more likely to be used by paramedics. Future studies could test the clinical benefit and feasibility of adopting these approaches to guide paramedic interventions prospectively, or they could investigate the predictive ability of these same screening strategies for other important patient outcomes, as has been tested for predicting mortality with the qSOFA score.^{44,66}

Limitations

This study had some limitations. The update to our previous search, which we used to identify existing paramedic screening strategies, was not systematic and was limited by the use of only 1 reviewer and restriction of the search to English articles published in peer-reviewed journals.

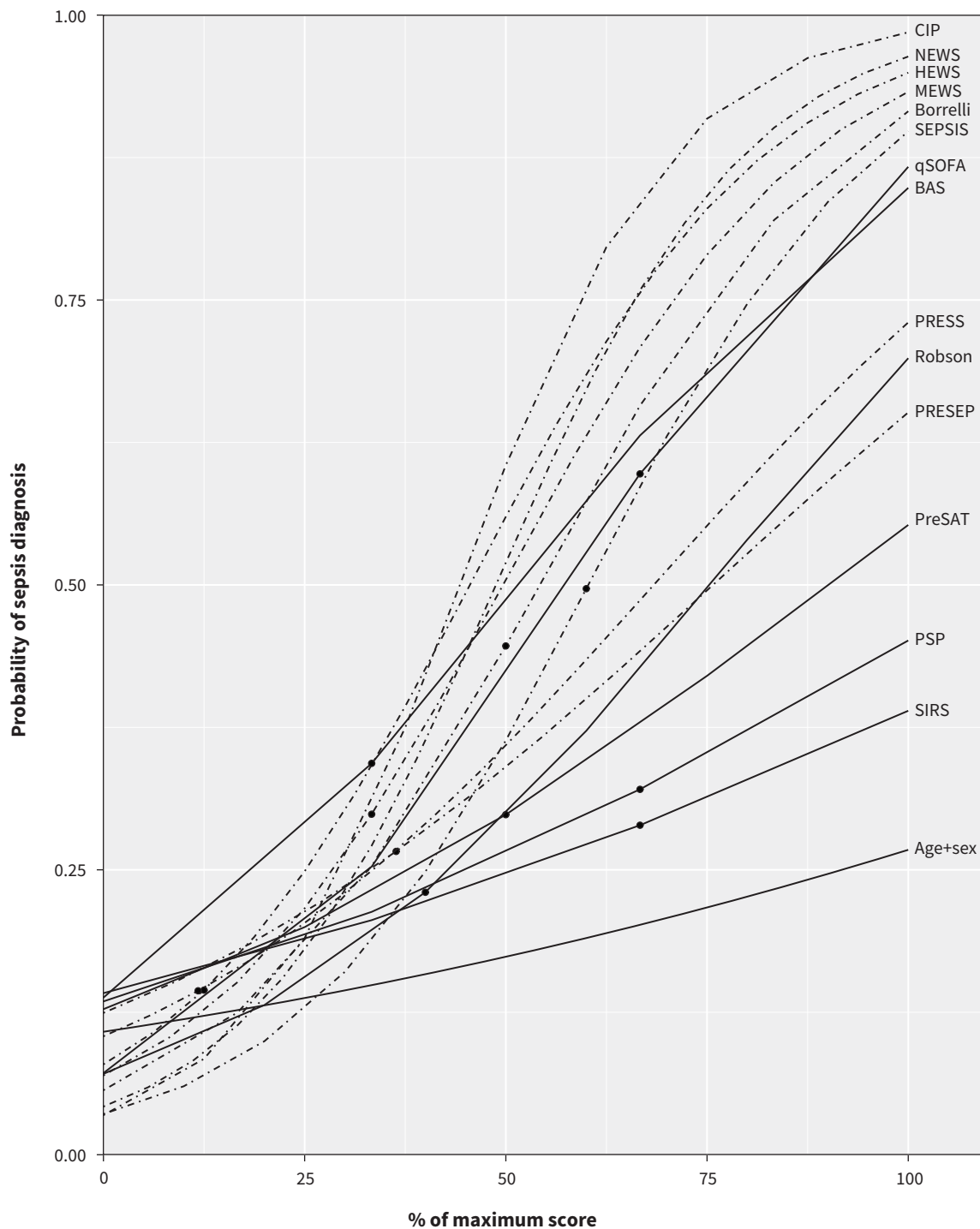


Figure 2: Comparison of the predicted probability of a diagnosis of sepsis according to 14 numeric scores for patients with infection diagnosed in the emergency department. Note: The x axis was scaled from minimum to maximum for each score to allow for comparison of scores with a different number of points. Broken lines indicate scores with more than 5 possible points. The dot on each line represents the recommended threshold score for a diagnosis of sepsis (if available). BAS = Blood Pressure *Andningsfrekvens* (“respiratory rate” in Swedish) Saturation, Borrelli = Borrelli et al.,²⁹ CIP = Critical Illness Prediction score, HEWS = Hamilton Early Warning Score, MEWS = Modified Early Warning Score, NEWS = National Early Warning Score, PreSAT = Prehospital Sepsis Assessment Tool, PRESEP = Prehospital Early Sepsis Detection, PRESS = Prehospital Severe Sepsis, PSP = Prehospital Sepsis Project, qSOFA = Quick Sepsis-Related Organ Failure Assessment, Robson = Robson score (originally validated by Wallgren et al.⁵³), SEPSIS = Screening to Enhance Prehospital Identification of Sepsis, SIRS = Systemic Inflammatory Response Syndrome.

In the absence of a gold standard for sepsis diagnosis, we adopted a validated approach that aligns with the most recent consensus definition.¹³ Previous studies have found consistent

undercoding of sepsis when administrative databases are used;¹⁵ in our study, such undercoding of sepsis would result in missed cases and more conservative estimates of predictive ability overall.

