HHS Public Access

Author manuscript

Crit Care Med. Author manuscript; available in PMC 2019 August 01.

Published in final edited form as:

Crit Care Med. 2018 August ; 46(8): 1247–1254. doi:10.1097/CCM.0000000000003184.

Mortality measures to profile hospital performance for patients with septic shock

Allan J. Walkey, MD, MSc[1], **Meng-Shiou Shieh, PhD**[2], **Vincent X. Liu, MD, MS**[3], and **Peter K. Lindenauer, MD, MSc**[2]

[1]Department of Medicine, Division of Pulmonary, Allergy and Critical Care, Center for Implementation and Improvement Sciences, Boston University School of Medicine

[2]Institute for Healthcare Delivery and Population Science and Department of Medicine, University of Massachusetts Medical School – Baystate, Springfield MA, and Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA USA

[3]Kaiser Permanente Division of Research; Oakland CA

Abstract

Objective—Sepsis care is becoming a more common target for hospital performance measurement, but few studies have evaluated the acceptability of sepsis or septic shock mortality as a potential performance measure. In the absence of a gold standard to identify septic shock in claims data, we assessed agreement and stability of hospital mortality performance under different case definitions.

Design—Retrospective cohort study.

Setting—United States acute care hospitals.

Patients—Hospitalized with septic shock on admission, identified by either implicit diagnosis criteria (charges for antibiotics, cultures and vasopressors), or by explicit *International Classification of Diseases, 9th Revision* (*ICD-9)* codes.

Interventions—None

Measurements and Main Results—We used hierarchical logistic regression models to determine hospital risk-standardized mortality rates and hospital performance outliers. We assessed agreement in hospital mortality rankings when septic shock cases were identified by either explicit *ICD-9* codes or implicit diagnosis criteria. Kappa statistics and intra-class correlation coefficients (ICC) were used to assess agreement in hospital risk-standardized

Corresponding Author: Allan J. Walkey, MD, MSc; Boston University School of Medicine, 72 East Concord Street, R-304; Boston, MA 02118, Phone: (617) 638-4860, Fax: (617) 638-4896, alwalkey@bu.edu.

Conception and design: AJW, PKL
Data Acquisition and Analysis: MSS, AJW, PKL
Interpretation of data for the work: AJW, PKL, MSS, VXL
Drafting the work and revising for important intellectual content: AJW, PKL, MSS, VXL

Conflicts of Interest: None.
Reprints will not be ordered.

Dr. Shieh has disclosed that he does not have any potential conflicts of interest.

mortality and hospital outlier status, respectively. 56,673 patients in 308 hospitals fulfilled at least one case definition for septic shock, while 19,136 (33.8%) met both the explicit *ICD-9* and implicit septic shock definition. Hospitals varied widely in risk-standardized septic shock mortality (interquartile range of implicit diagnosis mortality: 25.4-33.5%; ICD-9 diagnosis: 30.2-38.0%). The median absolute difference in hospital ranking between septic shock cohorts defined by *ICD-9* vs. implicit criteria was 37 places (IQR16-70), with an ICC of 0.72, p<0.001; agreement between case definitions for identification of outlier hospitals was moderate [kappa 0.44, (95% CI 0.30-0.58)].

Conclusions—Risk-standardized septic shock mortality rates varied considerably between hospitals, suggesting that septic shock is an important performance target. However, efforts to profile hospital performance were sensitive to septic shock case definitions, suggesting that septic shock mortality is not currently ready for widespread use as a hospital quality measure.

MeSH keywords

Sepsis; Health Services Research; Outcome assessment

Sepsis is the most common reason for non-elective hospitalization and hospital readmissions in the United States,(1, 2) with short-term mortality rates of approximately 20-30%(3) and significant long-term morbidity and mortality.(4) The major public heath burdens of sepsis, paired with large unexplained variation in practice patterns,(5–7) make sepsis a logical target for hospital performance profiling and quality improvement. In response, a sepsis performance measure (SEP-1) was recently introduced by the Centers of Medicare and Medicaid Services (CMS).(8) CMS SEP-1 measures multiple processes of care across the first 6 hours of sepsis management.(9) However, the high costs,(10) large administrative burdens, and poor quality evidence linking individual SEP-1 process measures to improved patient outcomes(3, 11) have elicited substantial controversy.(12)

In contrast to the resource-intensive and evidence-dependent nature of process performance measures, outcome measures represent a more patient-centered and efficient method of measuring hospital performance.(13, 14) Mortality measures allow hospitals to tailor quality approaches to local contexts in order to achieve outcome improvements. However, outcome-based performance measures face substantial challenges that include the need for risk adjustment,(15–18) and sensitivity to variation in documentation and coding between hospitals.(19, 20) In the case of sepsis and septic shock, the lack of a gold standard definition(21, 22) coupled with changes in Consensus definitions(23) and hospital variation in patterns of claims coding(24) have the potential to substantially affect risk-adjusted outcomes such as mortality rates.

Multiple conditions are the subject of hospital outcome performance assessments, such as validated CMS measures for myocardial infarction, heart failure and pneumonia mortality.(8) Prior to approval by organizations such as the National Quality Forum, potential performance measures must demonstrate clinically meaningful opportunities for performance improvement and establish their “scientific acceptability”.(25) While sepsis is of substantial clinical importance, with evidence of performance gaps between hospitals for sepsis risk-adjusted mortality seen at the state level,(26) little evidence supports the

scientific acceptability of traditional methods used for comparing outcomes between hospitals for sepsis as a national outcome measure. Because an outcome measure for hospital sepsis management could overcome limitations of SEP-1, we sought to evaluate the effect of using different definitions of septic shock upon hospital mortality performance assessments.

