OPEN ACCESS

Validation of automated sepsis surveillance based on the Sepsis-3 clinical criteria against physician record review in a general hospital population: observational study using electronic health records data

John Karlsson Valik ,[1,2] Logan Ward ,[3,4] Hideyuki Tanushi,[2] Kajsa Müllersdorf,[1,2] Anders Ternhag,[1,2] Ewa Aufwerber,[2] Anna Färnert,[1,2] Anders F Johansson ,[5] Mads Lause Mogensen,[3] Brian Pickering,[6] Hercules Dalianis,[7] Aron Henriksson,[7] Vitaly Herasevich,[6] Pontus Nauclér[1,2]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmjqs-2019-010123).

For numbered affiliations see end of article.

Correspondence to
Dr John Karlsson Valik, Division of Infectious Diseases, Department of Medicine, Solna (MedS), Karolinska Institutet, Stockholm, 171 77 Solna, Sweden;
john.karlsson.valik@ki.se

Received 23 July 2019
Revised 19 January 2020
Accepted 21 January 2020

Check for updates

To cite: Valik JK, Ward L, Tanushi H, *et al*. *BMJ Qual Saf* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bmjqs-2019-010123

ABSTRACT

Background Surveillance of sepsis incidence is important for directing resources and evaluating quality-of-care interventions. The aim was to develop and validate a fully-automated Sepsis-3 based surveillance system in non-intensive care wards using electronic health record (EHR) data, and demonstrate utility by determining the burden of hospital-onset sepsis and variations between wards.

Methods A rule-based algorithm was developed using EHR data from a cohort of all adult patients admitted at an academic centre between July 2012 and December 2013. Time in intensive care units was censored. To validate algorithm performance, a stratified random sample of 1000 hospital admissions (674 with and 326 without suspected infection) was classified according to the Sepsis-3 clinical criteria (suspected infection defined as having any culture taken and at least two doses of antimicrobials administered, and an increase in Sequential Organ Failure Assessment (SOFA) score by ≥2 points) and the likelihood of infection by physician medical record review.

Results In total 82 653 hospital admissions were included. The Sepsis-3 clinical criteria determined by physician review were met in 343 of 1000 episodes. Among them, 313 (91%) had possible, probable or definite infection. Based on this reference, the algorithm achieved sensitivity 0.887 (95% CI: 0.799 to 0.964), specificity 0.985 (95% CI: 0.978 to 0.991), positive predictive value 0.881 (95% CI: 0.833 to 0.926) and negative predictive value 0.986 (95% CI: 0.973 to 0.996). When applied to the total cohort taking into account the sampling proportions of those with and without suspected infection, the algorithm identified 8599 (10.4%) sepsis episodes. The burden of hospital-onset sepsis (>48 hour after admission) and related in-hospital mortality varied between wards.

Conclusions A fully-automated Sepsis-3 based surveillance algorithm using EHR data performed well compared with physician medical record review in non-intensive care wards, and exposed variations in hospital-onset sepsis incidence between wards.

INTRODUCTION

Sepsis, a severe organ dysfunction induced by infection, is a leading cause of morbidity and death worldwide.[1–3] The true burden of sepsis has been difficult to assess, mainly due to the absence of a generalisable gold standard. About one third of sepsis episodes are considered healthcare-associated and the problem of sepsis needs to be addressed as a patient safety concern.[4]

Surveillance with feedback to healthcare personnel and policy makers is the backbone of most quality improvement programmes for healthcare-associated infections.[5] To be useful, such surveillance systems require standardised case-definitions free from subjective interpretations and appropriate denominator data.[6] Surveillance systems based on clinical data are preferred to administrative data, as these are more objective, reproducible and stable over time.[7 8] Fully-automated surveillance systems with data from electronic health records (EHR) could replace surveillance relying on manual chart review and generate continuous data from large populations, but needs thorough validation before

implementation.[5] As mandatory reporting of sepsis is becoming increasingly common, defining a surveillance method that produces objective high quality data is important.[9] Reliable sepsis surveillance data can benefit large patient groups by allowing clinical resources to be directed to where they are most needed. Continuous incidence monitoring can also be used to evaluate quality of care interventions down to the ward level, and for benchmarking sepsis prediction models and artificial intelligence tools integrated with the EHR.[8 10 11]

In 2016 the Third International Consensus Definition for Sepsis and Septic Shock (Sepsis-3) was introduced.[12 13] An advantage of the new definition (here denoted Sepsis-3 clinical criteria) lies in its objective case-definition, which requires that the patient has a suspected infection in combination with a newly developed organ dysfunction.[13] Yet, the performance of an automated surveillance system using these criteria has not been evaluated. In 2018, the Center of Disease Control and Prevention (CDC) presented an additional sepsis definition called Adult Sepsis Event (ASE) aimed specifically for surveillance purposes.[14] The ASE differs from the Sepsis-3 clinical criteria with regards to both classifications of suspected infection and organ dysfunction. The focus of ASE is on patients with more severe disease, with more than half of cases admitted to the intensive care unit (ICU), and it has been shown to underestimate the burden of sepsis cases defined by Sepsis-3 criteria.[15]

The primary aim of this study was to develop and validate a fully-automated EHR-based surveillance algorithm against physician medical record review in non-intensive care wards using the Sepsis-3 clinical criteria. A secondary aim was to demonstrate the algorithm's utility by determining the burden of hospital-onset sepsis in a general hospital population.

METHODS

Design, data source and study population

This was an observational study performed at an academic centre with 1350 beds divided between two hospitals and serving a population of 2.3 million inhabitants. Data was obtained from routinely prospectively entered information in the EHR system, stored in a research databank called Health Bank—Swedish Health Record Research Infrastructure.[16] The database structure is a duplicate of the operating EHR system, where each subject can be followed over time, and consists of all medical records from more than 2 million anonymised patients that received care at the hospital between 2006 and 2013. Due to improved recording in the EHR during the later years, analyses were restricted to July 2012 until December 2013, except for information about International Classification of Diseases (ICD) codes that were used to estimate the presence of co-morbidities in patients, which were retrieved up to 5 years before inclusion. Data collection included demographics, hospital administrative data, vital parameters, laboratory findings, microbiological data, medications and in-hospital mortality.

Patients >18 years admitted to the hospital for ≥24 hours were included, and followed until first sepsis episode, discharge or death. Patients were excluded if admitted to an obstetric ward and censored during ICU-care, due to lack of data on vital parameters and medication for these wards.

Sepsis-3 surveillance case definition

The rule-based algorithm was based on the operational Sepsis-3 clinical criteria: a suspected infection in combination with an increase in Sequential Organ Failure Assessment (SOFA) score by ≥2 points compared with the baseline.[13]

Suspected infection was defined as having any culture taken and at least two doses of antimicrobials administered. If the patient was admitted to the ICU prior to 24 hours, or died prior to 48 hours from the first dose of antimicrobials, they were deemed to have a suspected infection despite only being given one dose. Cultures had to be performed within 24 hours after the start of antimicrobial treatment. Antimicrobial treatment had to be started within 72 hours after culture. Onset of infection was determined based on which of these events occurred first.[13] Sensitivity analyses were done using different definitions of suspected infections: only blood cultures and two doses of antimicrobials, any culture and four calendar days of antimicrobials or only blood cultures and four calendar days of antimicrobials, of which the last being equivalent to the ASE definition (online supplementary methods 1).[14]

Organ dysfunction was measured as the maximum SOFA score 48 hours before to 24 hours after onset of infection and compared with a baseline SOFA score measured separately (online supplementary methods 1 and online supplementary figure 1). Similar to the study that developed the Sepsis-3 clinical criteria, missing values during the 72 hours window were considered to be normal.[13] Since we studied a non-ICU population, some modification to the SOFA score was done. The most important changes were, (i) if PaO2 was not available it was calculated from peripheral capillary oxygen saturation (SpO2), (ii) if Glasgow Coma Scale (GCS) was not available, structured data on 'alert' (interpreted as GCS score 15 points) or 'not alert' (interpreted as GCS score 14 points) was used, and (iii) urine output was not used due to data being unavailable. For each component of the SOFA score, the baseline was defined as the latest value measured before the 72 hours time window, and was assumed to be zero in patients not known to have a pre-existing organ dysfunction. Pre-existing organ dysfunction was based on measured parameters (coagulation, liver, renal and respiration) within the previous three months or a specific ICD-code (chronic dialysis or home oxygen

therapy) within the last year. For SOFA cardiovascular and central nervous system (CNS) scores, only values measured during the current hospital episode was used. Onset of sepsis was when the patient fulfilled the organ dysfunction criteria.

To capture the burden of healthcare-associated sepsis, hospital-onset (HO) sepsis was defined as onset of suspected infection and organ dysfunction 48 hours after admission, or re-admission with sepsis within 48 hours of discharge. All other episodes were defined as community-onset (CO) sepsis. A patient could have several suspected infections, but only the first episode of sepsis was considered for each hospital episode. Algorithm classification based on clinical data was compared with classification using the following ICD-10 codes indicating sepsis: A02.1, A22.7, A26.7, A32.7, A39.2, A39.4, A40.x, A41.x, A42.7, A48.3, B37.7, M72.6, R57.2, R65.1 and R65.9.

Validation using medical record review

To evaluate the performance of the surveillance algorithm, two validation sets including a total of 1000 hospital admissions were selected from the entire hospital cohort for medical record review. In the first validation set, 674 hospital admissions were randomly sampled from patients with suspected infection (540 CO and 134 HO episodes). Medical records including demographics, hospital administrative data, free text notes, medications, microbiological cultures, laboratory and radiological findings were reviewed by two trained infectious diseases physicians to classify whether the patient fulfilled the Sepsis-3 clinical criteria. The first 10 patients were reviewed together as a run-in period, and further reviewing was performed independently with an overlap of 100 patients. There was substantial agreement between reviewers, with Cohen's kappa 0.75 for sepsis classification. Complicated cases were classified using a consensus decision. The reviewers were blinded from the results of the developed surveillance algorithm. In the second validation set, 326 episodes were randomly sampled from hospital admissions without a suspected infection. Full medical records were assessed by one of the reviewers and classified according to the Sepsis-3 clinical criteria.

The medical records of subjects that fulfilled the Sepsis-3 clinical criteria by physician review were assessed in further detail for likelihood and source of infection. The categorisation followed previously validated criteria based on CDC and The International Sepsis Forum definitions.[17–20] Accordingly, episodes were divided by source and classified on a four-graded scale as *no infection*, *possible infection*, *probable infection* and *definite infection*. For details regarding the exact definitions used we refer to a previously published study by Klein Klouwenberg *et al*.[18] One minor modification to the criteria was done in this study. We added unknown source, defined as patients (i) with symptoms of an infection, (ii) the symptoms indicated an infection according to the attending physician, and (iii) the patient received a full course of anti-infective treatment, but (iv) no source could be determined. Unknown source could only be classified as *possible infection*. In the assessment of the sensitivity and specificity of the surveillance algorithm, patients had to fulfil both the Sepsis-3 clinical criteria and the *possible*, *probable* or *definite infection* criteria to be classified as true sepsis.

Statistical analyses

To assess algorithm performance in the intended target population of all patients admitted to the hospital, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated by generalising the proportions from validation to the entire cohort, as previously described by Rhee *et al*.[15] CI for sensitivity, specificity, PPV and NPV were calculated as the 2.5th and 97.5th percentiles of point estimates obtained from 10 000 bootstrap samples for each of the two validation sets (n=674 and n=326) using the 'boot' package of R. To account for uncertainty, the bootstrapping was performed before extrapolating the proportions from validation to the entire hospital cohort. The extrapolation accounted for the nested selection of suspected CO and HO infections, as well as for the proportion of sepsis cases from the population of patients who did not have a suspected infection episode. In sensitivity analyses of different definitions of suspected infection, the proportion of patients that were falsely categorised as true negatives due to not fulfilling the tested suspected infection definition was also accounted for. When assessing algorithm performance for CO infections, HO infections were omitted and vice versa.