Our use of patient measures to identify patients with organ dysfunction due to sepsis might have introduced some incorporation bias into the assessment of screening strategies that used these same measures to diagnose sepsis. However, our use of the strategy results rather than the individual measures would have reduced this bias; there was no direct incorporation of the measures into the model as predictors. We considered that this approach was superior to the alternative of excluding patients with clinically important organ dysfunction from the sepsis outcome classification due to the consistent undercoding discussed above.

A delay occurs between paramedic assessments and determination of sepsis in the emergency department. Sepsis is a syndrome that may progress during this period, and screening strategies that paramedics apply at initial assessment may not correctly identify patients whose condition will continue to worsen and in whom sepsis is subsequently diagnosed by emergency physicians. This limitation could decrease the apparent accuracy of the prehospital screening strategies for use by paramedics.

We did not have access to information about the location from which patients were transferred; this information might have helped us to identify populations at higher risk (e.g., nursing home residents).

In our sensitivity analysis, the rate of identification of infections by paramedics was low. The application of the screening strategies for detecting sepsis depends on recognition of infections; therefore, paramedics should be trained to improve infection recognition in the prehospital setting.

Conclusion

Validation of the predictive ability of available screening strategies for possible sepsis in the prehospital setting showed that certain scores identified patients with both low and high probability for sepsis diagnosis, despite poor sensitivity or specificity at recommended thresholds. The CIP, NEWS and qSOFA scores each identified patients with low probability of sepsis at low scores and high probability of sepsis at high scores. The qSOFA score is the simplest of these to calculate and should be considered for use by paramedics.