Methods

Data source

We performed a retrospective cohort study of hospital mortality rates using de-identified, enhanced administrated claims data from Premier, Inc. using hospitalizations during the year 2014. Premier data include fields available in traditional hospital administrative claims such as patient demographics, *ICD-9* codes for diagnoses and procedures, and patient hospital discharge status (ie., hospital mortality), as well as detailed, time-stamped information on all charges during the hospitalization, including medications, laboratory, and imaging orders. Premier data uniquely allowed evaluation and comparison of patient cohorts defined using traditional claims as well as using cohorts defined using medications and laboratory orders. Hospitals voluntarily submit data to Premier for benchmarking and quality assessment and Premier data represent a non-random, approximate 20% sample of all hospitalizations in the United States.

Septic Shock Cohorts

We chose septic shock as the condition of interest for a hospital outcome measure because of its high mortality rate and the ability to detect septic shock using two different methods, both of which have been previously validated against chart data and/or and Consensus criteria.(23) The National Quality Forum defines a quality measure as “valid” if it agrees with another authoritative source of quality measurement.(25) Because no gold standard definition of septic shock exists,(21) we evaluated the acceptability of a septic shock outcome measure by assessing the agreement between two different definitions of “septic shock”: 1) A cohort with an implicit diagnosis of septic shock on admission (based upon any duration of charges for intravenous antibiotics, blood cultures and vasopressors during the first two days of hospitalization, previously validated against medical chart criteria with sensitivity of 88% and specificity of 92%);(27) and 2) an explicit *ICD-9* code-based definition defined by a diagnosis for septicemia (*ICD-9* 038) present on admission along with *ICD-9* 785.52 for septic shock(23) [i.e., explicit “*ICD-9* definition”, an approach with a lower sensitivity (48%) and higher specificity 99%].(27)

Patients were excluded if they were discharged alive within 48 hours of admission, were transferred from another hospital, left against medical advice, had unknown gender or vital status, or were admitted to a hospital with fewer than 25 cases of septic shock, similar to CMS methods for other conditions.(28) An index hospitalization was chosen randomly for patients with more than one hospitalization.(28)

Statistical Analysis

We used multivariable-adjusted hierarchical logistic regression models (SAS proc GLIMMIX) to calculate predicted-to-expected mortality ratios for each hospital as a measure of risk-standardized hospital mortality rates (RSMR).(29, 30) All models included patient age, sex, 29 Elixhauser/Gagne comorbid conditions present on hospital admission, (31–33) and sepsis-associated acute organ failures present on admission (*ICD-9* codes for renal, respiratory, hepatic, metabolic, and neurological failure)(34, 35) (eTable 2 in Supplemental Digital Content), with a random intercept calculated for each hospital to account for correlated outcomes within each hospital.(34, 35) RSMRs were calculated as previously described(8, 29, 30) from the ratio of the number of predicted deaths to the number of expected deaths at each hospital, multiplied by the average hospital mortality rate in the sample, with confidence intervals calculated through 500-fold bootstrapping. We described hospital variation in mortality through 1) interquartile range of RSMR for each cohort, and 2) calculation of median odds ratios (MOR) of mortality based upon the hospital to which a patient was admitted.(36) The MOR is the median odds of mortality for similar patients who were admitted to two randomly selected hospitals from the sample, and provides an estimate of the median increased odds of mortality that would occur if a patient moved from a lower-risk hospital to a higher-risk hospital. Variation in hospital outcomes was shown using caterpillar plots of the RSMR for hospitals ranked in order of RSMR.

We used three different approaches to compare hospital RSMRs generated from each septic shock cohort definition. First, we calculated the absolute difference in hospital ranking between cohorts and reported the median and interquartile range of the differences in hospital ranking. Second, we calculated intra-class correlation coefficients (ICC class 3,1 per Shrout-Fleiss nomenclature), a measure of inter-rater reliability for continuous variables, to assess agreement between hospital RSMRs generated from each cohort.(18, 37) ICCs were similar to Pearson correlation coefficients in all analyses. Third, we reported the agreement between statistically significant hospital outliers identified between cohorts. An outlier hospital was defined by the presence of a RSMR 95% confidence interval that did not include the average cohort mortality rate.(28) Because power to detect statistically significant outliers differs between cohorts based upon sample size, we also divided hospitals into quartiles of RSMR and used modified kappa statistics to evaluate agreement between quartiles of RSMR between cohorts.(38) As per convention, we *a priori* defined ICC and kappa statistics >0.8 as “strong agreement”. (18, 37, 38)

Sensitivity Analyses

Sepsis can be triggered by different sources of infection with different risks of mortality; thus, we performed a sensitivity analysis adjusting for source of infection in the primary hierarchical logistic regression models. Infections present on admission were classified as previously described using *ICD-9* codes for pneumonia, urinary tract infection, skin and soft tissue infection, ischemic bowel, and other intra-abdominal infections.(39) We performed an additional sensitivity analysis using an implicit sepsis cohort defined by any duration of blood cultures, antibiotics and vasopressors during the first two days with a less sensitive (75%), but more specific (97%) implicit septic shock cohort defined using more narrow criteria for antibiotics (at least 4 consecutive days) and vasopressors (at least 2 consecutive

days, unless patient died prior to the consecutive days criteria).(27) Finally, in order to evaluate agreement between two different methods of case identification using ICD-9 codes, we compared hospital rankings derived from a cohort identified with a broad ICD-9 code approach to identifying cases of sepsis (ICD-9 0.38x present on admission)(34) to the subset of patients with an ICD-9 code for septic shock (ICD-9 785.52).

SAS version 9.4 (Cary, NC, USA) was used for all statistical analyses. Study procedures were deemed not to be human subjects research by the Institutional Review Board of Baystate Medical Center.