CIs for incidence densities of CO sepsis per 100 admissions and HO sepsis per 1000 patient days at risk were calculated using the R package 'compeir', assuming a log-normal distribution. Cumulative incidence function (CIF) for the probability of HO sepsis was calculated using the Aalen-Johansen estimator and taking into account competing risks: ICU admission, discharge or death.[21] Pairwise comparison of CIF between wards was calculated using the R package 'cmprsk', which is based on the Fine and Gray (1999) formula.[22] Cumulative incidence regression analysis for in-hospital death stratified by likelihood of infection (none-possible vs probable-definite) was determined with discharge as a competing event using the 'stcrreg' command in STATA. Confounders associated with both the likelihood of infection and in-hospital death were considered using a directed acyclic graph based on a priori clinical expertise and available literature. Accordingly, adjustments were made for age, Charlson comorbidity index and CO/HO onset, but not for severity of disease since this covariate could be on the causal pathway between likelihood of infection and in-hospital death.

Results are presented as median (med) and IQR or numbers (No.) and percentages as appropriate. Differences between categorical variables and between continuous variables were assessed using the Fisher exact and Mann-Whitney U tests, respectively. Missing data was not considered missing at random, but rather owing to clinical decisions. Based on the assumption that data would still be not missing at random even after controlling for other observed variables, we did not perform multiple imputation for individual SOFA score variables. Data handling and automated sepsis annotation was performed in Python V.3.6. Statistical analyses were done in R V.3.4.3 and STATA V.14.2.

RESULTS

In total 144 179 hospital admissions of 99 864 patients were recorded during the study period, of which 95 858 admissions fulfilled the inclusion criteria (patients >18 years admitted to the hospital for ≥24 hours). Twelve thousand nine hundred-fifty admissions to the obstetrical wards, 214 admissions to paediatric wards and 41 direct transfers between hospital ICUs were excluded. Finally, a total of 82 653 hospital admissions of 54 884 patients were included in the analysis. There was a suspected infection in 19 479 (23.6%) of the admissions and no suspected infection in 63 174 (76.4%) of the admissions. The median patient age was 64 years, 50.9% were women and the median length of stay was 3.8 days.

Validation of the surveillance case definition

In total, 340 of 674 patients with suspected infection (50.4%) and three of 326 without suspected infection (0.9%) fulfilled Sepsis-3 criteria according to physician medical record review (table 1). Among subjects that fulfilled Sepsis-3 criteria, 109/343 (31.8%) fulfilled the criteria for *possible infection*, 87/343 (25.4%) for *probable infection*, and 117/343 (34.1%) for *definite infection*. In total, 30/343 (8.7%) were classified as *no infection*. Hence, 311 of 674 patients (46.1%) with suspected infection and two of 326 patients (0.6%) without suspected infection were finally considered as true sepsis by reviewers. In subjects with suspected infection, the algorithm classified 288 true positive, 39 false positive, 324 true negative and 23 false negative sepsis cases. In subjects without suspected infection, the algorithm classified zero sepsis cases resulting in 324 true negative and two false negative cases. Based on the medical record reviewed reference, the algorithm achieved sensitivity 0.887 (95% CI: 0.799 to 0.964), specificity 0.985 (95% CI: 0.978 to 0.991), PPV 0.881 (95% CI: 0.833 to 0.926) and NPV 0.986 (95% CI: 0.973 to 0.996) when extrapolating proportions to the entire hospital cohort (table 2). When assessed only in subjects with suspected infection, the algorithm achieved sensitivity 0.926 (95% CI: 0.896 to 0.955), specificity 0.893 (95% CI: 0.859 to 0.923), PPV 0.881 (95% CI: 0.833 to 0.926) and NPV 0.934 (95% CI: 0.895 to 0.969) (table 2). The most common reasons for misclassification resulting in reduced sensitivity was respiratory or CNS dysfunction only being mentioned in free text, followed by overestimation of pre-existing organ dysfunction or development of infection related organ dysfunction outside of the 72 hours time window (online supplementary table 1). Reasons for imperfect specificity was due to episodes judged by reviewers as *no infection*, misclassification of baseline SOFA score or obvious measurement errors of vital parameters in the EHR.

For CO-sepsis the algorithm achieved sensitivity 0.910 (95% CI: 0.825 to 0.984), specificity 0.987 (95% CI: 0.982 to 0.991), PPV 0.881 (95% CI: 0.844 to 0.917) and NPV 0.990 (95% CI: 0.980 to 0.998). For HO-sepsis the algorithm achieved sensitivity 0.794 (95% CI: 0.683 to 0.889), specificity 0.997 (95% CI: 0.995 to 0.999), PPV 0.877 (95% CI: 0.782 to 0.966) and NPV 0.994 (95% CI: 0.991 to 0.997). Restricting analyses to hospital admissions without ICU admission (n=78 318) resulted in slightly decreased sensitivity 0.879 (0.793–0.952) but increased specificity 0.988 (0.983–0.992) and PPV 0.895 (0.860–0.931). For hospital admissions with ICU admission (n=4335), sensitivity was higher 0.952 (0.881–1.000) at the expense of decreased specificity 0.938 (0.907–0.969) and PPV 0.800 (0.712–0.894) (online supplementary table 2).

Classification of infection as *probable* or *definite* was not associated with a significantly different in-hospital mortality compared with subjects with *no* or *possible infection* (online supplementary figure 2). The most common source of infection in true sepsis patients was respiratory (n=119/313, 38.0%), followed by urogenital (n=54/313, 17.3%), unknown source (n=42/313, 13.4%), bloodstream (35/313, 11.2%), skin, bone and joint (30/313, 9.6%), abdominal (n=26/313, 8.3%) and other infectious sources (7/313, 2.2%) (table 1 and online supplementary figure 3). Among patients classified as having an unknown source of infection, 15/42 (35.7%) had neutropenia.

The burden of sepsis

The surveillance algorithm identified 8599 sepsis episodes (10.4% of all hospital admissions), of which 7493 (87.1%) were CO sepsis and 1106 (12.9%) were HO sepsis (table 3). The most common SOFA score triggers for sepsis were respiratory and renal dysfunction (online supplementary figure 4). Availability of data to calculate SOFA score during suspected infection ranged between 92.0%–95.0% (coagulation, renal, respiratory and cardiovascular) and 38.3%–55.0% (liver and CNS) for community-onset episodes, compared with 73.2%–86.2% (coagulation, respiratory, renal and cardiovascular) and 3.0%–30.2% (CNS and liver) for hospital-onset episodes (online supplementary table 3). Assumptions of normal baseline SOFA score were almost exclusively done in community-onset episodes

Table 1 Characteristics of patients fulfilling Sepsis-3 clinical criteria* according to physician review of medical records

		Likelihood of infection			
	All	None	Possible	Probable	Definite
Patients, No. (% of all)	343 (100.0)	30 (8.7)	109 (31.8)	87 (25.4)	117 (34.1)
Female sex, No. (%)	158 (46.1)	13 (43.3)	43 (39.4)	34 (39.1)	68 (58.1)
Age, med (IQR)	71 (60–81)	68 (52–81)	70 (60–79)	72 (60–83)	72 (62–81)
Length of stay, med (IQR)	11 (6–20)	10 (6–14)	10 (6–19)	11 (5–24)	12 (7–22)
Charlson Comorbidity Index†, med (IQR)	2 (0–3)	1 (0–2)	2 (1–3)	1 (0–2)	2 (0–3)
Risk factors at sepsis onset					
Prior surgery (30 days)‡, No. (%)	82 (23.9)	6 (20.0)	32 (29.4)	11 (12.6)	33 (28.2)
Central venous catheter, No.(%)	87 (25.4)	8 (26.7)	33 (30.3)	13 (14.9)	33 (28.2)
Urinary catheter, No. (%)	80 (23.4)	11 (36.7)	28 (25.7)	18 (20.7)	23 (19.7)
Sepsis characteristic					
Community-onset sepsis§, No. (%)	277 (80.8)	27 (90.0)	78 (71.6)	77 (88.5)	95 (81.2)
Hospital-onset sepsis§, No. (%)	66 (19.2)	3 (10.0)	31 (28.4)	10 (11.5)	22 (18.8)
ICU admission¶, No. (%)	52 (15.2)	10 (33.3)	14 (12.8)	9 (10.3)	19 (16.2)
SOFA baseline, med (IQR)	0 (0–1)	0(0)	0 (0–1)	0 (0–1)	0 (0–1)
SOFA Max, med (IQR)	4(2–5)	3(2–4)	3(2–4)	4(2–5)	4(3–5)
Shock**, No. (%)	30 (8.7)	0 (0)	10 (9.2)	7 (8.0)	13 (11.1)
Neutropenia††, No. (%)	46 (13.4)	2 (6.7)	19 (17.4)	9 (10.3)	16 (13.7)
Bloodstream infection,‡‡ No. (%)	65 (19.0)	0 (0)	2 (1.8)	7 (8.0)	56 (47.9)
Complete course of antimicrobials, No. (%)	317 (92.4)	12 (40.0)	105 (96.3)	85 (97.7)	115 (98.3)
Source of infection					
Respiratory infection, No. (%)	119 (34.7)	0 (0)	45 (41.3)	58 (66.7)	16 (13.7)
Urogenital infection, No. (%)	54 (15.7)	0 (0)	9 (8.3)	20 (23.0)	25 (21.4)
Unknown source, No. (%)	42 (12.2)	0 (0)	42 (38.5)	0 (0)	0 (0)
Bloodstream infection, No. (%)	35 (10.2)	0 (0)	0 (0)	1 (1.1)	34 (29.1)
Skin, bone and joint infection, No. (%)	30 (8.7)	0 (0)	4 (3.7)	0 (0)	26 (22.2)
Abdominal, No. (%)	26 (7.6)	0 (0)	7 (6.4)	6 (6.9)	13 (11.1)
Other infection, No. (%)	7 (2.0)	0 (0)	2 (1.8)	2 (2.3)	3 (2.6)
In-hospital mortality, No. (%)	40 (11.6)	1 (3.3)	16 (14.7)	13 (14.9)	10 (8.5)
ICD-10 code for sepsis,§§ No. (%)	59 (17.2)	1 (3.3)	4 (3.7)	13 (14.9)	41 (35.0)

*Sepsis-3 defined as any culture taken and administration of two doses antimicrobials combined with change in SOFA score of two points or more during 48 hours before and 24 hours after onset of infection compared with baseline SOFA score calculated before this time window.
†Weighted Charlson comorbidity index.[50] Total modified Charlson score: 0–24.
‡Prior surgery was generated using procedure codes. Time zero was sepsis onset.
§Hospital-onset defined as an episode 48 hours after admission or if the patient was readmitted with sepsis within 48 hours of discharge. All other episodes were defined as community-onset.
¶ICU admission at any time during hospitalisation.
**Shock defined as patients receiving vasopressors.
††Neutropenia defined as absolute neutrophil count less than $0.5x10^9$/L.
‡‡Bloodstream infection at any time during hospitalisation (online supplement methods 2).
§§International Classification of Diseases (ICD)-10 codes including A02.1, A22.7, A26.7, A32.7, A39.2, A39.4, A40.x, A41.x, A42.7, A48.3, B37.7, M72.6, R57.2, R65.1 and R65.9.
ICU, intensive care unit; med, median; SOFA, Sequential Organ Failure Assessment.

(range between 20.2%–46.8% for liver, renal, coagulation and respiration), but a large portion of patients had a measured baseline value (range between 18.1%–56.0% for liver, respiration, coagulation and renal) (online supplementary table 4). In hospital-onset suspected infection, only 0.4%–3.2% (all SOFA score components) had an assumed normal baseline value.

