IMPORTANCE Estimates from claims-based analyses suggest that the incidence of sepsis is increasing and mortality rates from sepsis are decreasing. However, estimates from claims data may lack clinical fidelity and can be affected by changing diagnosis and coding practices over time.

OBJECTIVE To estimate the US national incidence of sepsis and trends using detailed clinical data from the electronic health record (EHR) systems of diverse hospitals.

DESIGN, SETTING, AND POPULATION Retrospective cohort study of adult patients admitted to 409 academic, community, and federal hospitals from 2009-2014.

EXPOSURES Sepsis was identified using clinical indicators of presumed infection and concurrent acute organ dysfunction, adapting Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) criteria for objective and consistent EHR-based surveillance.

MAIN OUTCOMES AND MEASURES Sepsis incidence, outcomes, and trends from 2009-2014 were calculated using regression models and compared with claims-based estimates using International Classification of Diseases, Ninth Revision, Clinical Modification codes for severe sepsis or septic shock. Case-finding criteria were validated against Sepsis-3 criteria using medical record reviews.

RESULTS A total of 173,690 sepsis cases (mean age, 66.5 [SD, 15.5] y; 77,660 [42.4%] women) were identified using clinical criteria among 2,901,019 adults admitted to study hospitals in 2014 (6.0% incidence). Of these, 26,061 (15.0%) died in the hospital and 10,731 (6.2%) were discharged to hospice. From 2009-2014, sepsis incidence using clinical criteria was stable (+0.6% relative change/y [95% CI, −2.3% to 3.5%], P = .67) whereas incidence per claims increased (+10.3%/y [95% CI, 7.2% to 13.3%], P < .001). In-hospital mortality using clinical criteria declined (−3.3%/y [95% CI, −5.6% to −1.0%], P = .004), but there was no significant change in the combined outcome of death or discharge to hospice (−1.3%/y [95% CI, −3.2% to 0.6%], P = .19). In contrast, mortality using claims declined significantly (−7.0%/y [95% CI, −8.8% to −5.2%], P < .001), as did death or discharge to hospice (−4.5%/y [95% CI, −6.1% to −2.8%], P < .001). Clinical criteria were more sensitive in identifying sepsis than claims (69.7% [95% CI, 52.9% to 92.0%] vs 32.3% [95% CI, 24.4% to 43.0%], P < .001), with comparable positive predictive value (70.4% [95% CI, 64.0% to 76.8%] vs 75.2% [95% CI, 69.8% to 80.6%], P = .23).

CONCLUSIONS AND RELEVANCE In clinical data from 409 hospitals, sepsis was present in 6% of adult hospitalizations, and in contrast to claims-based analyses, neither the incidence of sepsis nor the combined outcome of death or discharge to hospice changed significantly between 2009-2014. The findings also suggest that EHR-based clinical data provide more objective estimates than claims-based data for sepsis surveillance.
Sepsis is a major public health problem. It is among the most expensive conditions treated in US hospitals and a leading cause of death.\(^1\)\(^2\) Numerous studies suggest that the incidence of sepsis is increasing over time, offsetting declining case-fatality rates.\(^3\)\(^6\)

Despite its importance, reliably measuring sepsis incidence and trends is challenging. Most studies have used claims data, but increasing clinical awareness, changes in diagnosis and coding practices, and variable definitions have led to uncertainty about the accuracy of reported trends as well as marked heterogeneity in incidence and mortality rates.\(^7\)\(^8\) Analyses in a limited set of hospitals using clinical data have also suggested that sepsis incidence and outcomes may be more stable than previously thought.\(^12\)\(^13\)

The increasing use of electronic health record (EHR) systems allows for the possibility of widespread sepsis surveillance using consistent clinical criteria for concurrent infection and organ dysfunction rather than claims data. In this study, Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) criteria were adapted for public health surveillance and optimized for applicability across different EHR systems. This definition was then applied to EHR data from a diverse set of hospitals to estimate US sepsis incidence and trends from 2009-2014 and to compare with trends estimated from claims data.