References

- 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.
- Executive Board, 140th session. *Improving the prevention, diagnosis and clinical management of sepsis. Report by the Secretariat*. Geneva: World Health Organization; 2017.
- Rhodes A, Evans LE, Alhazzani W, et al. Surviving Sepsis Campaign: international guidelines for management of sepsis and septic shock: 2016. *Crit Care Med* 2017;45:486-552.
- Wang HE, Weaver MD, Shapiro NI, et al. Opportunities for emergency medical services care of sepsis. *Resuscitation* 2010;81:193-7.
- Lane D, Ichelson RI, Drennan IR, et al. Prehospital management and identification of sepsis by emergency medical services: a systematic review. *Emerg Med J* 2016;33:408-13.
- Smyth MA, Brace-McDonnell SJ, Perkins GD. Identification of adults with sepsis in the prehospital environment: a systematic review. *BMJ Open* 2016;6:e011218.
- Smyth MA, Brace-McDonnell SJ, Perkins GD. Impact of prehospital care on outcomes in sepsis: a systematic review. *West J Emerg Med* 2016;17:427-37.
- Herlitz J, Bång A, Wireklint-Sundström B, et al. Suspicion and treatment of severe sepsis. An overview of the prehospital chain of care. *Scand J Trauma Resusc Emerg Med* 2012;20:42.
- Hlatky MA, Mark DB, Harrell FE Jr, et al. Rethinking sensitivity and specificity. *Am J Cardiol* 1987;59:1195-8.
- Harrell FE Jr, Shih YC. Using full probability models to compute probabilities of actual interest to decision makers. *Int J Technol Assess Health Care* 2001; 17:17-26.
- Elm von E, Altman DG, Egger M, et al.; STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008; 61:344-9.
- Benchimol EI, Smeeth L, Guttman A, et al. The reporting of studies conducted using observational routinely-collected health data (RECORD) statement. *PLoS Med* 2015;12:e1001885.
- Lane DJ, Blanchard IE, Cheskes S, et al. Strategy to identify paramedic transported sepsis cases in an emergency department administrative database. *Prehosp Emerg Care* 2020;24:23-31.
- Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet* 1974;2:81-4.
- Jolley RJ, Quan H, Jetté N, et al. Validation and optimisation of an ICD-10-coded case definition for sepsis using administrative health data. *BMJ Open* 2015;5:e009487-11.
- Hendriksen JMT, Geersing GJ, Moons KGM, et al. Diagnostic and prognostic prediction models. *J Thromb Haemost* 2013;11:129-41.
- Harrell F. *Regression modeling strategies*. Cham (Switzerland): Springer; 2015:1.
- Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology* 2010;21:128-38.
- Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models. *JAMA* 2017;318:1377-84.
- Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005;43:1130-9.
- Yoshida K, Chipman JJ, Bohn J, et al. *Tableone: create "table 1" to describe baseline characteristics*. R package version 0.10.0. Comprehensive R Archive Network (CRAN); 2015. Available: <https://CRAN.R-project.org/package=tableone> (accessed 2020 Feb. 16).
- Wasey JO. *icd: comorbidity calculations and tools for ICD-9 and ICD-10 codes*. Version 4.0.6. Comprehensive R Archive Network (CRAN); 2017. Available: <https://CRAN.R-project.org/package=icd> (accessed 2020 Feb. 16).
- Alam N, Doerga KB, Hussain T, et al. Epidemiology, recognition and documentation of sepsis in the pre-hospital setting and associated clinical outcomes: a prospective multicenter study. *Acute Med* 2016;15:168-75.
- Alam N, Oskam E, Stassen PM, et al. Prehospital antibiotics in the ambulance for sepsis: a multicentre, open label, randomised trial. *Lancet Respir Med* 2018;6:40-50.
- Asayama K, Aikawa N. Evaluation of systemic inflammatory response syndrome criteria as a predictor of mortality in emergency patients transported by ambulance. *Keio J Med* 1998;47:19-27.
- Baez AA, Cochon L; Acute Care Diagnostics Collaboration. Assessment of a Bayesian clinical decision model integrating the Prehospital Sepsis Score and point-of-care lactate. *Am J Emerg Med* 2016;34:193-6.
- Barbara P, Graziano CG, Caputo WC, et al. The quick sequential organ failure assessment (qSOFA) identifies septic patients in the out-of-hospital setting. *Am J Emerg Med* 2018;36:1022-6.
- Bayer O, Schwarzkopf D, Stumme C, et al. An early warning scoring system to identify septic patients in the prehospital setting: the PRESEP score. *Acad Emerg Med* 2015;22:868-71.
- Borrelli G, Koch EK, Sterk ES, et al. Early recognition of sepsis through emergency medical services pre-hospital screening. *Am J Emerg Med* 2018;22:1-5.
- Dorsett M, Kroll M, Smith CS, et al. qSOFA has poor sensitivity for prehospital identification of severe sepsis and septic shock. *Prehosp Emerg Care* 2017; 21:489-97.
- Green RS, Travers AH, Cain E, et al. Paramedic recognition of sepsis in the prehospital setting: a prospective observational study. *Emerg Med Int* 2016;2016: 6717261.
- Guerra WF, Meyers MS, Cloutre AE, et al. Early detection and treatment of patients with severe sepsis by prehospital personnel. *J Emerg Med* 2013; 44:1116-25.
- Harada M, Takahashi T, Haga Y, et al. Comparative study on quick sequential organ failure assessment, systemic inflammatory response syndrome and the shock index in prehospital emergency patients: single-site retrospective study. *Acute Med Surg* 2019;6:131-7.
- Hunter CL, Silvestri S, Ralls G, et al. A prehospital screening tool utilizing end-tidal carbon dioxide predicts sepsis and severe sepsis. *Am J Emerg Med* 2016; 34:813-9.
- Hunter C, Silvestri S, Ralls G, et al. Comparing quick sequential organ failure assessment scores to end-tidal carbon dioxide as mortality predictors in prehospital patients with suspected sepsis. *West J Emerg Med* 2018;19:446-51.