Only 13.4% of sepsis episodes had an ICD-10 code indicating sepsis. The in-hospital mortality was 8.6% for all sepsis episodes, 8.0% for CO sepsis and 12.7% for HO-sepsis, compared with 2.4% in the entire hospital cohort. The incidence was 9.1 (95% CI: 8.9 to 9.3) per 100 admissions for CO sepsis and 2.6 (95% CI: 2.4 to 2.8) per 1000 patient days for HO sepsis, with a CIF of 0.013 at day 30 for HO sepsis in the competing risk model (online supplementary figure 5). The cumulative incidence of HO sepsis varied significantly depending on type of hospital ward, with the highest risk in Transplant (CIF=0.078) and Haematology (CIF=0.061) wards, and the lowest risk in Orthopaedic (CIF=0.004) wards (figure 1 and online supplement figures 6 and 7). In-hospital mortality

Table 2 Performance of the surveillance algorithm using different definitions of suspected infection

Definition of suspected infection	Entire hospital cohort (n=82 653)				Suspected infection validation cohort (n=674)			
	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Any culture and two doses of antimicrobials (equivalent to Sepsis-3 clinical criteria)	0.887 (0.799 to 0.964)	0.985 (0.978 to 0.991)	0.881 (0.833 to 0.926)	0.986 (0.973 to 0.996)	0.926 (0.896 to 0.955)	0.893 (0.859 to 0.923)	0.881 (0.833 to 0.926)	0.934 (0.895 to 0.969)
Any culture and 4 days of antimicrobials	0.853 (0.761 to 0.935)	0.988 (0.981 to 0.993)	0.899 (0.848 to 0.945)	0.981 (0.968 to 0.992)	0.891 (0.855 to 0.925)	0.915 (0.885 to 0.942)	0.899 (0.848 to 0.945)	0.907 (0.866 to 0.945)
Blood culture and two doses of antimicrobials	0.739 (0.639 to 0.832)	0.990 (0.985 to 0.995)	0.906 (0.853 to 0.952)	0.967 (0.952 to 0.980)	0.772 (0.724 to 0.818)	0.931 (0.904 to 0.956)	0.906 (0.853 to 0.952)	0.826 (0.781 to 0.871)
Blood cultures and 4 days of antimicrobials (equivalent to Adult Sepsis Event definition)	0.718 (0.615 to 0.813)	0.992 (0.986 to 0.996)	0.917 (0.862 to 0.962)	0.965 (0.949 to 0.978)	0.749 (0.699 to 0.796)	0.942 (0.917 to 0.964)	0.917 (0.862 to 0.962)	0.814 (0.769 to 0.859)

NPV, negative predictive value; PPV, positive predictive value.

after HO sepsis was highest in internal medicine wards (17.3%) compared with only 2.1% in thoracic surgery wards (online supplementary table 5).

Sensitivity analysis demonstrating variations in number of sepsis episodes and in-hospital mortality using different definitions of suspected infection are shown in figure 2. Mortality in sepsis episodes fulfilling only the definition used in the Sepsis-3 clinical criteria (any culture and two doses of antimicrobials), but not the definition used in the ASE (blood cultures and four days of antimicrobials), was 8.4% (n=174/2066). This was not significantly different from mortality 8.6% (n=563/6533) in sepsis episodes fulfilling both definitions (p=0.78 for difference).

DISCUSSION

In this study, we show that it is possible to build a fully-automated sepsis surveillance system based on the Sepsis-3 clinical criteria that correctly captures almost 90% of sepsis episodes occurring outside the ICU and assigns the events to space (ward) and time (onset). The mortality and patient characteristics in our study were similar to the studies used when developing the Sepsis-3 definition, speaking in favour of our results being generalisable to the European and US setting.[13] The usefulness of the algorithm was shown by indicating variations in HO sepsis incidence and mortality depending on ward type, which can be used to inform infection prevention interventions and improve sepsis care.

The sepsis definition is based on the pathophysiological response to an infection and is neither constrained to a certain type of infection nor does it require that the infection is microbiologically confirmed.[12] Quality improvement initiatives focusing on education and sepsis care bundles have been associated with survival benefits, warranting structured approaches in sepsis care.[23 24] In our study, the majority of patients presented with sepsis on admission, but the burden of HO sepsis was still substantial. Recent data have associated HO sepsis with mortality approximately twice as high compared with CO sepsis.[25] Despite this, traditional surveillance programmes for healthcare-associated infections, such as CDC/National Healthcare Safety Network and European Centre for Disease Prevention and Control, do not include sepsis as a distinct entity.[19 26 27]

Initiatives to monitor sepsis incidence have often focussed on using administrative hospital data, such as discharge diagnosis, trigger based audits or reporting to clinical databases, all carrying risk of bias and making comparisons between hospitals difficult.[28–30] The use of ICD-codes for sepsis surveillance is associated with considerable uncertainty[31 32] and studies indicate that some of the increased incidence of sepsis during the last decade can be explained by changes in coding practices.[33–37] Overall, epidemiological surveillance based on explicit sepsis ICD-codes seems to underestimate

Table 3 Characteristics of fully-automated sepsis incidence surveillance in a general hospital population

	All	No sepsis	Sepsis-3 clinical criteria	Community-onset* Sepsis-3 clinical criteria	Hospital-onset* Sepsis-3 clinical criteria
Hospital admissions, No.	82 653	74 054	8599	7493	1106
Patients, No.	54 884	51 343	7286	6472	1055
Female sex, No. (%)	27 928 (50.9)	26 378 (51.4)	3213 (43.9)	2876 (44.2)	430 (41.4)
Age, med (IQR)	64.0 (47.0–75.0)	63.0 (47.0–74.0)	70.0 (59.0–80.0)	70.0 (60.0–81.0)	67.0 (54.0–76.0)
Length of stay, med (IQR)	3.8 (2.0–7.6)	3.3 (1.9–6.9)	8.0 (4.2–15.7)	7.0 (4.0–12.6)	23.0 (14.7–36.7)
Charlson Comorbidity Index†, med (IQR)	0 (0–2)	0 (0–2)	2 (0–3)	2 (0–2)	2 (0–3)
Comorbidities‡, No. (%)					
Chronic pulmonary disease	2370 (4.3)	2103 (4.1)	692 (9.5)	645 (10.0)	76 (7.2)
Cancer	14 036 (25.6)	13 094 (25.5)	2601 (35.7)	2240 (34.6)	514 (48.7)
Cerebral vascular disease	4061 (7.4)	3783 (7.4)	670 (9.2)	580 (9.0)	123 (11.7)
Chronic heart failure	2887 (5.3)	2643 (5.1)	792 (10.9)	708 (10.9)	113 (10.7)
Myocardial infarction	2638 (4.8)	2485 (4.8)	403 (5.5)	361 (5.6)	49 (4.6)
Connective tissue disease	1471 (2.7)	1360 (2.6)	287 (3.9)	259 (4.0)	42 (4.0)
Diabetes mellitus	1681 (3.1)	1541 (3.0)	442 (6.1)	394 (6.1)	65 (6.2)
HIV infection	54 (0.1)	46 (0.1)	18 (0.2)	17 (0.3)	2 (0.2)
Kidney disease	2222 (4.0)	1998 (3.9)	644 (8.8)	565 (8.7)	111 (10.5)
Liver disease	1281 (2.3)	1173 (2.3)	340 (4.7)	294 (4.5)	62 (5.9)
Prior surgery (30 days)§, No. (%)	12 274 (14.9)	10 091 (13.6)	2186 (25.4)	1458 (19.5)	728 (65.8)
Suspected infection¶, No. (%)	19 479 (23.6)	10 880 (14.7)	8599 (100.0)	7493 (100.0)	1106 (100.0)
Sepsis**, No. (%)					
Sepsis-3 clinical criteria	8599 (10.4)	0 (0.0)	8599 (100.0)	7493 (100.0)	1106 (100.0)
ICD-10 coded	2055 (2.5)	907 (1.2)	1148 (13.4)	939 (12.5)	209 (18.9)
Community-onset sepsis, No. (%)	7493 (9.1)	0 (0.0)	7493 (87.1)	7493 (100.0)	0 (0.0)
Hospital-onset sepsis, No. (%)	1106 (1.3)	0 (0.0)	1106 (12.9)	0 (0.0)	1106 (100.0)
ICU admission††, No. (%)	4335 (5.2)	3471 (4.7)	864 (10.0)	578 (7.7)	286 (25.9)
ICU days, med (IQR)	1.4 (0.9–4.1)	1.2 (0.9–3.5)	2.8 (1.1–6.3)	2.1 (1.0–4.6)	4.8 (2.0–11.4)
Bloodstream infection‡‡, No. (%)	2659 (3.2)	1104 (1.5)	1555 (18.1)	1279 (17.1)	276 (25.0)
In-hospital mortality, No. (%)	1953 (2.4)	1216 (1.6)	737 (8.6)	596 (8.0)	141 (12.7)

*Hospital-onset defined as a sepsis episode 48 hours after admission or if the patient was readmitted with sepsis within 48 hours of discharge. All other sepsis episodes were defined as community-onset.
†Weighted Charlson comorbidity index.[50] Total modified Charlson score: 0–24.
‡Comorbidity defined according to International Classification of Diseases (ICD)-10 codes registered within five years prior to hospitalisation.[51]
§Prior surgery was generated using procedure codes. For sepsis patients time zero was sepsis onset. For non-sepsis patients time zero was admission.
¶Defined as any culture taken and administration of two doses antimicrobials.
**Sepsis-3 defined as suspected infection combined with change in SOFA score of two points or more. ICD-10 codes including A02.1, A22.7, A26.7, A32.7, A39.2, A39.4, A40.x, A41.x, A42.7, A48.3, B37.7, M72.6, R57.2, R65.1 and R65.9.
††Intensive care unit admission any time during hospitalisation.
‡‡Bloodstream infection any time during hospitalisation (online supplement methods 2).
ICU, intensive care unit; med, median; SOFA, Sequential Organ Failure Assessment.

the incidence of sepsis compared with using clinical data,[15] and in our study only 13.4% of sepsis patients had an ICD-code indicating sepsis. Similar findings have been observed in studies comparing medical record review to ICD-codes,[38 39] but manual medical record review is both resource intensive and associated with subjectivity and limited inter-rater agreement.[40 41] Recently, a case definition, ASE, was developed by CDC to facilitate automated sepsis surveillance using clinical data from EHR.[14] Compared with Sepsis-3, the ASE algorithm is based on different criteria for both suspected infection and organ dysfunction and tends to capture a patient population with higher mortality than the Sepsis-3 criteria.[15] The sensitivity and specificity of the ASE definition, when using Sepsis-3 as the reference standard, was 69.7% and 98.1% in a US

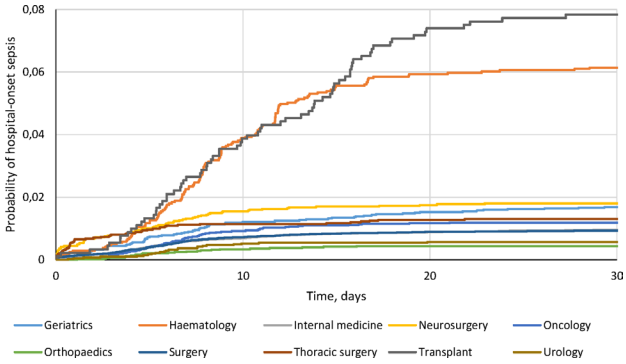


Figure 1 Cumulative incidence function (CIF) curves of hospital-onset sepsis stratified by ward type and taking into account competing risks ICU-admission, discharge or death. The CIF curves differed significantly in pairwise comparison (online supplementary figure 6). ICU, intensive care unit.

hospital setting,[15] compared with 88.7% and 98.5% for our algorithm in a European hospital. Using the ASE definition of suspected infection (blood culture and four days of antimicrobials) in our cohort resulted in 71.8% sensitivity and 99.2% specificity.