## Methods

### Study Design, Data Sources, and Population

This was a retrospective cohort study using EHR and administrative data from 409 academic, community, and federal acute care hospitals gathered from 7 independent data sets: Brigham and Women’s Hospital, Cerner HealthFacts, Emory Healthcare, Hospital Corporation of America, Institute for Health Metrics, University of Pittsburgh Medical Center health care system, and the Veterans Affairs (VA) hospital system (described further in eMethods 1 in the Supplement). These data sets were chosen to include a broad mix of hospitals that resemble all US acute care hospitals in terms of geographic mix, teaching status, and hospital size (eTable 1 in the Supplement).

The study included adults 20 years or older admitted as inpatients or under observation status or who died in the emergency department in calendar years 2009-2014. Fixed categories of race/ethnicity, as reported by patients in the EHR systems of each health care system, were included to characterize the generalizability of the study dataset and because of previously reported associations between race/ethnicity and sepsis incidence and outcomes.\(^3\) The study was approved with a waiver of informed consent by the institutional review boards at Harvard Pilgrim Health Care Institute, Partners HealthCare, University of Pittsburgh, Emory University, and Ann Arbor VA.

### Sepsis Clinical Surveillance Definition

As per Sepsis-3 criteria, sepsis was defined as concurrent infection and organ dysfunction.\(^5\)\(^6\) Suspected infection criteria and the Sequential Organ Failure Assessment (SOFA) score,\(^7\) however, were modified to facilitate widespread retrospective surveillance using routinely collected EHR data (Box). Presumed serious infections were defined as a blood culture draw and sustained administration of new antibiotics. Four or more antibiotic days, including at least 1 intravenous antibiotic, were required to identify those most likely to have serious infections and to eliminate patients treated empirically for 48 to 72 hours before culture results were obtained. Fewer than 4 antibiotic days were allowed if death or discharge to hospice or another acute care hospital occurred before 4 days elapsed.

Sepsis criteria were met if patients had at least 1 concurrent acute organ dysfunction, defined by initiation of vasopressors or mechanical ventilation, elevated lactate level, or significant changes in baseline creatinine level, bilirubin level, or platelet count. The first antibiotic day and organ dysfunction were required to occur within ±2 calendar days of the blood culture draw.

Organ dysfunction thresholds were selected to generally yield an increase in SOFA score of 2 or more points, to parallel Sepsis-3 criteria. The Glasgow Coma Scale score was not included because it was not measured in most patients and was variably assessed across hospitals.\(^18\) Vital signs were also not included because they are not available in all EHRs and are susceptible to transient perturbations and measurement errors. Although not part of the SOFA score, lactate levels of 2.0 mmol/L or greater were included for the analysis of sepsis incidence, outcomes, and clinical characteristics in 2014, given the central role of lactate levels in identifying and risk stratifying sepsis.\(^19\)\(^20\) As discussed below, however, the lactate criterion was excluded from the primary trends analysis because lactate testing rates are rapidly increasing over time and may thus introduce ascertainment bias.\(^21\)

Sepsis was defined as hospital onset (vs present on admission) if infection and organ dysfunction criteria first occurred on or after hospital day 3. Baseline laboratory values for creatinine, bilirubin, and platelets were estimated using the best value during hospitalization for infection present on admission, or best values within ±2 days of the blood culture for hospital-onset infection. The entire hospitalization was considered a single case of sepsis if surveillance criteria were met.
multiple times. Septic shock was defined as presumed serious infection concurrent with vasopressors and serum lactate level 2.0 mmol/L or greater.