Results

Comparison of cohorts

Among 6.4 million hospitalizations during the year 2014, 56,673 patients in 308 hospitals fulfilled inclusion criteria and at least one of the primary cohort definitions for septic shock, and 19136 (33.8%) met both the explicit *ICD-9* and implicit antibiotics/vasopressor/culture septic shock definitions (Figure 1). Characteristics of patients included in the *ICD-9* and implicit septic shock cohorts are shown in eTable 2 in Supplemental Digital Content. Multivariable models showed similar associations between patient characteristics and hospital mortality, regardless of cohort definition (eTable 3 in Supplemental Digital Content).

Hospital mortality rates for septic shock

The median unadjusted hospital mortality rate for the implicit septic shock cohort was 27.8% (IQR 23.9-33.7), with a median RSMR of 28.8% (25.4-33.5), whereas median unadjusted mortality and RSMR for the *ICD-9* septic shock cohort were 32.9 (27.4-39.0) and 34.3 (30.2-38.0), respectively. The median odds ratio between hospitals for mortality was 1.34 (95% CI 1.30-1.37) for the implicit diagnosis cohort and was also 1.34 (1.31-1.38) for the *ICD-9* cohort. Variation of RSMRs for septic shock across hospitals using the implicit septic shock case definition criteria is demonstrated in Figure 2.

The median absolute difference in hospital ranking between the *ICD-9* and implicit diagnosis septic shock cohorts was 37 (IQR16-70), with an ICC of 0.72, $p<0.001$ (Figure 3), signifying moderate agreement. There were 23 (7.5%) low performing outliers and 24 (7.8%) high-performing outliers identified in the implicit cohort, whereas there were 13 (4.2%) low-performing outliers and 20 (6.5%) high performing outliers in the *ICD 9* cohort (Table 1). There was moderate agreement in identifying hospitals as outliers between the ICD-9 and implicit septic shock cohorts (kappa 0.44, 95% CI 0.30-0.58), with 6 of 13 (46%) high performing outliers and 8 of 20 (40%) low-performing outliers identified by *ICD-9* criteria not characterized as outliers using implicit criteria. Figure 2 demonstrates the location of rankings for each outlier hospital identified using *ICD-9* criteria on the x-axis of the caterpillar plot for hospital rankings and outlier hospitals as defined by the implicit criteria. Agreement between cohorts for RSMR quartiles was similar to agreement for outliers (kappa 0.53, 95% CI 0.47-0.60, eFigure 1 in Supplemental Digital Content), with

39% and 32% of hospitals differing in classification within the top or bottom quartiles based upon use of *ICD-9* or implicit septic shock criteria, respectively.

We evaluated agreement between the implicit septic shock cohorts defined by charges for any duration of antibiotics and vasopressors compared with a narrower cohort requiring at least 4 days of antibiotics and 2 days of vasopressors (and blood culture). Although adjusted hospital mortality in the cohort requiring consecutive days of antibiotics and vasopressors was expectedly higher (median 49.0%, IQR 45.4-52.9%), the median odds ratio for hospital variation was similar (1.38, 95% CI 1.34-1.43) to other sepsis cohorts, as was ICC ($\rho = 0.76$, $p<0.001$) and kappa statistic of (0.52, (95% CI 0.46-0.59) see Table 2) between the two cohorts identifying septic shock using implicit diagnosis criteria. Sensitivity analyses comparing ICD-9-based definitions of identifying sepsis and septic shock (Table 2) and the two implicit septic shock cohorts after adjustment for infection source [kappa statistic of 0.53 (0.47-0.60) and ICC of 0.76] did not substantively change the results.

Discussion

Much of the growth in measuring hospital performance of the past decade has involved the use of validated outcome measures based upon cohorts defined using *ICD-9* codes. Given the large public health burden and the complexity of current process quality measures focused on sepsis, developing a measure of hospital risk-standardized sepsis mortality would seem to be a logical choice for an outcome measure. However, the extent of variation in sepsis mortality between hospitals has been unclear and the validity of potential sepsis outcome measures has been underexplored. Across multiple methods of defining cases hospitalized with septic shock, we identified wide variation in hospital mortality, indicating that sepsis represents an important outcome measure and target for performance improvement. However, hospital performance rankings for septic shock mortality were not robust to use of different algorithms to define “septic shock”. Because the same hospitals were ranked differently based upon the manner in which septic shock was defined, our findings call into question whether a sepsis mortality measure is ready for implementation at this time.

Compared with established hospital outcome performance measures such as myocardial infarction and pneumonia, septic shock had larger between-hospital variation in performance. For example, reported interquartile ranges of RSMR for myocardial infarction (14.3-15.8%) and pneumonia (10.3-12.6%), were smaller than interquartile ranges of RSMR for septic shock (27.4-39.0%) identified in the present study,(40–43) even after accounting for higher median baseline mortality rates. Median odds ratio results further show that, depending upon hospital of admission, similar patients with septic shock would have a 30-40% increased odds of death when presenting to a low vs. high performing hospital. The wide between-hospital variation in risk-standardized mortality for septic shock – likely greater than current hospital mortality performance measures such as pneumonia and myocardial infarction – suggests opportunities to improve sepsis outcomes at lower performing hospitals.

However, the substantial changes in an individual hospital's ranking based upon the current method used to identify septic shock casts doubt on the acceptability of septic shock mortality performance measures. For example, studies of performance measures for myocardial infarction and pneumonia demonstrated substantially higher correlation (ρ>0.90) (40, 44, 45) between hospital rankings based upon chart-abstraction and *ICD-9* codes as compared with methods used to identify septic shock in the present study (ICC ρ=0.72). Further, only 33.8% of patients met both the explicit *ICD-9* and implicit antibiotics, vasopressor, and culture septic shock definitions; and more than 1 in 3 hospitals classified as outliers differed based upon the method used to identify septic shock cases. Thus, although septic shock mortality is an important target for quality improvement, the lack of adequate agreement in hospital rankings between different published methods of identifying septic shock suggests that more research is required to develop consistent sepsis case definitions that can generate reliable hospital mortality performance measurement.