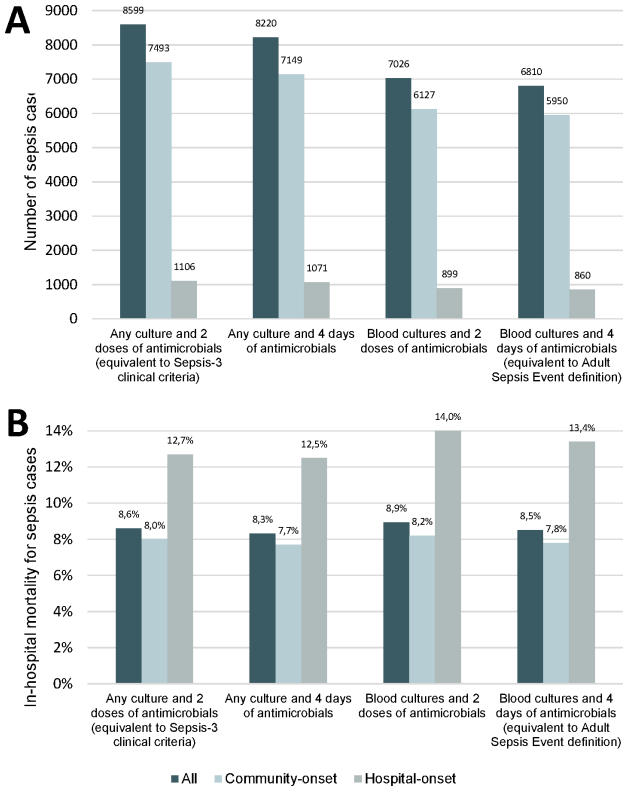


Figure 2 Effect on number of sepsis episodes and in-hospital mortality depending on different definitions of suspected infection. (2A) shows number of sepsis episodes per definition of suspected infection. (2B) shows in-hospital mortality (%) for sepsis cases per definition of suspected infection. 'Any culture and two doses of antimicrobials' is equivalent to the definition of suspected infection used in the Sepsis-3 clinical criteria. only 'blood cultures and four days of antimicrobials' is equivalent to the definition of suspected infection used in the Adult Sepsis Event (ASE) criteria. Note that in some episodes, time of onset of infection differed depending on the definition of suspected infection. This affected the time window for assessing organ dysfunction, which in a few cases resulted in differences in the classification of sepsis.

In this study, 91% of patients with sepsis according to the Sepsis-3 clinical criteria had either a *possible*, *probable* or *definite infection* as determined by physician review of medical records in post-hoc assessment, which is similar to a previous report from the ICU.[17] This suggests that the Sepsis-3 criteria perform well in capturing a patient population where clinicians maintain a suspicion of infection also after the initial treatment phase. Organ dysfunction in the Sepsis-3 clinical criteria is determined by SOFA score and concerns have been raised that this is not suited for EHR-based surveillance due to the inclusion of parameters not frequently measured in most patients.[42] However, integration of automated SOFA score calculators in EHR systems have shown strong agreement with manual score calculations,[43] limiting the need to use other criteria for organ dysfunctions. The SOFA score is based on assessment of six organ systems, compared with ASE that assesses five organ systems (CNS dysfunction is omitted). For respiratory and cardiovascular dysfunction, the ASE requires initiation of mechanical ventilation and vasopressor treatment. This biases sepsis surveillance towards patients eligible for aggressive treatment and access to ICU care, limiting generalisability to all hospitalised patients. One of the arguments for abandoning the Sepsis-3 definition in ASE was to facilitate widespread use to hospitals with limited collection of EHR data. However, the only additional data used in our surveillance case definition was vital parameters, which are routinely collected in many hospitals.

We show that it is feasible to use a surveillance algorithm based on the Sepsis-3 clinical criteria to automatically identify sepsis with high sensitivity and predictive values in non-ICU wards, and thus keeping a uniform sepsis definition for EHR surveillance. The objective of such surveillance is not early bedside sepsis recognition, but rather making continuously collected data on disease burden and patient management easily available. The possible use-cases of such surveillance data are multifaceted. First, incidence data presented down to the single-ward level as shown in this study, creates important feedback loops, which can guide quality improvement interventions, such as education programmes, systems for earlier sepsis recognition, treatment bundles and targeted infection control measures. Second, since Sepsis-3 based surveillance criteria do not require four days of antimicrobial treatment, but two doses, feedback on patients that have developed sepsis can be presented to clinicians early in the treatment course. This facilitates optimisation of care beyond the very initial treatment phase, such as better source control, adequate diagnostics, optimised antimicrobial treatment, infectious diseases specialist consultation and targeted rehabilitation, all of which have the potential to improve patient outcomes.[44 45] In

addition, 9% of patients fulfilling the Sepsis-3 clinical criteria did not have an infection, and 32% had only a possible infection, indicating the possibility to use this type of surveillance system as part of an antimicrobial stewardship programme, to balance the empirical broad spectrum antimicrobial treatment imposed by guidelines such as the Surviving Sepsis Campaign Bundle.[46–48]

Strength and limitations

A strength of our study is the use of a large clinical dataset representative of the population in a defined catchment area. This is, to our knowledge the first report of a sepsis surveillance system using the Sepsis-3 clinical criteria as case-definition, which overlaps better with other standards used for early sepsis recognition. This enables the integration of surveillance data in the direct clinical care of individual patients which can encourage clinicians to use the data, as opposed to implementing criteria developed exclusively for retrospective surveillance and thus risking to disconnect surveillance from the everyday clinical work. When developing and validating the algorithm, we used a duplicate of the EHR system, to ensure that our model can be implemented using real-time patient data. We could follow each subject over time and were not limited to data from the current hospital admission. This improves proper calculation of baseline organ dysfunction, which has been a limitation in previous methods.[14] Furthermore, we performed medical record review, showing that a rule-based surveillance algorithm performed well in non-ICU wards where data is usually of lower resolution and quality. This demonstrates that automated sepsis surveillance using the Sepsis-3 clinical criteria can be done without the need for complex computational methods such as text mining of unstructured data in EHR notes.

Limitations of fully-automated surveillance systems include possible misclassification of sepsis since both the algorithm and validation with medical record review depends on correct and accessible data in the EHR system. Not all hospital admissions with suspected infection contained the measurements necessary to assess a complete SOFA score, leading to missing data. By definition in the Sepsis-3 clinical criteria, missing values of SOFA score components were assumed to be normal, which may have affected correct classification of organ dysfunction and sepsis. Even though our validation sample included hospital admissions with and without suspected infection, our reference standard was based on infections recognised by clinicians and we may have missed sepsis cases among patients where an infection passed unnoticed. Sepsis classification can be affected by updates and changes in the EHR system, as well as by differences in recordings and access of data between wards, which could have influenced our results. This may explain the decreased algorithm specificity and PPV when restricting analyses to only hospital admissions including an ICU admission, from where we did not have access to data on medications and vital parameters. Since we did not include patient risk-time while in ICUs or obstetrical wards, our results cannot be generalised to such settings and inference on the true sepsis incidence is uncertain and should be interpreted with caution. It is also possible that patients' characteristics, such as organ dysfunction and source of infection, may be different for sepsis developing in these wards. Yet, in the ICU, documentation is usually both extensive and of good quality and a similar surveillance system has performed well in this setting.[49] The algorithm also showed lower sensitivity for HO sepsis compared with CO sepsis. This was primarily due to organ dysfunction only mentioned in free text, which indicates that improved recording of oxygen therapy and vital parameters such as GCS could result in better algorithm performance in surveillance of HO sepsis. Moreover, an implemented surveillance system requires continuous maintenance and validation. Although we used an exact duplicate of the EHR system, our algorithm has not yet been implemented and also needs evaluation in a real-world scenario. Finally, the study was limited to a single centre and needs confirmation within different EHR systems in different hospitals.

Conclusion

Based on data from EHR, it is feasible to automatically monitor sepsis incidence with good validity compared with physician medical record review in non-intensive care wards using the Sepsis-3 clinical criteria as surveillance definition. The algorithm exposed variations in hospital-onset sepsis incidence depending on ward type, which can be used to tailor infection prevention interventions and improve sepsis care.

Author affiliations
[1]Division of Infectious Diseases, Department of Medicine, Solna (MedS), Karolinska Institutet, Stockholm, Sweden
[2]Department of Infectious Diseases, Karolinska University Hospital, Stockholm, Sweden
[3]Treat Systems ApS, Aalborg, Denmark
[4]Center for Model-based Medical Decision Support, Department of Health Science and Technology, Aalborg University, Aalborg, Denmark
[5]Department of Clinical microbiology and the Laboratory for Molecular Infection Medicine (MIMS), Umeå University, Umeå, Sweden
[6]Department of Anesthesiology and Perioperative medicine, Mayo Clinic, Rochester, Minnesota, USA
[7]Department of Computer and Systems Sciences, Stockholm University, Kista, Sweden

Contributors Concept and design: PN, JKV, LW, VH, BP, AF, AT, AFJ, MLM, HD, AH. Acquisition, analysis, or interpretation of data: JKV, LW, PN, EA, HT, KM, MLM, HD, AH. Drafting of the manuscript: JKV, PN, LW. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: LW, JKV, PN. Obtained funding: PN, AF, AFJ, MLM, HD, JKV, BP, VH.

Funding The work was supported by Vinnova (grant 2016-00563). JKV was supported by Region Stockholm (combined clinical residency and PhD training program). PN was supported by Region Stockholm (clinical research

appointment). JKV, PN, VH and BP received the Mayo Clinic-Karolinska Institutet Collaborative Travel Award 2017.

Competing interests LW and MLM are employees of Treat Systems ApS (Aalborg, Denmark). Treat Systems produces medical decision support systems for antimicrobial and microbiological diagnostic stewardship.

Patient consent for publication Not required.

Ethics approval The study was approved by the Regional Ethical Review Board in Stockholm under permission no. 2016/2309-32 and 2012/1838-31/3

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data from deidentified electronic health records are not freely available due to protection of the personal integrity of the participants. Access to patient level data requires a Swedish ethical permit and an agreement with the research organisation, Department of Computer and Systems Sciences, Stockholm University, holder of the data.

ORCID iDs
John Karlsson Valik http://orcid.org/0000-0002-4521-1886
Logan Ward http://orcid.org/0000-0002-2697-3922
Anders F Johansson http://orcid.org/0000-0003-0548-594