Blood culture draws were used to anchor the primary definition because they are an important marker of suspected sepsis and are relatively simple to identify in EHRs. However, a sensitivity analysis was performed using any clinical culture as a broader definition of presumed infection (eAppendix C in the Supplement). These definitions were developed through consensus discussions among the investigative team, informed by prior work using similar approaches.21-23

Implementation
Case-finding code was created in SAS and SQL and distributed to partners to execute locally against their EHR data arrayed according to a common data specification (detailed description provided in eAppendices A-F in the Supplement). Patient comorbidities were derived from International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes using the Charlson method.23

Validation Using Medical Record Reviews
Study physicians reviewed full text medical records from 510 randomly selected hospitalizations, stratified into those that did and did not meet EHR sepsis surveillance criteria. These hospitalizations were drawn from 3 academic centers and 2 community hospitals in Massachusetts and Georgia. Hospitalizations were classified by reviewers as sepsis-positive if there was definite or possible infection and an increase in SOFA score by 2 or more points from baseline as a result of infection, as per Sepsis-3 criteria.15 “Definite infection” required positive cultures or radiography and a compatible clinical syndrome. “Possible infection” required documentation that the medical team presumed the patient’s illness was due to infection, treated for infection, and did not find an alternative etiology. Reviewers were blinded to whether cases were flagged by EHR surveillance criteria and patients’ ICD-9-CM codes (additional details in eMethods 2 in the Supplement).

Trends
Annual trends in sepsis incidence and in-hospital mortality from 2009-2014 were calculated in the subset of hospitals with historical data (at least 84 hospitals per year) (eTable 2 in the Supplement). In light of rising rates of hospice utilization nationwide,24 the combined outcome of in-hospital death or discharge to hospice was also examined to obtain a more complete picture of trends in sepsis outcomes.

An a priori decision was made to exclude the lactate criterion when assessing whether sepsis incidence and outcomes are changing over time. This was done in light of steady increases in lactate testing rates, which might create ascertainment bias and inflate perceived changes in sepsis incidence.21,25 However, a secondary analysis was performed with the lactate criterion to test this hypothesis.

We also applied 2 claims-based definitions: “explicit” ICD-9-CM codes for severe sepsis (995.92) and septic shock (785.52) and an “implicit” method requiring at least 1 infection code and 1 organ dysfunction code (Angus method) or explicit codes for severe sepsis or septic shock.26,27 To examine whether recognition and coding for sepsis is increasing, we calculated the annual proportion of hospitalizations flagged by the primary EHR surveillance definition that also received sepsis codes.

Statistical Analyses
To account for different hospitals contributing data in different years, 2009-2014 trends were modeled using generalized estimating equations to fit Poisson regression models, adjusting for hospital characteristics (institution, region, teaching status, bed count, and annual admissions) and case mix (median age of hospitalized patients, sex and race/ethnicity distributions, and proportion of intensive care unit [ICU] vs total admissions). Generalized estimating equations were used to account for correlations in the data over time as well as hospital-level clustering.
Table 1. Characteristics and Case Mix of Study Hospitals in 2014