The difficulties presented by septic shock as an outcome performance measure are likely due, in part, to the effects of variation in sepsis recognition. Prior studies have demonstrated only moderate agreement between clinicians in recognizing patients with sepsis (kappa 0.7-0.8)(27, 46) and wide variation between institutions in *ICD-9* coding patterns for sepsis and infection.(19, 20) While hospital variation in *ICD-9* codes for identification of septic shock may seem to argue for use of more objective implicit diagnoses abstracted from hospital charge data, hospital performance rankings were not robust to small changes in implicit criteria that varied duration of antibiotics and vasopressors used to define septic shock.

At least two potentially feasible avenues exist to develop future sepsis mortality performance measures. The first approach could involve extraction of highly granular electronic medical record data to develop and standardize sepsis identification for high dimensional risk-adjustment. Such an approach is not currently in routine use for other medical conditions and requires a greater harmonization across electronic medical records than currently exists. (24, 47) A second approach would acknowledge the poor agreement across case definitions and pool hospital rankings across multiple cohort definitions – for example, averaging RSMRs calculated from the *ICD-9* and implicit sepsis cohorts. In a pooled performance metric, hospitals consistently ranked as outliers would remain as outliers, whereas hospitals that switch rankings based upon cohort definition would likely be classified within average performance.(48) Such an approach would capture hospital performance and likely be more resistant to measure gaming.

Evaluation of sepsis as a mortality performance measure is subject to a number of additional limitations. Though modeled on CMS methods, our methods differed from mortality performance measures used by entities such as CMS. We did not have data for post-hospitalization mortality available to calculate 30-day mortality rates traditionally used by CMS, and thus used in-hospital mortality. Although differences in hospital discharge practices may increase variation in in-hospital mortality as compared with 30-day mortality rates, our within-hospital comparisons in hospital performance rankings are unlikely to be affected by between-hospital differences in discharge practices. In addition, we did not use comorbid conditions identified via the CMS hierarchical condition category approach, but

rather, identified comorbidities via a method developed by Gagne at al.(33) Although CMS hierarchical condition category approach may improve risk adjustment over other methods, (49) application of the same comorbidity risk adjustment approach to different cohorts limits the effect of comorbidity risk-adjustment approaches on reliability of sepsis outcome measures. Further, we did not test reliability of ICD-10-based definitions of septic shock; however, prior studies have found similar performance of ICD-9 and ICD-10 codes for identifying sepsis.(50–52) Finally, we acknowledge that gold standard definitions of septic shock to ascertain performance measure validity and optimal measures of test-retest reproducibility to ascertain measure reliability are unclear and require further study.

Conclusion

Although substantial variation between hospitals in septic shock outcomes argues for the need for a sepsis outcome measure, traditional methods to evaluate and compare hospital outcomes were not robust to different methods of identifying septic shock cases. Novel methods, such as pooling hospital rankings across case definitions (e.g., ICD-9 and 'implicit' electronic medical record methods), are likely necessary to produce valid outcome measures for sepsis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

AJW received support from NIH/NHLBI K01HL116768 and R01 HL136660. PKL received support from K24HL 132008. VXL was supported by K23GM112018.

Copyright form disclosure: Drs. Walkey, Liu, and Lindenauer received support for article research from the National Institutes of Health. Dr. Liu's institution received funding from the National Institute of General Medicine Sciences. Dr. Lindenauer disclosed that he receives salary support to develop and maintain hospital outcome measures through a subcontract with the Center for Outcomes Research and Evaluation and the Centers for Medicare and Medicaid Services.

References

1. Elixhauser A, , Friedman B, , Stranges E. Septicemia in US hospitals, 2009 HCUP Agency for Healthcare Research and Quality, editorRockville, MD: 2011
2. Mayr FB, Talisa VB, Balakumar V, et al. Proportion and cost of unplanned 30-day readmissions after sepsis compared with other medical conditions. JAMA. 2017
3. PRISM Investigators. Early, goal-directed therapy for septic shock – A patient-level meta-analysis. N Engl J Med. 2017
4. Prescott HC, Langa KM, Liu V, et al. Increased one-year health care utilization in survivors of severe sepsis. Am J Respir Crit Care Med. 2014
5. Marik PE, Linde-Zwirble WT, Bittner EA, et al. Fluid administration in severe sepsis and septic shock, patterns and outcomes: An analysis of a large national database. Intensive Care Med. 2017; 43:625–632. [PubMed: 28130687]
6. Peltan ID, Mitchell KH, Rudd KE, et al. Physician variation in time to antimicrobial treatment for septic patients presenting to the emergency department. Crit Care Med. 2017; 45:1011–1018. [PubMed: 28426466]
7. Lagu T, Rothberg MB, Nathanson BH, et al. Variation in the care of septic shock: The impact of patient and hospital characteristics. J Crit Care. 2012; 27:329–336. [PubMed: 22300489]