REFERENCES

1 Fleischmann C, Scherag A, Adhikari NKJ, *et al*. Assessment of global incidence and mortality of Hospital-treated sepsis. current estimates and limitations. *Am J Respir Crit Care Med* 2016;193:259–72.
2 Mellhammar L, Wullt S, Lindberg Åsa, *et al*. Sepsis incidence: a population-based study. *OFIDS* 2016;3:ofw207.
3 Gotts JE, Matthay MA. Sepsis: pathophysiology and clinical management. *BMJ* 2016;353:i1585.
4 Page DB, Donnelly JP, Wang HE. Community-, Healthcare-, and hospital-acquired severe sepsis hospitalizations in the University HealthSystem Consortium. *Crit Care Med* 2015;43:1945–51.
5 van Mourik MSM, Perencevich EN, Gastmeier P, *et al*. Designing surveillance of healthcare-associated infections in the era of automation and reporting mandates. *Clin Infect Dis* 2018;66:970–6.
6 Sips ME, Bonten MJM, van Mourik MSM. Automated surveillance of healthcare-associated infections. *Curr Opin Infect Dis* 2017;30:425–31.
7 van Mourik MSM, van Duijn PJ, Moons KGM, *et al*. Accuracy of administrative data for surveillance of healthcare-associated infections: a systematic review. *BMJ Open* 2015;5:e008424.
8 Rhee C, Dantes RB, Epstein L, *et al*. Using objective clinical data to track progress on preventing and treating sepsis: CDC's new 'Adult Sepsis Event' surveillance strategy. *BMJ Qual Saf* 2019;28:305–9.
9 Rhee C, Gohil S, Klompas M. Regulatory mandates for sepsis care — reasons for caution. *N Engl J Med* 2014;370:1673–6.
10 Cohen J, Vincent J-L, Adhikari NKJ, *et al*. Sepsis: a roadmap for future research. *Lancet Infect Dis* 2015;15:581–614.
11 Reinhart K, Daniels R, Kissoon N, *et al*. Recognizing sepsis as a global health priority — a who resolution. *N Engl J Med* 2017;377:414–7.
12 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.
13 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.
14 Centers for Disease Control and Prevention. Hospital toolkit for adult sepsis surveillance, 2018. Available: https://www.cdc.gov/sepsis/clinicaltools/index.html [Accessed 10 Jul 2019].
15 Rhee C, Dantes R, Epstein L, *et al*. Incidence and trends of sepsis in US hospitals using clinical vs claims data, 2009-2014. *JAMA* 2017;318:1241–9.
16 Dalianis H, Henriksson A, Kvist M, *et al*. HEALTH BANK - A Workbench for Data Science Applications in Healthcare. *Proceedings of the CAiSE-2015 Industry Track co-located with 27th Conference on Advanced Information Systems Engineering* 2015;1381:18–1.
17 Klein Klouwenberg PM, Cremer OL, van Vught LA, *et al*. Likelihood of infection in patients with presumed sepsis at the time of intensive care unit admission: a cohort study. *Crit Care* 2015;19:319.
18 Klein Klouwenberg PM, Ong DSY, Bos LDJ, *et al*. Interobserver agreement of centers for disease control and prevention criteria for classifying infections in critically ill Patients*. *Crit Care Med* 2013;41:2373–8.
19 Horan TC, Andrus M, Dudeck MA. CDC/NHSN surveillance definition of health care–associated infection and criteria for specific types of infections in the acute care setting. *Am J Infect Control* 2008;36:309–32.
20 Calandra T, Cohen J. The International sepsis forum consensus conference on definitions of infection in the intensive care unit. *Crit Care Med* 2005;33:1538–48.
21 Wolkewitz M, Cooper BS, Bonten MJM, *et al*. Interpreting and comparing risks in the presence of competing events. *BMJ* 2014;349:g5060.
22 Fine JP, Gray RJ. A proportional hazards model for the Subdistribution of a competing risk. *J Am Stat Assoc* 1999;94:496–509.
23 Ferrer R, Artigas A, Levy MM, *et al*. Improvement in process of care and outcome after a multicenter severe sepsis educational program in Spain. *JAMA* 2008;299:2294–303.
24 Seymour CW, Gesten F, Prescott HC, *et al*. Time to treatment and mortality during mandated emergency care for sepsis. *N Engl J Med* 2017;376:2235–44.
25 Rhee C, Wang R, Zhang Z, *et al*. Epidemiology of Hospital-Onset versus community-onset sepsis in U.S. hospitals and association with mortality: a retrospective analysis using electronic clinical data. *Crit Care Med* 2019;47:1169–76.
26 European Centre for Disease Prevention and Control (ECDC). Point prevalence survey of healthcare-associated infections and antimicrobial use in European acute care hospitals – protocol version 5.3 2016.
27 Center for Disease Control and Prevention (CDC). National healthcare safety network (NHSN). Available: https://www.cdc.gov/nhsn/acute-care-hospital/index.html [Accessed 10 Jul 2019].
28 Prescott HC, Cope TM, Gesten FC, *et al*. Reporting of sepsis cases for performance measurement versus for reimbursement in New York State*. *Crit Care Med* 2018;46:666–73.

29 Paoli CJ, Reynolds MA, Sinha M, *et al*. Epidemiology and costs of sepsis in the United States—An analysis based on timing of diagnosis and severity Level*. *Crit Care Med* 2018;46:1889–97.
30 Rhee C, Jentzsch MS, Kadri SS, *et al*. Variation in identifying sepsis and organ dysfunction using administrative versus electronic clinical data and impact on hospital outcome Comparisons*. *Crit Care Med* 2019;47:493–500.
31 Gaieski DF, Edwards JM, Kallan MJ, *et al*. Benchmarking the incidence and mortality of severe sepsis in the United States*. *Crit Care Med* 2013;41:1167–74.
32 Wilhelms SB, Huss FR, Granath G, *et al*. Assessment of incidence of severe sepsis in Sweden using different ways of abstracting International classification of diseases codes: difficulties with methods and interpretation of results. *Crit Care Med* 2010;38:1442–9.
33 Walkey AJ, Lagu T, Lindenauer PK. Trends in sepsis and infection sources in the United States. A population-based study. *Ann Am Thorac Soc* 2015;12:216–20.
34 Fleischmann-Struzek C, Mikolajetz A, Schwarzkopf D, *et al*. Challenges in assessing the burden of sepsis and understanding the inequalities of sepsis outcomes between National health systems: secular trends in sepsis and infection incidence and mortality in Germany. *Intensive Care Med* 2018;44:1826–35.
35 Jafarzadeh SR, Thomas BS, Marschall J, *et al*. Quantifying the improvement in sepsis diagnosis, documentation, and coding: the marginal causal effect of year of hospitalization on sepsis diagnosis. *Ann Epidemiol* 2016;26:66–70.
36 Rhee C, Murphy MV, Li L, *et al*. Improving documentation and coding for acute organ dysfunction biases estimates of changing sepsis severity and burden: a retrospective study. *Crit Care* 2015;19:338.
37 Kadri SS, Rhee C, Strich JR, *et al*. Estimating Ten-Year Trends in Septic Shock Incidence and Mortality in United States Academic Medical Centers Using Clinical Data. *Chest* 2017;151:278–85.
38 Wang HE, Addis DR, Donnelly JP, *et al*. Discharge diagnoses versus medical record review in the identification of community-acquired sepsis. *Crit Care* 2015;19:42.
39 Henriksen DP, Laursen CB, Jensen TG, *et al*. Incidence rate of community-acquired sepsis among hospitalized acute medical Patients—A population-based Survey*. *Crit Care Med* 2015;43:13–21.
40 Rhee C, Kadri SS, Danner RL, *et al*. Diagnosing sepsis is subjective and highly variable: a survey of intensivists using case vignettes. *Crit Care* 2016;20:89.
41 Rhee C, Brown SR, Jones TM, *et al*. Variability in determining sepsis time zero and bundle compliance rates for the centers for Medicare and Medicaid services SEP-1 measure. *Infect Control Hosp Epidemiol* 2018;39:994–6.
42 Rhee C, Zhang Z, Kadri SS, *et al*. Sepsis surveillance using adult sepsis events simplified eSOFA criteria versus Sepsis-3 sequential organ failure assessment Criteria*. *Crit Care Med* 2019;47:307–14.
43 Aakre C, Franco PM, Ferreyra M, *et al*. Prospective validation of a near real-time EHR-integrated automated SOFA score calculator. *Int J Med Inform* 2017;103:1–6.
44 Paulsen J, Solligård E, Damås JK, *et al*. The Impact of Infectious Disease Specialist Consultation for *Staphylococcus aureus* Bloodstream Infections: A Systematic Review. *Open Forum Infect Dis* 2016;3:ofw048.
45 Schmitt S, McQuillen DP, Nahass R, *et al*. Infectious diseases specialty intervention is associated with decreased mortality and lower healthcare costs. *Clin Infect Dis* 2014;58:22–8.
46 Levy MM, Evans LE, Rhodes A. The surviving sepsis campaign bundle: 2018 update. *Crit Care Med* 2018;46:997–1000.
47 Fitzpatrick F, Tarrant C, Hamilton V, *et al*. Sepsis and antimicrobial stewardship: two sides of the same coin. *BMJ Qual Saf* 2019;28:758–61.
48 Kalil AC, Gilbert DN, Winslow DL, *et al*. Infectious diseases Society of America (IDSA) position statement: why IDSA did not Endorse the surviving sepsis campaign guidelines. *Clin Infect Dis* 2018;66:1631–5.
49 Johnson AEW, Aboab J, Raffa JD, *et al*. A comparative analysis of sepsis identification methods in an electronic Database*. *Crit Care Med* 2018;46:494–9.
50 Quan H, Li B, Couris CM, *et al*. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge Abstracts using data from 6 countries. *Am J Epidemiol* 2011;173:676–82.
51 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.

SUPPLEMENTS

Validation of automated sepsis surveillance based on the Sepsis-3 clinical criteria against physician record review in a general hospital population: observational study using electronic health records data

Valik JK, *et al. BMJ Qual Saf* 2020;0:1–11. doi: 10.1136/bmjqs-2019-010123

Supplementary Methods 1. Rule-based algorithm definitions

Suspected infection

We defined suspected infection as having any culture taken and at least 2 doses of antimicrobial treatment (ATC code J01 and J04) newly administrated either by the oral or parental route within 6-48 hours between the doses[1]. Treatment with the antimicrobials pivmecillinam, nitrofurantoin and trimethoprim were excluded from the definition since they are solely used to treat lower urinary tract infections and not sepsis. Cultures included body fluid from: abdomen, blood, bone, bronchoalveolar lavage, cerebral spinal fluid, catheters/devices, nasopharynx, pleural space, skin/tissue, sputum, stool, synovial fluid and urinary tract. Culture types included only bacterial culture and testing for *C. difficile* toxin, *Mycoplasma pneumoniae* DNA, EHEC DNA and Legionella antigen in urine. Fungal cultures were included only if collected from blood. Viral and parasitic samples were excluded.

If the patient was admitted to the ICU prior to 24 hours, or died prior to 48 hours from the first dose of antimicrobial treatment, they were deemed to have suspected infection despite of only 1 dose given. Cultures had to be performed within 24 hours after the start of antimicrobial treatment. Antimicrobial treatment had to be started within 72 hours after culture[1]. Onset of infection was defined as which of these events occurred first.

Sensitivity analyses were done with different definitions of suspected infections: only blood cultures and 2 doses of antimicrobial treatment, any culture and 4 calendar days of antimicrobial treatment and only blood cultures and 4 calendar days of antimicrobial treatment. To fulfil the 4 calendar day antimicrobial treatment criteria, we used a modified version of the Center of Disease Control and Prevention (CDC) Hospital Tool Kit for Adult Sepsis Surveillance[2]. At least one dose had to be administered intravenously within the window period. Subsequent calendar days could be the same antimicrobial, or a different antimicrobial as long as the first dose of each antimicrobial in the sequence was new. A new antimicrobial did not have to be started within the window period to be counted as part of the 4 calendar days. A gap of a single calendar day between administrations of the same antibiotic was counted as part of the 4 calendar days as long as the gap was not greater than 1 day. If the patient were admitted to the ICU, died or was discharged prior to 4 calendar days of antimicrobial treatment, they were deemed to have suspected infection anyway.

Sequential Organ Failure Assessment (SOFA) score

Organ dysfunction was measured with SOFA score counted during a time window beginning 48 hours before (limited by time of data availability) until 24 hours after onset of infection (limited by death or discharge)[1]. Worst values were registered and missing values were considered to be normal. The baseline SOFA score was defined as the latest value measured before the 72-hour time window, and was assumed to be 0 in patients not known to have pre-existing organ dysfunction (see below).

- ***SOFA respiration:*** Calculated from PaO2/FiO2 (mm Hg). If PaO2 was not available it was calculated from peripheral capillary oxygen saturation (SpO2) obtained from pulse oximetry via a conversion table which has been previously validated[3]. Prior studies have demonstrated feasibility to impute SpO2 when calculating the SOFA respirator score[4]. If no data on oxygen therapy was registered, FiO2 was be assumed to be 0.21. FiO2 values for patients receiving supplemental oxygen were estimated assuming each 1 L/min of oxygen flow rate increased FiO2 by 0.03 for the first L/min and 0.04 for consecutively L/min over room air. We used scoring cut-offs: ≥400 mm Hg for 0 points, <400 mm Hg for 1 point, <300 mm Hg for 2 points, <200 mm Hg for 3 points and <100 mm Hg for 4 points. For baseline SOFA respiratory, the latest measured PaO2 or SaO2, prior to the suspected infection window, during the last 3 months was used. Registration of home oxygen and/or ventilator treatment (ICD-codes DG008 and DG009) during the previous 1 year was considered default SOFA respiratory 2 points for both baseline and suspected infection window.