<table>
<thead>
<tr>
<th>Hospital Characteristic</th>
<th>Distribution Among Study Hospitals, No. (%) (N = 409)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Geographic region</strong></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>56 (13.7)</td>
</tr>
<tr>
<td>South</td>
<td>205 (50.1)</td>
</tr>
<tr>
<td>Midwest</td>
<td>60 (14.7)</td>
</tr>
<tr>
<td>West</td>
<td>88 (21.5)</td>
</tr>
<tr>
<td><strong>Teaching status</strong></td>
<td></td>
</tr>
<tr>
<td>Teaching</td>
<td>152 (37.2)</td>
</tr>
<tr>
<td>Nonteaching</td>
<td>257 (62.8)</td>
</tr>
<tr>
<td><strong>AHA hospital size</strong></td>
<td></td>
</tr>
<tr>
<td>Small (&lt;200 beds)</td>
<td>220 (53.8)</td>
</tr>
<tr>
<td>Medium (200–499 beds)</td>
<td>155 (37.9)</td>
</tr>
<tr>
<td>Large (≥500 beds)</td>
<td>34 (8.3)</td>
</tr>
<tr>
<td><strong>Annual admissions</strong></td>
<td></td>
</tr>
<tr>
<td>0–5000</td>
<td>127 (45.4)</td>
</tr>
<tr>
<td>5001–10 000</td>
<td>53 (18.9)</td>
</tr>
<tr>
<td>10 001–20 000</td>
<td>76 (27.1)</td>
</tr>
<tr>
<td>&gt;20 000</td>
<td>24 (8.6)</td>
</tr>
</tbody>
</table>
| **Characteristics of hospitalized patients, median (IQR)**
| Age, y                  | 62 (58-66)                                            |
| Women, %                | 41.3 (37.7-44.9)                                      |
| Race/ethnicity, %       |                                                       |
| White                   | 80.5 (62.9-90.8)                                      |
| Black                   | 6.6 (1.4-14.9)                                        |
| Hispanic                | 1.4 (0.1-8.6)                                         |
| Asian                   | 0.6 (0.3-1.5)                                         |
| Other                   | 1.7 (0.6-4.3)                                         |
| Median proportion of ICU vs total admissions, % | 12.4 (8.0-16.7)                                      |

Abbreviations: AHA, American Hospital Association; ICU, intensive care unit; IQR, interquartile range.

* Data on annual number of admissions were not available for Veterans Affairs (VA) hospitals (n = 129); thus, the denominator for this category includes 280 hospitals.

* The data shown represent the median value of hospital-level medians among all 409 hospitals, along with the IQRs of these medians. Individual hospital-level data were unavailable for VA hospitals; thus, VA hospitals were considered a single institution, and median values in the entire VA dataset were calculated.

Adjusted rates for 2009-2013 were generated by creating binary indicators for each year in the model, with 2014 as the reference year. The overall change from 2009 to 2014 was summarized as the exponentiated slope of the line between these 2 years on the log scale. All percentages were presented as relative annual changes. VA hospitals were excluded from trends analyses because only data from 2014 were available.

The national weighted incidence and mortality of sepsis among hospitalized adults in 2014 was estimated by projecting study hospital case counts into stratifications of US hospitals by region, size, and teaching status (eMethods 3 in the Supplement).

Missing laboratory data for organ dysfunction calculations were assumed to be normal. Multiple imputation was not performed because missing data were presumed to be missing owing to clinical decisions not to order certain tests rather than missing at random. Data completeness was assessed by inspecting each data element required for the EHR sepsis definition on a hospital-by-hospital-level basis and aggregate rates of organ dysfunction and sepsis incidence and outcomes across each dataset (eFigures 1 and 2 in the Supplement).

Continuous variables were expressed as means. Non-normally distributed variables were expressed as medians of the medians among the 7 datasets. Analyses were conducted using SAS version 9.3 (SAS Institute) and R version 3.3.1 (http://www.r-project.org). For all analyses, P < .05 (2-sided) was considered statistically significant.

Results

**Sepsis Incidence, Clinical Characteristics, and Outcomes in 2014**

The 2014 study cohort included 2,901,019 adult encounters in 409 hospitals, representing approximately 10% of all US adult hospitalizations. Study hospitals’ characteristics are reported in Table 1. There were 423,758 patients with presumed serious infection (14.6% incidence [95% CI, 14.6% to 14.7%], of which 32,574 (7.7% [95% CI, 7.6% to 7.8%]) died in the hospital. There were 173,690 patients with sepsis (overall hospital incidence, 6.0% [95% CI, 6.0% to 6.0%]). The clinical characteristics of patients with sepsis are reported in Table 2. Mean age was 66.5 years (SD, 15.5), and 42.4% (95% CI, 42.2% to 42.6%) were women. Comorbidities were common, including diabetes (35.7% [95% CI, 35.5% to 36.0%]), pulmonary disease (30.9% [95% CI, 30.7% to 31.2%]), renal disease (26.8% [95% CI, 26.7% to 27.0%]), and cancer (19.7% [95% CI, 19.5% to 19.9%]). Most sepsis cases (86.8% [95% CI, 86.7% to 87.0%]) were present on admission. Among hospitals (n = 280) for which culture results were available, 17.2% (95% CI, 17.0% to 17.4%) of patients with sepsis had positive blood cultures.