8. Centers for Medicare & Medicaid Services. CMS measures inventory https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityMeasures/CMS-Measures-Inventory.html. Edition
9. QualityNet. Specifications manual, version 5–0b, section 2–2– severe sepsis and septic shock https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1228774725171 Edition
10. Wall MJ, Howell MD. Variation and cost-effectiveness of quality measurement programs. the case of sepsis bundles. Ann Am Thorac Soc. 2015; 12:1597–1599. [PubMed: 26540416]
11. Seymour CW, Gesten F, Prescott HC, et al. Time to treatment and mortality during mandated emergency care for sepsis. N Engl J Med. 2017; 376:2235–2244. [PubMed: 28528569]
12. ACCP, AACN, GNYHA, MWCCC, NAMDRC, SAEM, SHM. Letter of appeal on 0500 severe sepsis and septic shock: Management bundle www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=73103 Edition
13. Klompas M, Rhee C. The CMS sepsis mandate: Right disease, wrong measure. Ann Intern Med. 2016; 165:517–518. [PubMed: 27294338]
14. Krumholz HM, Normand SL, Spertus JA, et al. Measuring performance for treating heart attacks and heart failure: The case for outcomes measurement. Health Aff (Millwood). 2007; 26:75–85. [PubMed: 17211016]
15. Iezzoni LI. The risks of risk adjustment. JAMA. 1997; 278:1600–1607. [PubMed: 9370507]
16. Glance LG, Li Y, Dick AW. Quality of quality measurement: Impact of risk adjustment, hospital volume, and hospital performance. Anesthesiology. 2016; 125:1092–1102. [PubMed: 27662399]
17. Glance LG, Osler TM, Dick AW. Identifying quality outliers in a large, multiple-institution database by using customized versions of the simplified acute physiology score II and the mortality probability model II0. Crit Care Med. 2002; 30:1995–2002. [PubMed: 12352032]
18. Shahian DM, Wolf RE, Iezzoni LI, et al. Variability in the measurement of hospital-wide mortality rates. N Engl J Med. 2010; 363:2530–2539. [PubMed: 21175315]
19. Lindenauer PK, Lagu T, Shieh M, et al. Association of diagnostic coding with trends in hospitalizations and mortality of patients with pneumonia, 2003-2009. JAMA: The Journal of the American Medical Association. 2012; 307:1405–1413. [PubMed: 22474204]
20. Rothberg MB, Pekow PS, Priya A, et al. Variation in diagnostic coding of patients with pneumonia and its association with hospital risk-standardized mortality rates: A cross-sectional analysis. Ann Intern Med. 2014; 160:380–388. [PubMed: 24723078]
21. Angus DC, Seymour CW, Coopersmith CM, et al. A framework for the development and interpretation of different sepsis definitions and clinical criteria. Crit Care Med. 2016; 44:e113–21. [PubMed: 26901559]
22. 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 89-016-1266-9.
23. Singer M, Deutschman CS, Seymour CW, et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). JAMA. 2016; 315:801–810. [PubMed: 26903338]
24. Rhee C, Murphy MV, Li L, et al. Comparison of trends in sepsis incidence and coding using administrative claims versus objective clinical data. Clin Infect Dis. 2014
25. National Quality Forum. Measure evaluation criteria http://www.qualityforum.org/Measuring_Performance/Submitting_Standards/Measure_Evaluation_Criteria.aspx Edition
26. New York State Department of Public Health. New york state report on sepsis care improvement initiative: Hospital quality performance accessed August 4, 2017 Edition. https://www.health.ny.gov/press/reports/docs/2015_sepsis_care_improvement_initiative.pdf, June 2017
27. 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; 151:278–285. [PubMed: 27452768]
28. Centers for Medicare & Medicaid Services. Measure methodology https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Measure-Methodology.html Edition

29. Normand ST, Glickman ME, Gatsonis CA. Statistical applications for profiling providors of medical care: Issues and applications. Journal of the American Statistical Association. 1997; 92:803–814.
30. Bratzler DW, Normand SL, Wang Y, et al. An administrative claims model for profiling hospital 30-day mortality rates for pneumonia patients. PLoS One. 2011; 6:e17401. [PubMed: 21532758]
31. Elixhauser A, Steiner C, Harris DR, et al. Comorbidity measures for use with administrative data. Med Care. 1998; 36:8–27. [PubMed: 9431328]
32. Glance LG, Dick AW, Osler TM, et al. Does date stamping ICD-9-CM codes increase the value of clinical information in administrative data? Health Serv Res. 2006; 41:231–251. [PubMed: 16430609]
33. Gagne JJ, Glynn RJ, Avorn J, et al. A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol. 2011; 64:749–759. [PubMed: 21208778]
34. Martin GS, Mannino DM, Eaton S, et al. The epidemiology of sepsis in the united states from 1979 through 2000. N Engl J Med. 2003; 348:1546–1554. [PubMed: 12700374]
35. Angus DC, Linde-Zwirble WT, Lidicker J, et al. Epidemiology of severe sepsis in the united states: Analysis of incidence, outcome, and associated costs of care. Crit Care Med. 2001; 29:1303–1310. [PubMed: 11445675]
36. Merlo J, Chaix B, Ohlsson H, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: Using measures of clustering in multilevel logistic regression to investigate contextual phenomena. J Epidemiol Community Health. 2006; 60:290–297. [PubMed: 16537344]
37. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull. 1979; 86:420–428. [PubMed: 18839484]
38. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]
39. 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
40. McNamara RL, Wang Y, Partovian C, et al. Development of a hospital outcome measure intended for use with electronic health records: 30-day risk-standardized mortality after acute myocardial infarction. Med Care. 2015; 53:818–826. [PubMed: 26225445]
41. Suter LG, Li SX, Grady JN, et al. National patterns of risk-standardized mortality and readmission after hospitalization for acute myocardial infarction, heart failure, and pneumonia: Update on publicly reported outcomes measures based on the 2013 release. J Gen Intern Med. 2014; 29:1333–1340. [PubMed: 24825244]
42. Bernheim SM, Grady JN, Lin Z, et al. National patterns of risk-standardized mortality and readmission for acute myocardial infarction and heart failure. update on publicly reported outcomes measures based on the 2010 release. Circ Cardiovasc Qual Outcomes. 2010; 3:459–467. [PubMed: 20736442]
43. Lindenauer PK, Bernheim SM, Grady JN, et al. The performance of US hospitals as reflected in risk-standardized 30-day mortality and readmission rates for medicare beneficiaries with pneumonia. J Hosp Med. 2010; 5:E12–8.
44. Krumholz HM, Wang Y, Mattera JA, et al. An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with an acute myocardial infarction. Circulation. 2006; 113:1683–1692. [PubMed: 16549637]
45. Lindenauer PK, Normand SL, Drye EE, et al. Development, validation, and results of a measure of 30-day readmission following hospitalization for pneumonia. J Hosp Med. 2011; 6:142–150. [PubMed: 21387551]
46. Iwashyna TJ, Odden A, Rohde J, et al. Identifying patients with severe sepsis using administrative claims: Patient-level validation of the angus implementation of the international consensus conference definition of severe sepsis. Med Care. 2012
47. 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
48. California Office of Statewide Health Planning and Development, HealthcareCommunity-acquired pneumonia: Hospital outcomes in california, 1999–2001, sacramento CA: California office of