Valik JK, *et al. BMJ Qual Saf* 2020;0:1–11. doi: 10.1136/bmjqs-2019-010123

- ***SOFA cardiovascular:*** Calculated from the mean arterial blood pressure (MAP) (mm Hg). The MAP was calculated from systolic blood pressure (SBP) and simultaneously measured diastolic blood pressure (DBP) using formula (2*DBP+SBP)/3. We used scoring cut-offs: ≥70 mm Hg for 0 points and <70 mm Hg for 1 point. Since surveillance was performed outside ICUs, treatment with vasopressors were not used for the definition, meaning maximum score was 1. The baseline SOFA cardiovascular was the latest measured MAP before the suspected infection window. Only values measured during current hospitalization was used.

- ***SOFA central nervous system (CNS):*** Calculated from Glasgow Coma Scale (GCS). If GCS was not available, we used structured data on "alert" (interpreted as GCS score 15 points) or "not alert" (interpreted as GCS score 14 points). We used scoring cut-offs: GCS 15 for 0 points, GCS 13-14 for 1 point, GCS 10-12 for 2 points, GCS 6-9 for 3 points and GCS <6 for 4 points. The baseline SOFA CNS was the latest measured value before the suspected infection window. Only values measured during current hospitalization was used.

- ***SOFA coagulation:*** Calculated from platelets ($x10^3/\mu L$). We used scoring cut-offs: ≥150 $x10^3/\mu L$ for 0 points, <150 $x10^3/\mu L$ for 1 point, <100 $x10^3/\mu L$ for 2 points, <50 $x10^3/\mu L$ for 3 points and <20 $x10^3/\mu L$ for 4 points. For baseline SOFA coagulation, the latest measured platelets value, prior to the suspected infection window, during last 3 months was used.

- ***SOFA liver:*** Calculated from bilirubin (µmol/L). We used scoring cut-offs: <20 µmol/L for 0 points, 20-32 µmol/L for 1 point, 33-101 µmol/L for 2 points, 102-204 µmol/L for 3 points and >204 µmol/L for 4 points. For baseline SOFA liver, the latest measured bilirubin value, prior to the suspected infection window, during the last 3 months was used.

- ***SOFA renal:*** Calculated from creatinine (µmol/L). We used scoring cut-offs: <110 µmol/L for 0 points, 110-170 µmol/L for 1 point, 171-299 µmol/L for 2 points, 300-440 µmol/L for 3 points and >440 µmol/L for 4 points. For baseline SOFA renal, the latest measured creatinine value, prior to the suspected infection window, during last 3 months was used. Registration of chronic dialysis treatment (ICD-codes Z99.2, Z49.0, Z49.1 and Z49.2) during the previous 1 year was considered default SOFA renal 4 points for both baseline and suspected infection window. Urine output was not used as a measure due to data availability.

Supplementary Methods 2. Definition of significant bloodstream infection

All pathogens were regarded as bloodstream infection except pre-define contaminants species, if these were isolated in only one bottle or only one set if more than one set of blood cultures were collected within 24 hours. One set was defined as 1 anaerobe blood culture bottle and 1 aerobe blood culture bottle.

List of possible blood culture contaminants:

- *Alloiococcus otitis*
- Anaerobic bacteria
- *Bacillus cereus*
- *Bacillus species*
- *Bifidobacterium species*
- Coagulase-negative staphylococcus (CoNS)
- *Corynebacterium jeikeium*
- *Corynebacterium species*
- *Dermabacter hominis*
- *Desulfovibrio species*
- *Gardnerella vaginalis*
- *Gemella sanguinis*
- Gram negative coccus, anaerobe
- *Lactobacillus acidophilus*
- *Lactobacillus casei*
- *Lactobacillus gasseri*
- *Lactobacillus species*
- *Lactococcus lactis*
- *Leptotrichia species*
- *Leuconostoc lactis*
- *Leuconostoc species*
- *Micrococcus luteus*
- *Micrococcus species*
- *Propionibacterium acnes*
- *Propionibacterium species*
- *Staphylococcus epidermidis*

Supplementary Table 1. Reasons for imperfect algorithm performance of the surveillance algorithm in the validation sets

Reasons for imperfect sensitivity in patients with suspected infection (n=674)	**Number of cases (total n=23)**
Organ dysfunction only mentioned in free text	16[a]
Misclassified baseline SOFA and/or development of organ dysfunction related to the infection outside of 72-h suspected infection window	7
Reasons for imperfect specificity in patients with suspected infection (n=674)	**Number of cases (total n=39)**
No infection	29
Wrong baseline SOFA	7
Obvious measurement error of vital parameters in EHR	3
Reasons for imperfect sensitivity in patients without suspected infection (n=326)	**Number of cases (total n=2)**
Blood cultures performed by advanced home care services before arrival to the emergency department	1
Antimicrobial treatment not registered in the EHR medications module	1

[a]Among these, 6 was due to SOFA respiration, 5 was due to SOFA cns and 5 was due to combinations of SOFA respiration, cns and cardiovascular.

Supplementary Table 2. Performance of the surveillance algorithm stratified by ICU admission

	Entire hospital cohort (n=82 653)		**Suspected infection validation cohort (n=674)**	
	Episodes without ICU admission (n=78 318)	**Episodes with ICU admission (n=4335)**	**Episodes without ICU admission (n=603)**	**Episodes with ICU admission (n=71)**
Sensitivity [95% CI]	0.879 [0.793-0.952]	0.952 [0.881-1.000]	0.922 [0.888-0.952]	0.952 [0.881-1.000]
Specificity [95% CI]	0.988 [0.983-0.992]	0.938 [0.907-0.969]	0.913 [0.883-0.942]	0.655 [0.476-0.828]
PPV [95% CI]	0.895 [0.860-0.931]	0.800 [0.712-0.894]	0.895 [0.860-0.931]	0.800 [0.712-0.894]
NPV [95% CI]	0.985 [0.973-0.994]	0.987 [0.967-1.000]	0.936 [0.908-0.960]	0.905 [0.737-1.000]

Supplementary Table 3. Availability of SOFA score components in episodes with suspected infection

Table 4a. Baseline SOFA score availability[a] and timing, stratified by onset of infection

		SOFA respiration [b]	SOFA coagulation	SOFA cardio.	SOFA cns[c]	SOFA liver	SOFA renal
Community-onset	Percentage (%) of suspected infections	49.6	60.5	7.0	3.2	34.2	60.9
	Days before suspected infection for SOFA baseline measurement, med [IQR]	13.5 [3-35.2]	10.0 [2.5-26.5]	0.2 [0.1-0.5]	1.0 [0.4-1.5]	17.8 [6.2-39.4]	11.3 [3.3-28.0]
Hospital-onset	Percentage (%) of suspected infections	96.0	98.2	95.4	36.7	71.4	99.2
	Days before suspected infection for SOFA baseline measurement, med [IQR]	0.4 [0.1-2.0]	1.3 [0.4-3.3]	0.3 [0.1-0.9]	6.5 [3.5-13.1]	3.6 [1.1-10.3]	1.0 [0.3-2.2]

Table 4b. Suspected infection SOFA score (72-h window) availability[a], stratified by onset of infection

		SOFA respiration [b]	SOFA coagulation	SOFA cardio.	SOFA cns[c]	SOFA liver	SOFA renal
Community-onset	Percentage (%) of suspected infections	93.3	92.0	95.0	55.0	38.3	92.7
	Mean number of measurements	4.9	1.7	5.1	1.3	1.4	1.7
Hospital-onset	Percentage (%) of suspected infections	81.5	73.2	86.2	3.0	30.2	85.4
	Mean number of measurements	6.6	2.1	7.0	2.4	1.8	2.1

[a]Avaliability of SOFA score is presented as % of total number of suspected infections (n=21 201) with available data on the SOFA score component in 19 479 hospital admissions containing at least one suspected infection. Suspected infection episodes were registered up to and including an episode where there was sepsis, otherwise until discharge or death.

[b]In cases where baseline SOFA respiration was measured, SOFA respiration measurements were based on PaO2 in 0.012% and 0.025% of infections for CO and HO, respectively (1 case each). Within windows, SOFA respiration measurements were based on PaO2 in 0.019% and 0.087% of SOFA measurements for CO and HO infections, respectively (3 measurements each)

[c]In cases where baseline SOFA cns was measured, SOFA cns measurements were based on the Glasgow Coma Scale (GCS) in 7.0% and 6.1% of infections for CO and HO, respectively (38, 94 cases). Within windows, SOFA cns measurements were based on GCS in 8.9% and 5.5% of SOFA measurements for CO and HO infections, respectively (833, 7 measurements).

Supplementary Table 4. Paired missingness (baseline vs. suspected infection) for each of the SOFA components, stratified by place of acquisition

		SOFA respiration[b]	SOFA coagulation	SOFA cardio[c]	SOFA cns[c, d]	SOFA liver	SOFA renal
Community-onset	Measurement of SOFA score in both baseline and suspected infection (%)[a]	46.5	55.0	6.3	0.3	18.1	56.0
	Measurement of SOFA score in baseline only (%)[a]	3.4	5.5	0.7	2.9	16.0	5.0
	Measurement of SOFA score in suspected infection only (%)[a]	46.8	36.9	88.7	54.7	20.2	36.7
	Measurement of SOFA score in neither baseline nor suspected infection (%)[a]	3.3	2.5	4.4	42.1	45.6	2.3
Hospital-onset	Measurement of SOFA score in both baseline and suspected infection (%)[a]	79.8	72.5	84.3	2.0	27.0	85.1
	Measurement of SOFA score in baseline only (%)[a]	16.3	25.7	11.1	34.8	44.4	14.2
	Measurement of SOFA score in suspected infection only (%)[a]	1.8	0.6	1.9	1.1	3.2	0.4
	Measurement of SOFA score in neither baseline nor suspected infection (%)[a]	2.2	1.2	2.7	62.2	25.4	0.4

[a]Presented as % of total number of suspected infections (n=21 201) with available data on the SOFA score component in 19 479 hospital admissions containing at least one suspected infection. Suspected infection episodes were registered up to and including an episode where there was sepsis, otherwise until discharge or death.

[b]No infections had both baseline and suspected infection SOFA respiration measurements based on PaO2. For CO infections, 0.2% (1 case) of those with SOFA respiration at baseline only were based on PaO2 and 0.04% (3 cases) of those with SOFA respiration in suspected infection only were based on PaO2. For HO infections, 0.1% (1 case) of those with SOFA respiration at baseline only were based on PaO2 and 4.0% (3 cases) of those with SOFA respiration in suspected infection only were based on PaO2.

[c]Only measurements during the current hospital admission was used

[d]For CO infections, no infections had both baseline and suspected infection SOFA cns measurements based on GCS, 7.7% (38 cases) of those with SOFA cns at baseline only were based on GCS and 9.0% (833 cases) of those with SOFA cns in suspected infection only were based on GCS. For HO infections, 1.2% (1 case) of those with both baseline and suspected infection SOFA cns measurements were based on GCS, 6.3% (93 cases) of those with SOFA cns at baseline only were based on GCS and 13.0% (6 cases) of those with SOFA cns in suspected infection only were based on GCS.