Of the 173,690 patients with sepsis in study hospitals in 2014, 94,956 (54.7%) required ICU care during hospitalization, 27,502 (15.8%) had septic shock, 26,061 (15.0%) died in the hospital, and 10,731 (6.2%) were discharged to hospice (Table 2). Median ICU length of stay was 5 days (range, 2–6). Median hospital length of stay was 10 days (range, 8–12). Hospital mortality was 25.5% among patients with hospital-onset sepsis vs 13.4% for patients with sepsis present on admission (difference, 12.1% [95% CI, 11.5% to 12.7%]; P < .001). Sepsis was present during hospitalization in 34.7% of the 75,079 study patients who died in the hospital in 2014. When adjusting for hospital region, size, and teaching status, the estimated national weighted incidence of sepsis was 5.9% (95% CI, 5.5% to 6.3%), and the in-hospital mortality rate was 15.6% (95% CI, 14.8% to 16.5%) (strata-specific counts reported in eTable 3 in the Supplement). Mortality rates for septic patients were higher for older patients, men, teaching hospitals, and larger hospitals (eTable 4 in the Supplement).

The distribution of organ dysfunctions, and their associated mortality rates, are shown in Figure 1. Approximately 50% of patients with sepsis met at least 2 criteria for acute organ dysfunction.
Validation
On medical record reviews, EHR surveillance criteria had 69.7% sensitivity (95% CI, 52.9% to 92.0%), 98.1% specificity (95% CI, 97.7% to 98.5%), 70.4% positive predictive value (PPV) (95% CI, 64.0% to 76.8%), and 98.0% negative predictive value (95% CI, 95.9% to 99.6%), relative to Sepsis-3 criteria. Explicit sepsis codes had lower sensitivity (32.3% [95% CI, 24.4% to 43.0%] vs 69.7% [95% CI, 52.9% to 92.0%] for EHR criteria, P <.001) but with comparable PPV (75.2% [95% CI, 69.8% to 80.6%] vs 70.4% [95% CI, 64.0% to 76.8%] for EHR criteria, P = .23). The combination of explicit or implicit codes had comparable sensitivity compared with EHR criteria (66.0% [95% CI, 51.4% to 80.7%], P = .23) but lower PPV (31.0% [95% CI, 24.9% to 40.4%], P <.001) (eTable 5 in the Supplement).

Among the 13 Sepsis-3 cases missed by EHR criteria, the most common reason was hypoxemia causing an increase in SOFA score of 2 or more points without need for mechanical ventilation (Table 3 in the Supplement). None of these 13 patients died. There were 57 false-positives flagged by EHR criteria; most were because clinicians initially suspected infection but medical record reviewers ultimately deemed no infection was present or because patients’ organ dysfunction did not increase SOFA score by 2 or more points (typically patients with elevated lactate levels alone) (eTable 7 in the Supplement). If sepsis was defined as clinically suspected infection with organ dysfunction at the thresholds specified by EHR criteria, PPV was 87.9% (95% CI, 82.7% to 92.0%).

Trends From 2009-2014
The trends analysis included 7,801,624 hospitalizations between 2009-2014 (Figure 2). The annual incidence of sepsis (without the lactate criterion) was stable (+0.6% relative change/y [95% CI, −2.3% to 3.5%], P = .67). In-hospital mortality decreased (−3.3%/y [95% CI, −5.6% to −1.0%], P = .004), but there was no significant change in the combined outcome of death or discharge to hospice (−1.3%/y [95% CI, −3.2% to 0.6%], P = .19). Patients with sepsis were increasingly discharged to hospice over time (+6.3%/y [95% CI, 1.1% to 11.6%], P = .02).