36. Jouffroy R, Saade A, Carpentier A, et al. Triage of septic patients using qSOFA criteria at the SAMU regulation: a retrospective analysis. *Prehosp Emerg Care* 2018;12:84-90.
37. Jouffroy R, Saade A, Ellouze S, et al. Prehospital triage of septic patients at the SAMU regulation: comparison of qSOFA, MRST, MEWS and PRESEP scores. *Am J Emerg Med* 2018;36:820-4.
38. Koyama S, Yamaguchi Y, Gibo K, et al. Use of prehospital qSOFA in predicting in-hospital mortality in patients with suspected infection: a retrospective cohort study. *PLoS One* 2019;14:e0216560.
39. Lane DJ, Lin S, Scales DC. Classification versus prediction of mortality risk using the SIRS and qSOFA scores in patients with infection transported by paramedics. *Prehosp Emerg Care* 2019 June 19:1-8. doi: 10.1080/10903127.2019.1624901. [Epub ahead of print].
40. Miyamoto KM, Shibata NS, Nakashima TN, et al. Prehospital quick sequential organ failure assessment as a tool to predict in-hospital mortality. *Am J Emerg Med* 2018;36:1832-36.
41. Polito CC, Isakov A, Yancey AH, et al. Prehospital recognition of severe sepsis: development and validation of a novel EMS screening tool. *Am J Emerg Med* 2015;33:1119-25.
42. Roest AA, Stoffers J, Pijpers E, et al. Ambulance patients with nondocumented sepsis have a high mortality risk: a retrospective study. *Eur J Emerg Med* 2017;24:36-43.
43. Seymour CW, Kahn JM, Cooke CR, et al. Prediction of critical illness during out-of-hospital emergency care. *JAMA* 2010;304:747-54.
44. Seymour CW, Liu VX, Iwashyna TJ, et al. Assessment of clinical criteria for sepsis: for the third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016;315:762-74.
45. Silcock DJ, Corfield AR, Gowens PA, et al. Validation of the National Early Warning Score in the prehospital setting. *Resuscitation* 2015;89:31-5.
46. Shu E, Tallman CI, Frye W, et al. Pre-hospital qSOFA as a predictor of sepsis and mortality. *Am J Emerg Med* 2019;37:1273-8.
47. Skitch S, Tam B, Xu M, et al. Examining the utility of the Hamilton early warning scores (HEWS) at triage: retrospective pilot study in a Canadian emergency department. *CJEM* 2018;20:266-74.
48. Smyth MA, Gallacher D, Kimani PK, et al. Derivation and internal validation of the screening to enhance prehospital identification of sepsis (SEPSIS) score in adults on arrival at the emergency department. *Scand J Trauma Resusc Emerg Med* 2019;27:67.
49. Suffoletto B, Frisch A, Prabhu A, et al. Prediction of serious infection during prehospital emergency care. *Prehosp Emerg Care* 2011;15:325-30.
50. Tugul S, Carron P-N, Yersin B, et al. Low sensitivity of qSOFA, SIRS criteria and sepsis definition to identify infected patients at risk of complication in the pre-hospital setting and at the emergency department triage. *Scand J Trauma Resusc Emerg Med* 2017;25:108.
51. Vaittinada Ayar P, Delay M, Avondo A, et al. Prognostic value of prehospital quick sequential organ failure assessment score among patients with suspected infection. *Eur J Emerg Med* 2019;26:329-33.
52. Walchok JG, Pirralo RG, Furmanek D, et al. Paramedic-initiated CMS sepsis core measure bundle prior to hospital arrival: a stepwise approach. *Prehosp Emerg Care* 2017;21:291-300.
53. Wallgren UM, Castrén M, Svensson AEV, et al. Identification of adult septic patients in the prehospital setting. *Eur J Emerg Med* 2014;21:260-5.
54. van der Wekken LCW, Alam N, Holleman F, et al. Epidemiology of sepsis and its recognition by emergency medical services personnel in the Netherlands. *Prehosp Emerg Care* 2016;20:90-6.
55. Studnek JR. Assessing and comparing the validity of three screening tools to identify sepsis in the prehospital setting [abstract 556]. *Acad Emerg Med* 2017;24 Suppl 1:S197-8.
56. Scales DC. Paramedic initiated treatment of sepsis targeting out-of-hospital patients (PITSTOP). ClinicalTrials.gov: NCT03068741; 2017. Available: <https://clinicaltrials.gov/ct2/show/NCT03068741> (accessed 2020 Feb. 16).
57. Field paramedic application of sepsis triage (FAST). ClinicalTrials.gov: NCT03870789; 2019. Available: <https://clinicaltrials.gov/ct2/show/NCT03870789> (accessed 2020 Feb. 16).
58. Validation of early warning score and lactate in prehospital screening (VELPS). ClinicalTrials.gov: NCT02189096; 2014. Available: <https://clinicaltrials.gov/ct2/show/NCT02189096> (accessed 2019 Apr. 17).
59. Bullard MJ, Chan T, Brayman C, et al. Revisions to the Canadian Emergency Department Triage and Acuity Scale (CTAS) guidelines. *CJEM*. 2014;16:485-9.
60. Lane DJ, Wunsch H, Saskin R, et al. Assessing severity of illness in patients transported to hospital by paramedics: external validation of 3 prognostic scores. *Prehosp Emerg Care* 2019;July 15:1-9. doi: 10.1080/10903127.2019.1632998. [Epub ahead of print].
61. Prescott HC, Iwashyna JT. Improving sepsis treatment by embracing diagnostic uncertainty. *Ann Am Thorac Soc* 2019;16:426-9.
62. Feinstein AR. Misguided efforts and future challenges for research on "diagnostic tests." *J Epidemiol Community Health* 2002;56:330-2.
63. Vickers AJ, Basch E, Kattan MW. Against diagnosis. *Ann Intern Med* 2008;149:200-3.
64. Moons KGM, Harrell FE. Sensitivity and specificity should be de-emphasized in diagnostic accuracy studies. *Acad Radiol* 2003;10:670-2.
65. Alam N, Hobbelenk EL, van Tienhoven AJ, et al. The impact of the use of the Early Warning Score (EWS) on patient outcomes: a systematic review. *Resuscitation* 2014;85:587-94.
66. 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; 317:290-11.