statewide health planning and development; Feb, 2004 http://www.oshpd.ca.gov/HID/Products/PatDischargeData/ResearchReports/OutcomeRpts/CAP/Reports/1999_to_2001.html Edition

49. Li P, Kim MM, Doshi JA. Comparison of the performance of the CMS hierarchical condition category (CMS-HCC) risk adjuster with the charlson and elixhauser comorbidity measures in predicting mortality. BMC Health Serv Res. 2010; 10 245-6963-10-245.
50. Jolley RJ, Sawka KJ, Yergens DW, et al. Validity of administrative data in recording sepsis: A systematic review. Crit Care. 2015; 19 139-015-0847-3.
51. Jolley RJ, Quan H, Jette N, et al. Validation and optimisation of an ICD-10-coded case definition for sepsis using administrative health data. BMJ Open. 2015; 5 e009487-2015-009487.
52. Gedeborg R, Furebring M, Michaelsson K. Diagnosis-dependent misclassification of infections using administrative data variably affected incidence and mortality estimates in ICU patients. J Clin Epidemiol. 2007; 60:155–162. [PubMed: 17208121]

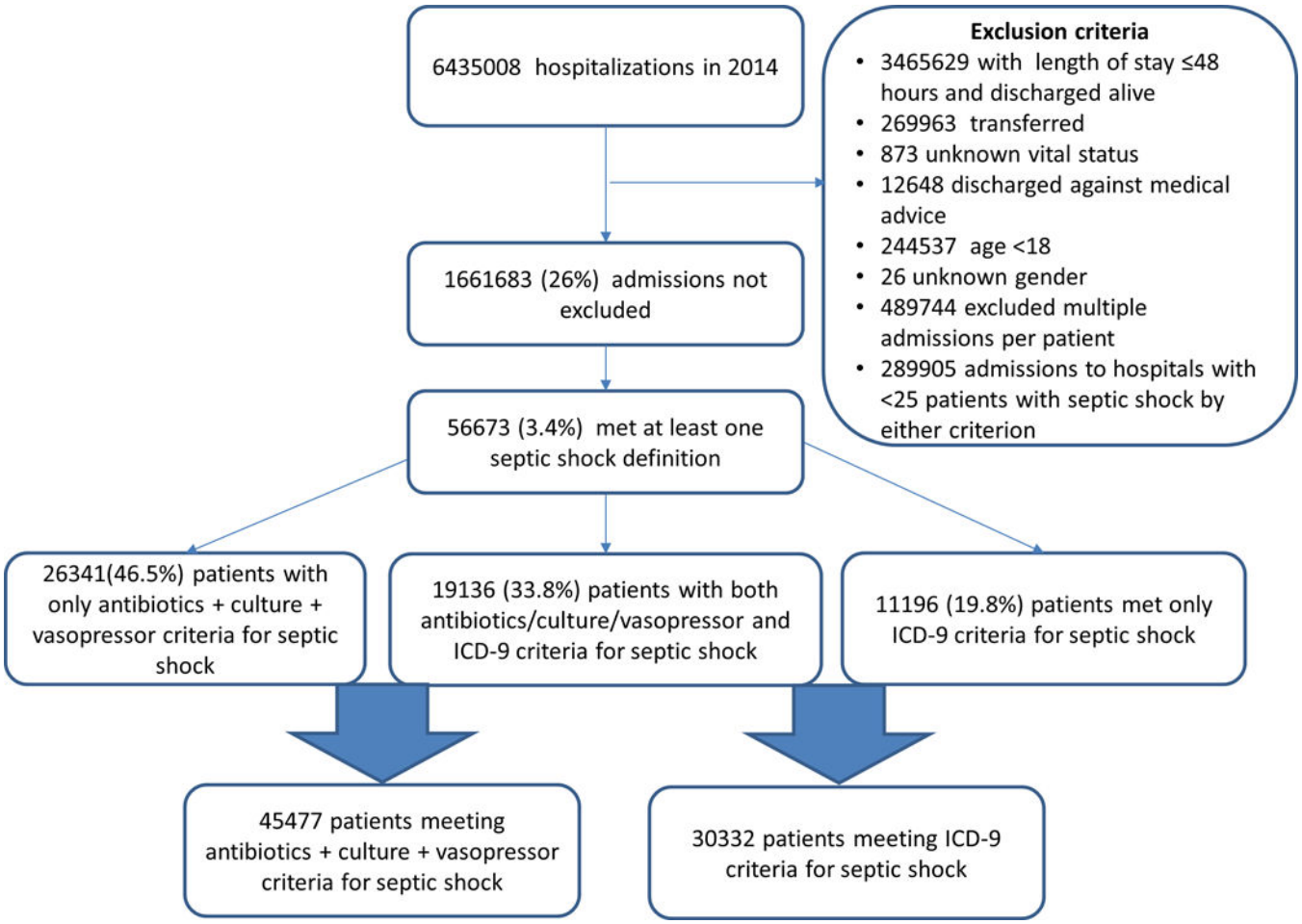

Figure 1.
Flow chart of primary cohort selection.