When including lactate as a criterion, sepsis incidence increased (+3.5%/y [95% CI, 0.7% to 6.4%], P = .02), mortality decreased (−5.0%/y [95% CI, −7.3% to −2.7%], P <.001), and death or discharge to hospice decreased (−2.0%/y [95% CI, −3.8% to −0.2%], P = .03).

Sepsis incidence using explicit severe sepsis/septic shock codes increased significantly (+10.3%/y [95% CI, 7.2% to 13.3%], P <.001), as did sepsis/septic shock defined using explicit codes or implicit codes (+7.3%/y [95% CI, 5.0% to 9.5%], P <.001). Mortality declined using explicit codes (−7.0%/y [95% CI, −8.8% to −5.2%], P <.001) and explicit or implicit codes combined (−6.6%/y [95% CI, −8.3% to −4.8%], P <.001). The combined outcome of death or discharge to hospice also decreased with explicit codes (−4.5% [95% CI, −6.1% to −2.8%], P <.001) and with explicit or implicit codes combined (−3.6% [95% CI, −4.9% to −2.3%], P <.001).

Among patients meeting EHR clinical criteria for sepsis, the proportion who received explicit sepsis codes increased from 24.9% in 2009 to 30.5% in 2014 (difference, 5.6% [95% CI, 4.9% to 6.3%], P < .001). The reported median organ dysfunctions, ICU length of stay, and hospital length of stay represent median of the median values in each of the 7 datasets, along with range of medians across the datasets.
Sensitivity Analyses

Sensitivity analyses restricted to hospitals that reported data each year from 2009-2014 yielded similar differences in EHR-based vs claims-based trends (eFigure 3 in the Supplement). When any clinical culture was used to indicate presumed infection rather than blood cultures alone, sepsis incidence rates in 2014 were higher (7.1% vs 6.0%; difference, 1.1% [95% CI, 0.8% to 1.2%; P < .001] and mortality rates lower (14.0% vs 15.0%; difference, 1.0% [95% CI, 0.8% to 1.2%]; P < .001). Clinical characteristics and trends were similar (eTables 8-9 in the Supplement).

Discussion

In this retrospective analysis of more than 2.9 million adults admitted to 409 US hospitals in 2014, clinical indicators of sepsis were present in 6% of hospitalized patients, of whom 21% died in the hospital or were discharged to hospice. Sepsis was present in 35% of all hospitalizations that culminated in death. In contrast to claims-based analyses, sepsis incidence rates using clinical data were stable from 2009-2014; in-hospital mortality rates declined, but there was no significant change in the combined outcome of death or discharge to hospice.

Reliable sepsis surveillance is essential given its high burden, the proliferation of treatment and prevention initiatives, and new national sepsis quality measures. Identifying sepsis using consistent clinical criteria through EHR data, rather than relying on explicit clinical diagnoses or hospital coders, enhances confidence in sepsis estimates because clinicians underrecognize sepsis and vary widely in their knowledge and application of sepsis definitions.28-30 Hospitals also vary significantly in how they assign codes for infection and organ dysfunction, and the presence of both of these codes at discharge does not guarantee that they occurred concurrently.31-33 EHR-based criteria were more sensitive than explicit sepsis codes on medical record review, with comparable PPV; EHR-based criteria had similar sensitivity to implicit or explicit codes combined but higher PPV.