Competing interests: Sheldon Cheskes has received investigator-initiated grant funding from Zoll Medical for several research programs (AED on the Fly, Community Responder Program for Out-of-Hospital Cardiac Arrest and Measuring Ventilation During Out-of-Hospital Cardiac Arrest). He also sits on the advisory board of Drone Delivery Canada. Damon Scales holds operating grants from the Canadian Institutes of Health Research. No other competing interests were declared.

This article has been peer reviewed.

Affiliations: Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health (Lane, Wunsch, Saskin, Lin, Scales), Interdepartmental Division of Critical Care (Wunsch, Scales), Division of Emergency Medicine, Department of Family and Community Medicine (Cheskes), and Division of

Emergency Medicine, Department of Medicine (Lin, Morrison), University of Toronto; Rescu, Li Ka Shing Knowledge Institute (Lane, Cheskes, Lin, Morrison), St. Michael's Hospital; Department of Critical Care Medicine (Wunsch) and Sunnybrook Centre for Prehospital Medicine (Cheskes), Sunnybrook Health Sciences Centre, Toronto, Ont.

Contributors: Daniel Lane and Damon Scales conceived of the study. All of the authors contributed to design of the study and interpretation of the data. Daniel Lane performed the primary analysis, supervised by Refik Saskin and Damon Scales. Daniel Lane and Damon Scales wrote the primary draft of the manuscript. All of the authors revised the manuscript for important intellectual content, approved the final version for publication and agreed to be accountable for the work.

Funding: No specific funding was received for this work. Laurie Morrison is the Robert and Dorothy Pitts Chair in Acute Care and Emergency Medicine at St. Michael's Hospital and the University of Toronto.

Data sharing: The data for this study were obtained under a research agreement with Alberta Health Services, Emergency Medical Services. Access to these data may be requested from Alberta Health Services through the process outlined on this site: www.albertahealthservices.ca/ems/Page13364.aspx

Accepted: Jan. 15, 2020

Correspondence to: Daniel Lane, djlane@ucalgary.ca