Published in final edited form as:

Crit Care Med. 2019 February ; 47(2): 303–305. doi:10.1097/CCM.0000000000003564.

Sepsis computable phenotypes in the service of observational research

Blake Martin, MD and Tellen D. Bennett, MD, MS

Pediatric Critical Care, University of Colorado School of Medicine, Aurora, CO

Keywords

sepsis; septic shock; pediatrics; intensive care; classification; phenotypes

Sepsis, “life-threatening organ dysfunction caused by a dysregulated host response to infection,”(1) is a leading cause of pediatric morbidity and mortality worldwide.(2) Despite its importance, pediatric sepsis is relatively rare at any one center. Multi-center studies are needed to accurately estimate incidence, describe disease trajectories, test interventions, and evaluate long-term outcomes. Prospective multi-center pediatric sepsis studies are expensive to conduct and uncommon. Randomized trials are rare. Leveraging multi-center observational data resources such as critical care registries to answer questions that do not require randomization has the potential to generate useful evidence at low cost.(3)

However, in order to study patients with sepsis, we have to find them. Identifying pediatric (and adult) patients with sepsis in existing data has been problematic.(4, 5) The “probes” used to detect patients with a particular condition in electronic health record (EHR) data are called computable phenotypes.(6)

Two kinds of sepsis computable phenotypes are currently in use. Phenotypes based on inpatient billing data such as International Classification of Diseases, 9th revision, Clinical Modification (ICD-9-CM) diagnosis codes look for patients with sepsis after a hospitalization has been completed. Examples include the Angus(7) and Martin(8) criteria and the “explicit” severe sepsis (995.92) and septic shock (785.52) ICD-9-CM codes.(9) These phenotypes might be used to find patients for quality improvement work or observational research.(10) A second kind of phenotype attempts to detect or predict sepsis within an EHR in real-time. These might be used to trigger clinical interventions or for prospective study recruitment. Examples include phenotypes based on systemic inflammatory response syndrome (SIRS)(11), Sequential Organ Failure Assessment scores (SOFA, quick SOFA)(1), and more complex machine learning algorithms.(12)

The Virtual Pediatric Systems (VPS, <https://www.myvps.org>) registry is a likely choice for pediatric sepsis research. It contains prospectively identified patients from more than 135

Pediatric Intensive Care Units (PICUs). Detailed procedural codes, simple functional status metrics, and illness severity measures make VPS data valuable for observational studies. Trained coders with (usually clinical) pediatric critical care expertise extract data from the EHR. One important part of this process is that the VPS coders assign ICD-9-CM-based diagnosis codes to patients using medical record review and a proprietary diagnosis rubric. Therefore, VPS diagnosis codes are not the ICD-9-CM codes assigned by hospital billers, available in aggregated administrative data sources and used in other computable phenotypes. Because the coders have clinical expertise, VPS diagnosis codes might be more accurate than billing codes.

In this issue of *Critical Care Medicine*(13), Lindell et al. sought to test that hypothesis by comparing the ability of the Angus and Martin criteria and the “explicit” ICD-9-CM diagnosis codes to identify children with severe sepsis or septic shock in the VPS database at one institution. They leveraged a prospective cohort identified using daily screening on PICU morning rounds with the Goldstein consensus criteria(11) for severe sepsis or septic shock. Screeners then reviewed the medical record of patients included in the prospective cohort to determine sepsis severity.

Lindell et al. used the prospective cohort as a gold standard against which to evaluate the computable phenotypes. Their study cohort included all 7,459 children admitted to a single large PICU from March 1st 2012 to March 31st 2014. Of those, 401 were prospectively identified as having severe sepsis or septic shock. VPS diagnosis codes from all 7,459 patients were grouped into the Angus, Martin, and “explicit” code criteria and tested to see how well they identified the 401 gold standard patients.

Not surprisingly, the explicit criteria performed best. One of the two ICD-9-CM diagnosis codes (995.92 or 785.52) was present for 515/7,459 patients. This represents a sensitivity of 0.73, a positive predictive value (PPV) of 0.57, and a F1 measure of 0.64. Specificity and percent correctly classified (agreement) are heavily influenced by a large number of true negatives, as is the case in this study. The more balanced metrics such as sensitivity/recall, PPV/precision, and F1 measure are more useful in evaluating overall performance. The patients selected by the explicit codes also most closely resembled the gold standard cohort with respect to age, illness severity, resource utilization, and PICU mortality. The Angus (sensitivity 0.77, PPV 0.22, F1 0.34) and Martin (sensitivity 0.82, PPV 0.34, F1 0.48) criteria detected severe sepsis or septic shock less effectively.