Supplementary Table 5. The burden of hospital-onset sepsis and in-hospital mortality depending on definition of suspected infection

	Any culture and 2 doses of antimicrobials				Any culture and 4 days of antimicrobials				Blood cultures and 2 doses of antimicrobials				Blood cultures and 4 days of antimicrobials			
Ward	**n/N[a]**	**/1000d[b]**	**CIF[c]**	**Mortality (%)**	**n/N**	**/1000d**	**CIF**	**Mortality (%)**	**n/N**	**/1000d**	**CIF**	**Mortality (%)**	**n/N**	**/1000d**	**CIF**	**Mortality (%)**
Haematology	156/2379	9.5	0.061	12.2	152/2387	9.1	0.061	9.9	159/2410	9.6	0.063	11.3	153/2416	9.1	0.060	9.2
Transplant	71/902	8.6	0.078	8.5	73/956	8.0	0.075	8.2	63/981	6.6	0.064	11.1	63/983	6.5	0.063	9.5
Neurosurgery	70/3393	4.3	0.018	2.9	67/3399	4.1	0.017	1.5	42/3455	2.4	0.011	7.1	39/3457	2.3	0.010	5.1
Thoracic surgery	47/2151	3.7	0.013	2.1	39/2171	3.0	0.010	2.6	25/2180	1.9	0.006	4.0	21/2190	1.6	0.005	4.8
Oncology	66/5540	2.6	0.012	15.2	61/5553	2.4	0.011	16.4	59/5604	2.2	0.011	13.6	57/5616	2.2	0.010	14.0
Surgery	217/22563	2.4	0.009	7.4	213/22608	2.3	0.009	7.5	177/22825	1.9	0.007	9.0	170/22861	1.8	0.007	8.2
Internal medicine	324/32456	2.0	0.009	17.3	314/32670	2.0	0.009	17.5	253/33315	1.5	0.007	18.6	243/33477	1.5	0.007	18.5
Urology	23/4013	1.7	0.006	13.0	21/4013	1.5	0.005	14.3	20/4075	1.4	0.005	15.0	18/4078	1.3	0.004	16.7
Geriatrics	75/4315	1.5	0.017	28.0	73/4340	1.5	0.016	27.4	54/4494	1.0	0.011	31.5	52/4503	1.0	0.011	30.8
Orthopaedics	23/5254	0.9	0.004	4.3	24/5310	0.9	0.005	4.2	18/5359	0.7	0.003	5.6	17/5367	0.6	0.003	5.9
All wards[d,e]	1106/82653	2.6	0.013	12.7	1071/82653	2.5	0.013	12.5	899/82653	2.0	0.011	14.0	860/82653	1.9	0.010	13.4

[a] Number of hospital-onset sepsis episodes/number of hospital admissions. Note that the denominator changes in the same ward depending on definition of suspected infection. This is due to the fact that alterations in definition of suspected infection also affects the number of community-onset sepsis episodes. In the study, only the first sepsis episode is recorded.

[b] Hospital-onset sepsis episodes per 1000 patient days at risk.

[c] Cumulative incidence function (CIF) at day 30 accounting for competing risks: ICU-admission, discharge or death. CIF-curves using the Sepsis-3 clinical criteria are presented in Figure 1

[d] The number of hospital-onset sepsis episodes/number of hospital admissions are not the exact sum of all cases above. This is due to the fact that some sepsis cases, which fulfilled the definition of hospital-onset sepsis, had not yet been assigned a specific hospital ward at onset of sepsis, and that it was possible for single hospital admissions to be counted in the denominator of more than one ward.

[e] Effect on number of sepsis episodes and in-hospital mortality depending on different definitions of *suspected infection* are showed for both hospital-onset and community-onset sepsis in Figure 3**.**

Supplementary Figure 1. Illustration of how Sequential Organ Failure Assessment (SOFA) score was calculated in the algorithm

A DATA INPUT FROM ELECTRONIC HEALTH RECORDS

SOFA	ICD CODE DEFAULT	BASELINE	TIME 0	TIME 1	TIME 2	TIME 3	TIME 4	TIME 5
RESPIRATION	2	2	1	1	1	1	1	1
CARDIOVASCULAR	-	0	1	0	0	0	0	0
CNS	-	0	0	0	0	0	0	0
COAGULATION	-	1	0	0	0	0	0	2
LIVER	-	3	1	1	1	1	3	1
RENAL	0	2	1	1	3	4	2	1
TOTAL	2	8	4	3	5	6	6	5

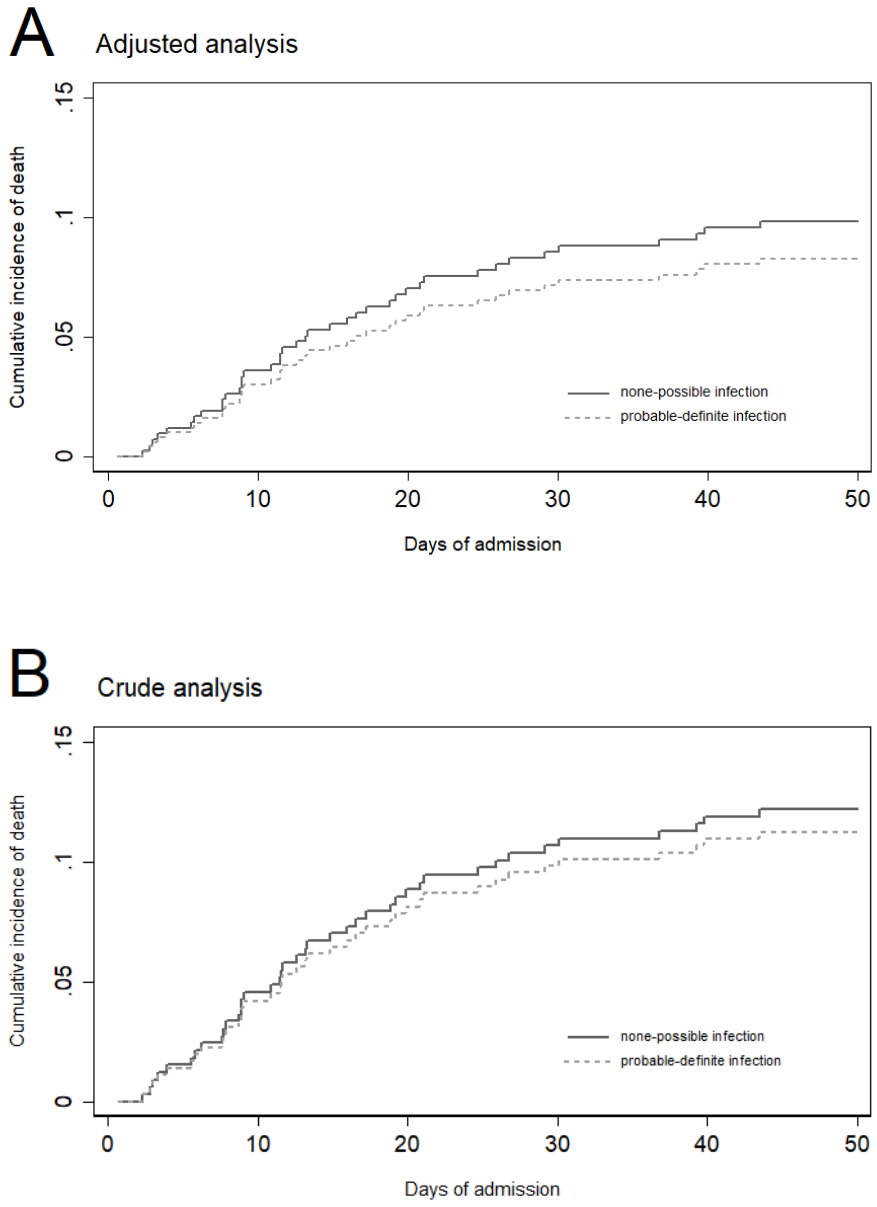
B TRANSFORMATION OF DATA INPUT TO FIT SEPSIS CRITERIA

SOFA	ICD CODE DEFAULT	BASELINE	TIME 0	TIME 1	TIME 2	TIME 3	TIME 4	TIME 5
RESPIRATION	2	2	2	2	2	2	2	2
CARDIOVASCULAR	-	0	1	1	1	1	1	1
CNS	-	0	0	0	0	0	0	0
COAGULATION	-	1	0	0	0	0	0	2
LIVER	-	3	1	1	1	1	3	3
RENAL	0	2	1	1	3	4	4	4
TOTAL	2	8	5	5	7	8	10*	12

* Sepsis criteria fulfilled

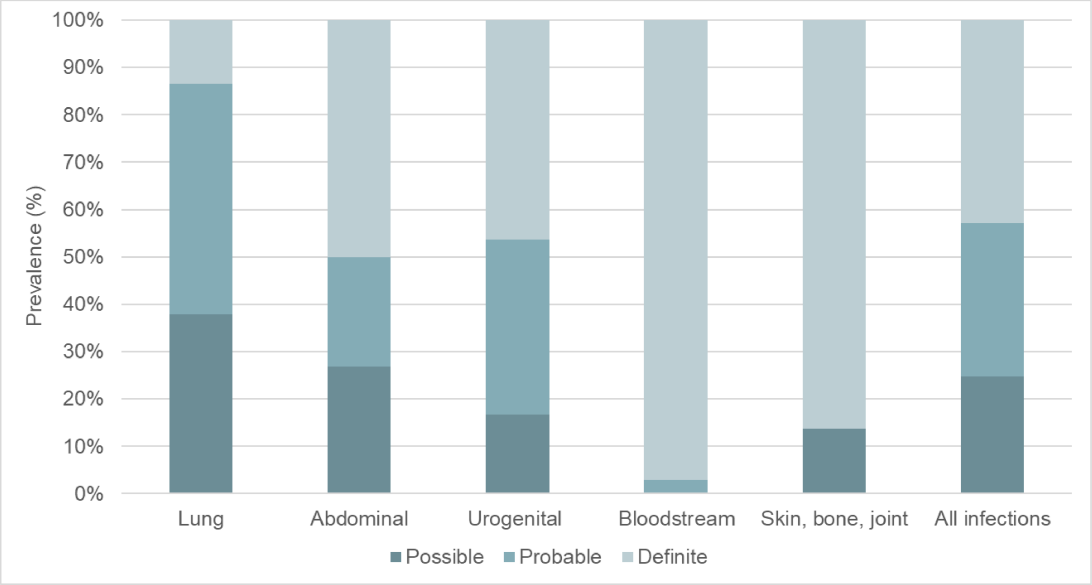
The example illustrates a mock-up patient. Figure A shows an example of how data input could be collected from the electronic health record. Figure B shows an example of how data input was transformed to calculate a maximum total SOFA score. SOFA score was calculated during two time periods, at baseline and during a time window beginning 48 hours (h) before until 24 h after onset of infection. The baseline SOFA score was defined as the latest value measured before the 72-h time window, and was assumed to be 0 in patients not known to have pre-existing organ dysfunction. For SOFA respiration, SOFA coagulation, SOFA liver and SOFA renal, measurements during the last 3 months were used to calculate the baseline SOFA. For SOFA cardiovascular and SOFA central nervous system (CNS) only measurements during the current hospitalization were used to calculate the baseline SOFA. To be able to calculate maximum SOFA score during the infection time window, worst values were carried forward and missing values were assumed to be 0. Pre-defined ICD-codes for SOFA respiratory and SOFA renal score during the previous year resulted in a default value.

Supplementary Figure 2. Cumulative incidence of death in sepsis cases stratified by likelihood of infection



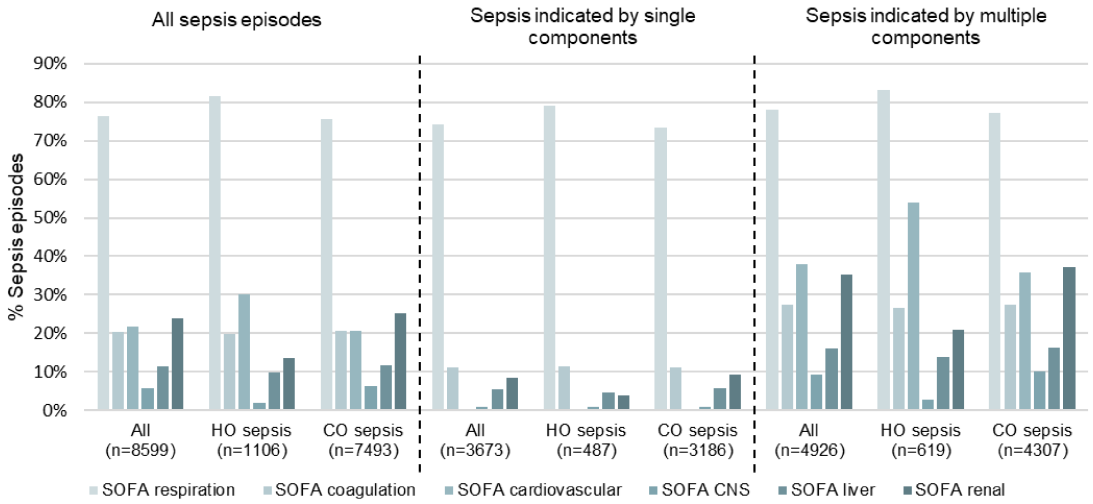
Cumulative incidence function (CIF) of death with discharge as competing risk and stratified by likelihood of infection (n=340 sepsis episodes from patients with suspected infection). Subjects were censored at day 50 (n=12). Figure A shows CIF adjusted for age, Charlson comorbidity index and community-/hospital-onset sepsis. Figure B shows the unadjusted CIF. In the adjusted model the CIF curves did not differ significantly (p=0.515).