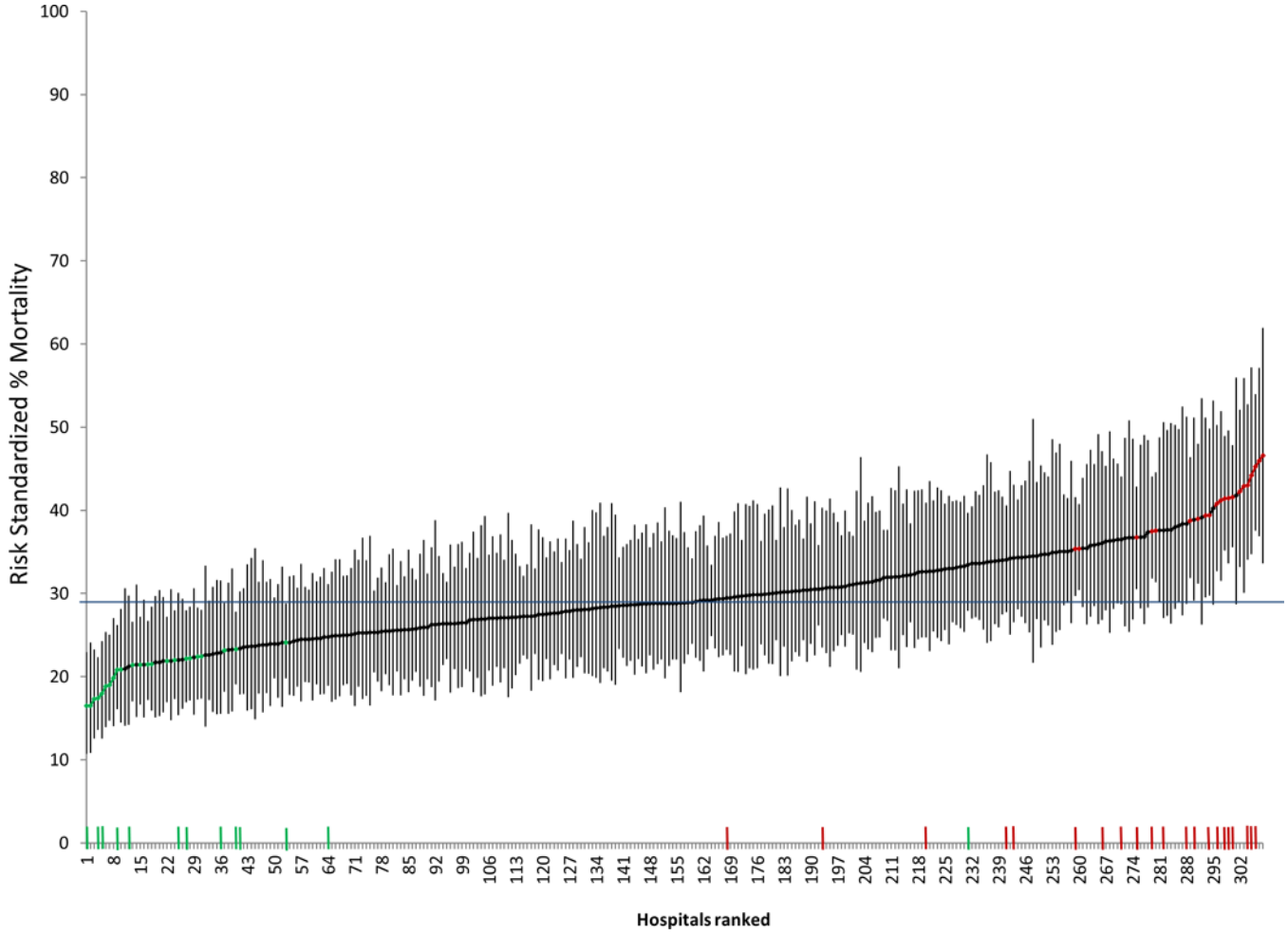

Figure 2.

Caterpillar plot of risk standardized hospital mortality rates for cohort of septic shock defined by charges for antibiotics/cultures/vasopressors (implicit diagnosis).

The horizontal line at 29% is the average mortality rate for the cohort. Points along the caterpillar plot colored green signify low mortality outlier hospitals and red points high mortality outliers. Tick marks along the x-axis demonstrate the hospitals identified as low mortality (green) and high mortality (red) outliers when patient cohorts were defined with *ICD-9* criteria for septic shock.