EHR-based clinical surveillance also provides more credible estimates of sepsis trends compared with claims, which can be biased by changing diagnosis and coding practices over time.9,10,13,34 Among patients with sepsis identified using EHR clinical criteria, there was an increase over time in the proportion assigned explicit and implicit sepsis codes, presumably reflecting ongoing efforts to improve sepsis awareness, documentation, and coding. Improving sepsis recognition, including less severe cases, likely accounts for the difference in clinical vs claims-based incidence trends as well as the greater mortality decline seen with claims. The apparent improvement in sepsis-associated mortality was also nullified when also considering discharge to hospice, underscoring the need to consider temporal changes in end-of-life care patterns when assessing trends in clinical outcomes.24

When including elevated lactate levels in the surveillance definition, mild increases in sepsis incidence and decreases in mortality or discharge to hospice were observed. This likely reflects increasingly aggressive lactate testing over time and identification of progressively less ill patients with sepsis, paralleling the increase in sepsis awareness and coding.21,25 Thus, although lactate surveillance can be useful for identifying sepsis
cases at a single point, it risks generating misleading impressions of changes in sepsis incidence and outcomes over time.

Ongoing improvements in sepsis coding suggest that claims may eventually reach the accuracy of EHR data. However, the variability of claims between hospitals and their susceptibility to shifting policy and reimbursement incentives limits their use to compare hospitals or assess trends—limitations particularly important in the new era of sepsis quality measures and regulations. Surveillance experience from other domains, such as health care–associated infections, speak to the risk of reimbursement policies changing coding practices and hence perceived rates that do not reflect true changes in infection rates.

The national weighted sepsis incidence of 5.9% among hospitalized patients and in-hospital mortality of 15.6% estimated in this study would translate into approximately 1.7 million US adult sepsis hospitalizations and 270,000 deaths in 2014. This falls within the wide incidence range of 900,000 to 3.1 million previously estimated using 4 different claims-based definitions. The observed mortality rate exceeds the 10% mortality reported in the Sepsis-3 derivation and validation studies. This may reflect the more stringent definition of presumed infection (requiring blood cultures and ≥4 days of antibiotics rather than a single dose) and SOFA score adaptations.

**Strengths and Limitations**

Strengths of the current study include the use of detailed clinical data from a large number of diverse hospitals that together account for approximately 10% of all US acute care hospitalizations in 2014. The data sets came from unrelated hospital networks, limiting the possibility of bias from idiosyncratic clinical, diagnosis, or coding patterns. A sensitivity analysis using a broader definition of presumed infection demonstrated relatively little change in sepsis incidence and outcomes and no difference in trends, supporting the robustness of the primary definition.

This study also has several limitations. First, EHR-based surveillance may be affected by differences in clinical practice between clinicians and hospitals as well as changes over time. This may be most evident with rising rates of lactate testing, but earlier initiation of vasopressors for hypotension and increasing use of noninvasive ventilation and high-flow nasal cannulas for respiratory failure could also affect estimates. The partial dependence on clinician-initiated interventions to measure organ dysfunction, however, is also a limitation of the SOFA score and thus the Sepsis-3 criteria.

Second, only adults in US acute care hospitals were included; future studies should address sepsis in neonates and children, as well as consider ways to identify sepsis outside hospitals and in countries without widespread EHR systems.

Third, not all hospitals contributed data each year for trends analyses. However, regression models were used to adjust for hospital-level differences.

Fourth, the study did not directly ascertain sepsis mortality that occurs after hospital discharge.

Fifth, medical record reviews were only conducted on a fraction of the study cohort. Cases were drawn from both academic and community hospitals, but they still may not be representative of all hospitals.