Supplementary Figure 3. Likelihood of infection in patients fulfilling Sepsis-3 clinical criteria divided by source of infection



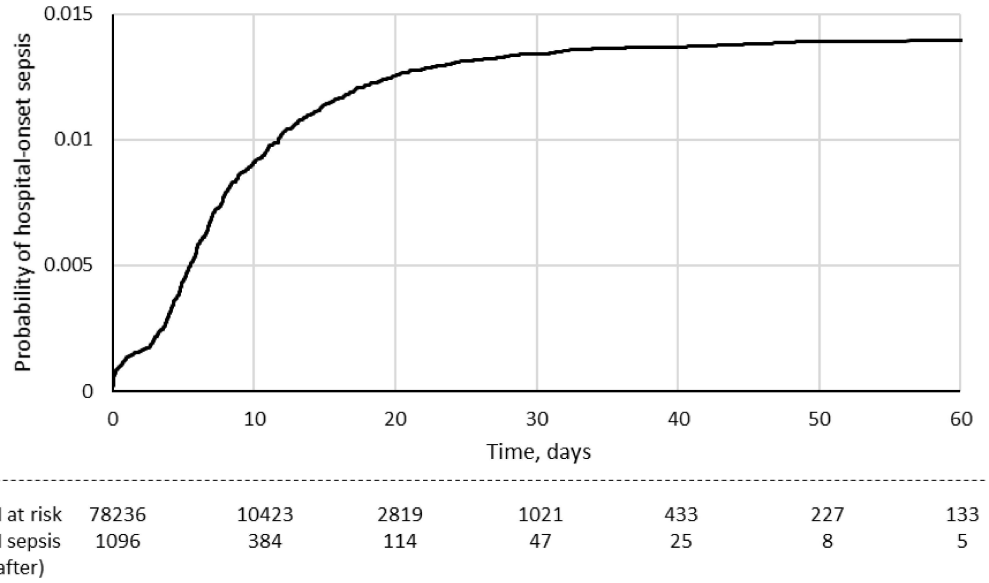
Likelihood of infection in patients fulfilling Sepsis-3 criteria in non-intensive care wards, judged by physician review of medical records. Only patients meeting the definitions for possible, probable and definite infection are presented (n=313/343). Likelihood of infection is divided by the most common sources of infection: lung (pneumonia, lung abscess/empyema, lower respiratory tract infections and upper respiratory tract infections), abdominal (peritonitis, biliary tract, intra-abdominal infection/abscess, pancreatic and gastrointestinal), urogenital (urinary tract infection and reproductive organs), bloodstream (primary bloodstream infection, vascular device infection and endocarditis), skin and bone infections (superficial skin infections, cellulitis, wound infection, bone and joint infection) and all infectious sources combined.

Supplementary Figure 4. Distribution of Sequential Organ Failure Assessment (SOFA) score triggers

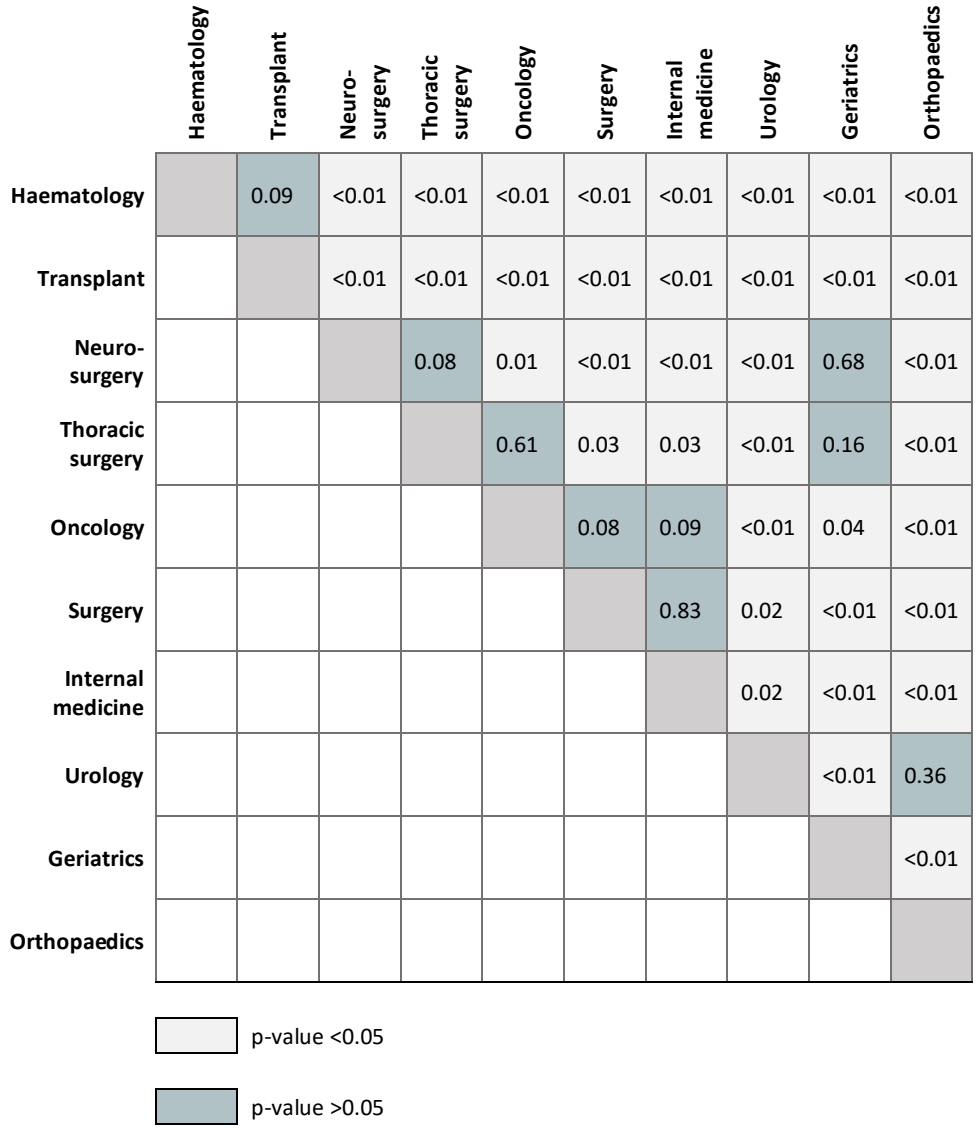


The percentage of sepsis episodes where each of the SOFA score components contributed to the sepsis classification are shown for all sepsis episodes (left), those where a single component was responsible for the sepsis classification (centre) and those where multiple components were responsible for the classification (right). Abbreviations: Hospital-onset (HO) and community-onset (CO).

Supplementary Figure 5. Cumulative incidence function curve of hospital-onset sepsis in all wards



Supplementary Figure 6. Table of p-values for pairwise significance testing of differences in cumulative incidence function between wards



Supplementary Figure 7. Monthly sepsis incidence surveillance in the whole hospital and selected wards

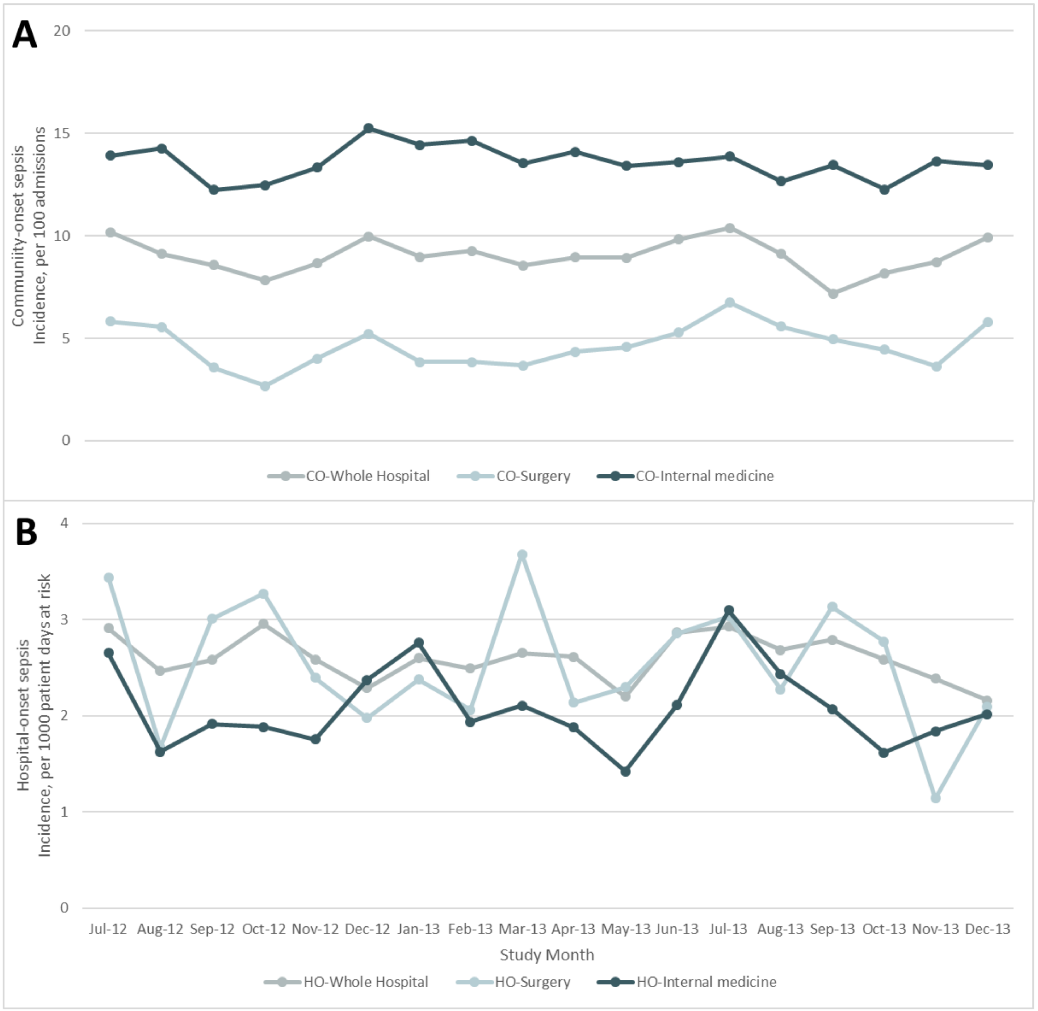


Illustration of changes in community-onset sepsis incidence (A) and hospital-onset sepsis incidence (B) during the study period. Community-onset sepsis incidence is presented per 100 admissions. Hospital-onset sepsis incidence is presented per 1000 patient days at risk. Monitoring curves are displayed for the whole hospital, surgical wards and internal medicine wards. Abbreviations: Community-onset (CO) and hospital-onset (HO).

Supplementary References

1. 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(8):762-774.
2. Centers for Disease Control and Prevention. Hospital Toolkit for Adult Sepsis Surveillance. 2018.https://www.cdc.gov/sepsis/clinicaltools/index.html (accessed 10 Jul 2019).
3. Aakre C, Franco PM, Ferreyra M, *et al.* Prospective validation of a near real-time EHR-integrated automated SOFA score calculator. *Int J Med Inform* 2017;103:1–6.
4. Pandharipande PP, Shintani AK, Hagerman HE, 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. *Crit Care Med*. 2009;37(4):1317-1321.