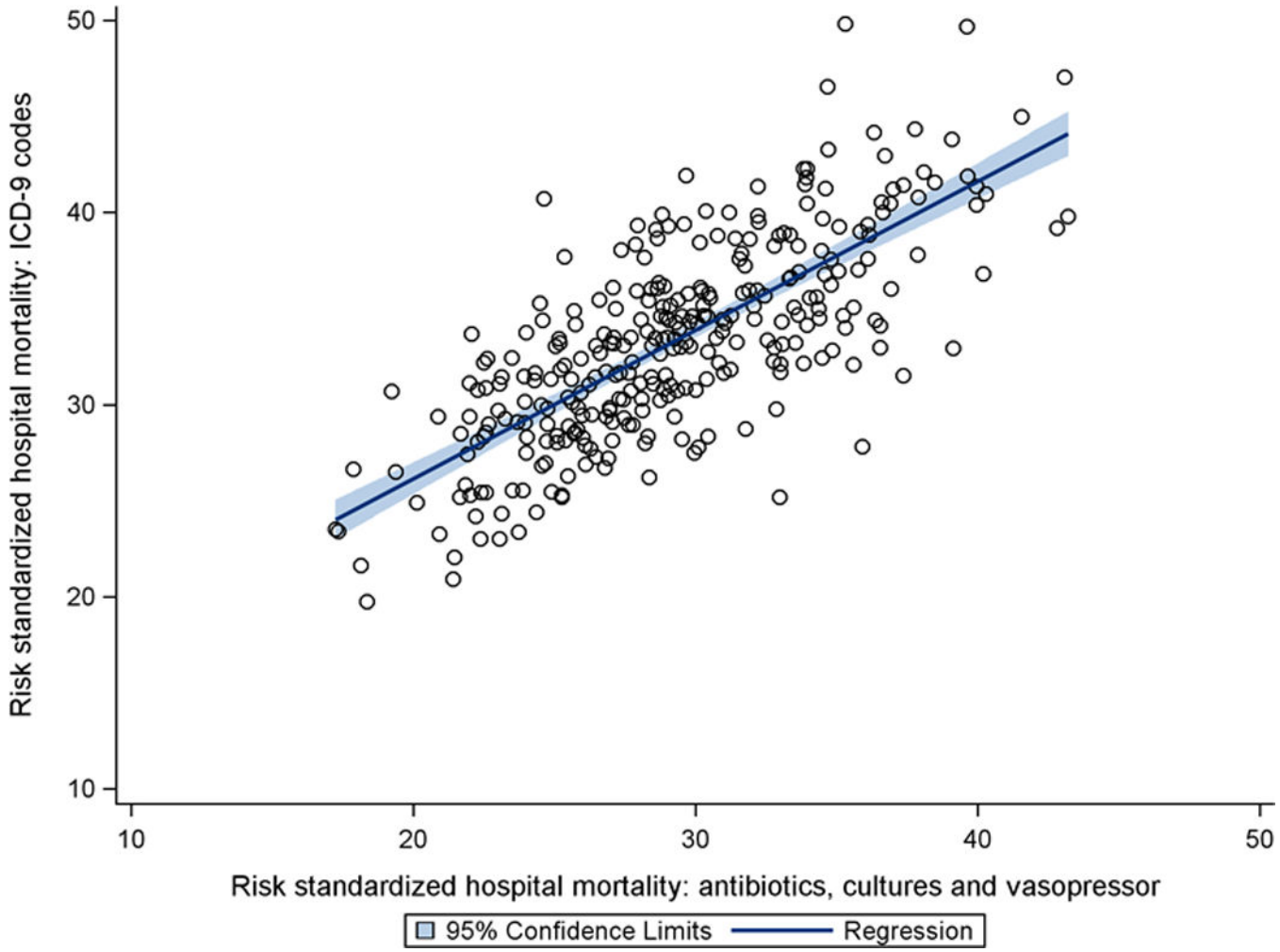

Figure 3.

Scatterplot of hospital risk-standardized mortality rates for septic shock defined by antibiotics/cultures/vasopressors as compared with ICD-9 codes. Intra-class correlation coefficient was 0.72.

Table 1

Comparison of hospital outliers as determined through septic shock cohort defined with ICD-9 codes vs. implicit diagnosis based upon charges for blood cultures, antibiotics, and vasopressors.

N (% of all hospitals)	ICD-9 septic shock cohort			
Charges for any blood cultures, antibiotics, and vasopressors (implicit criteria)		Low mortality outlier	Average hospital	High mortality outlier
	Low mortality outlier	7 (2.3)	16 (5.2)	0 (0)
	Average hospital	6 (2.0)	247 (80.2)	8 (2.6)
	High mortality outlier	0 (0)	12 (3.9)	12 (3.9)

Table 2

Agreement in hospital performance rankings between different methods of identifying “septic shock” within hospital claims data.

Septic shock cohort comparisons	Hospital risk-standardized mortality, Intra-class correlation coefficient, p-value	Hospital Ranking Outlier classification, Kappa statistic (95% CI)	Hospital Ranking Quartile classification, Kappa statistic (95% CI)
Primary analysis			
Implicit septic shock (any duration) vs. *ICD-9* septic shock	0.72, p<0.001	0.44 (0.30-0.58)	0.53 (0.47-0.60)
Sensitivity analyses			
Implicit septic shock (any duration) vs. implicit septic shock (minimum antibiotics and vasopressor duration)	0.76, p<0.001	0.60 (0.47-0.73)	0.52 (0.46-0.59)
Implicit septic shock (minimum antibiotics and vasopressor duration) vs. *ICD-9* septic shock	0.62, p<0.001	0.35 (0.19-0.50)	0.44 (0.37- 0.51)
ICD-9 septic shock vs. *ICD-9* sepsis	0.74, p<0.001	0.32 (0.23-0.42)	0.62 (0.56-0.68)

Implicit septic shock (any duration) definition: Any blood culture, antibiotics, and vasopressors charged during first 2 hospital days.

Implicit septic shock (minimum antibiotics and vasopressor duration) definition: Charges for blood culture, and at least 4 days of antibiotics and 2 days of vasopressors (unless death precedes minimum duration), starting within the first 2 hospital days.

***ICD-9* septic shock**: *ICD-9* 038.x for septicemia present on admission with *ICD-9* 785.52 for septic shock.

***ICD-9* sepsis**: *ICD-9* 038.x for septicemia present on admission