---

**Figure 2. Sepsis Trends From 2009-2014: Incidence, In-hospital Sepsis Mortality, and In-hospital Mortality or Discharge to Hospice.**

Adjusted rates from 2009-2013 calculated relative to observed 2014 rates. Error bars indicate 95% CIs. "Clinical criteria" indicates blood cultures + antibiotics + concurrent organ dysfunction (Box). "Clinical criteria without lactate" excludes the criterion for lactate level of 2.0 mmol/L or greater. Primary trends assessment was conducted using clinical criteria without lactate levels. "Explicit sepsis codes": discharge diagnoses of severe sepsis (995.92) or septic shock (785.52). "Implicit sepsis codes": at least 1 infection diagnosis and 1 organ dysfunction diagnosis. All trends adjusted for hospital characteristics (institution, region, teaching status, bed count, annual admissions) and case mix (median age of hospitalized patients, sex and race/ethnicity distributions, proportion of intensive care unit vs total admissions). Veterans Affairs hospitals not included in the trends analysis. Total number of sepsis cases per year was 30,744 (2009), 35,596 (2010), 34,445 (2011), 36,524 (2012), 144,322 (2013), and 145,236 (2014) for clinical criteria; 28,723 (2009), 32,175 (2010), 30,348 (2011), 32,019 (2012), 120,402 (2013), and 112,355 (2014) for clinical criteria without lactate level; 50,223 (2009), 61,483 (2010), 65,193 (2011), 76,208 (2012), 275,480 (2013), and 272,679 (2014) for implicit or explicit codes; and 9062 (2009), 12,688 (2010), 12,571 (2011), 15,309 (2012), 61,285 (2013), and 65,176 (2014) for explicit codes.
Sixth, the study dataset might not be representative of national data. However, study hospitals had characteristics similar to those of US hospitals overall and represented a substantial fraction of total admissions nationwide.

Seventh, medical record reviews suggest that EHR-based surveillance may miss up to 30% of patients with sepsis while misclassifying another 30%. Quantifying the accuracy of sepsis criteria is elusive, however, because there is no true gold standard for sepsis. Recognizing this problem, the Sepsis-3 task force developed and validated criteria based on associations with adverse outcomes. The EHR surveillance definition in this study carried high mortality rates (15%) compared with all encounters with presumed infection (8%), and mortality increased with increasing numbers of dysfunctional organs. On medical record reviews, the Sepsis-3 cases missed by EHR surveillance involved mild organ dysfunction, such as hypoxemia without need for mechanical ventilation, and no patients in that group died. False-positives were most often attributable to reviewers adjudicating the absence of infection despite patients receiving blood culture draws and antibiotics, or a noninfectious cause of organ dysfunction. When sepsis was defined as organ dysfunction concurrent with clinically suspected infection (as is common in practice), PPV of the surveillance definition increased to 88%.

Eighth, neither Sepsis-3 criteria nor EHR-based clinical surveillance can solve the challenge that clinicians routinely face in deciding whether their patient is infected and whether organ dysfunction is due to infection. Instead, EHR surveillance provides a consistent gauge to estimate sepsis incidence and outcomes using readily available, objective clinical data. This method cannot help clinicians identify sepsis at the bedside since it is retrospective, but it may be useful for public health surveillance, hospital evaluation, and assessing the effects of quality improvement efforts. EHR-based surveillance may further support these objectives by facilitating granular evaluation of the timing of sepsis onset and interventions.

Conclusions

In clinical data from 409 hospitals, sepsis was present in 6% of adult hospitalizations, and in contrast to claims-based analyses, neither the incidence of sepsis nor the combined outcome of death or discharge to hospice changed significantly between 2009-2014. The findings also suggest that EHR-based clinical data provide more objective estimates than claims-based data for sepsis surveillance.
the following study team members, who did receive compensation, for their assistance in obtaining and analyzing data needed for the study: Brenda Vincent, MS, Wyndy Wiltzala, PhD, Vanessa Dickerman, PhD, and Jennifer A. Davis, MHSA (Ann Arbor VA); Jonathan Holton, BS, and Edvin Music, MBA, MSIS (University of Pittsburgh Medical Center); Elizabeth Overton, MS (Emory Healthcare); Richard Schaaf, BS, David Fram, BA, and Karen Eberhardt, BS (Commonwealth Informatics); and Robert Jin, MS (Harvard Medical School/Harvard Pilgrim Health Care Institute).

REFERENCES