It is also not surprising that the Angus and Martin criteria identified a larger group of patients with lower illness severity and lower specificity. The Angus and Martin criteria were developed to identify patients with infection and organ dysfunction and sepsis of any severity, respectively. They are not, strictly speaking, severe sepsis or septic shock phenotypes. Using a retrospective gold standard, investigators from the same VPS site previously reported that the explicit codes identified pediatric severe sepsis or septic shock in standard billing diagnoses better (sensitivity 0.73, PPV 0.79) than combinations like the Angus criteria (sensitivity 0.15, PPV 0.18).(5) A study in adults showed similar results (PPV explicit codes 1.0 versus Angus 0.71).(9) As others have noted, explicit codes identify a

more “pure”(14) cohort of patients with severe sepsis or septic shock at the expense of missing many patients with lower-severity sepsis.

The recently released Sepsis-3 guidelines(1) have significantly changed the landscape of sepsis identification. First, they capture an even larger and more heterogeneous cohort of adults with sepsis than SIRS-based methods or the Angus criteria.(14) Second, the potentially confusing entity of severe sepsis is no longer present in Sepsis-3, which makes one of the explicit codes (995.92) less relevant going forward. Third, the Sepsis-3 definition is no longer based on SIRS. This is likely advantageous to those studying pediatric sepsis because so many children qualify as having SIRS when they are ill but not septic. SIRS-based pediatric sepsis definitions like the consensus criteria used by the current study(11) are the best available gold standard but can likely be improved upon. Unfortunately, Sepsis-3 currently lacks a pediatric sepsis definition, although investigators are working on adapting the qSOFA score to pediatric physiologic norms.(15) One very important contribution of Sepsis-3 is the estimation of timing of sepsis onset provided by the SOFA and qSOFA scores. Those scores and other sequential prediction algorithms, once optimized, have the potential to make sepsis detection and prediction possible.

One of Lindell et al.’s incidental findings reinforces the challenge of detecting patients with sepsis during a clinical care episode (in near real-time). Even with a daily standardized approach, the prospective gold standard cohort missed 66 patients that the VPS explicit criteria found. Had these patients been captured prospectively, the performance of the VPS explicit criteria would have been even better (PPV 0.70, F1 0.75). The authors’ careful chart review of discordant patients (those detected either prospectively or in VPS but not both) also revealed two interesting categories of patients who did not have sepsis and posed challenges for both SIRS-based and diagnosis code phenotypes: multiple organ dysfunction syndrome (MODS) and cytokine release syndrome (CRS) after chimeric antigen receptor T cell therapy. Because PICUs are likely to care for more patients with MODS and CRS in the future, this is an important limitation for sepsis investigators to understand and manage.

Strengths of the manuscript include the large and prospectively identified reference cohort and the careful exploratory analysis of the discordant patients. One limitation is that the phenotypes were tested in VPS data. The VPS database is available only to participating sites. That limitation is mitigated somewhat by the fact that more than 135 PICUs currently submit data to VPS. Other limitations include that the phenotypes were tested at only one VPS site. Despite guidance from the VPS registry administration, variance in diagnosis coding practices between VPS sites may exist. PICU admission criteria are also known to vary between hospitals, although most patients with severe sepsis or septic shock would be in a PICU at most centers. The phenotypes tested here are based on ICD-9-CM, which U.S. hospitals stopped using for billing in late 2015. ICD-10-based phenotypes will have to be tested in children for use with newer billing data.

Overall, the explicit severe sepsis and septic shock codes have only moderate accuracy in VPS data: sensitivity 0.73, PPV 0.57, and F1 measure 0.64. Investigators using these codes to identify cohorts for observational studies using VPS data should understand and report their limitations and perform careful sensitivity analyses to assess the robustness of their

conclusions. Validation at other VPS centers and using only billing ICD diagnosis codes would be valuable as well.

Increasing availability of high-dimensional EHR data is likely to support the development of accurate near real-time sepsis detection and prediction tools. Challenges going forward will include translation of models and prototypes into health systems where they can be trialed, interoperability of the tools across different EHRs, and integration of the tools into effective multidisciplinary sepsis workflows at each center.

Acknowledgments

TB was partially supported by NICHD R03 HD094912

Dr. Bennett's institution received funding from the National Institute of Child Health and Human Development, and he received support for article research from the National Institutes of Health.

References

1. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016;315(8):801–810. [PubMed: 26903338]
2. Weiss SL, Fitzgerald JC, Pappachan J, et al. Global epidemiology of pediatric severe sepsis: the sepsis prevalence, outcomes, and therapies study. *Am J Respir Crit Care Med* 2015;191(10):1147–1157. [PubMed: 25734408]
3. Bennett TD, Spaeder MC, Matos RI, et al. Existing data analysis in pediatric critical care research. *Front Pediatr* 2014;2:79. [PubMed: 25121079]
4. Weiss SL, Parker B, Bullock ME, et al. Defining pediatric sepsis by different criteria: discrepancies in populations and implications for clinical practice. *Pediatr Crit Care Med* 2012;13(4):e219–226. [PubMed: 22460773]
5. Balamuth F, Weiss SL, Hall M, et al. Identifying Pediatric Severe Sepsis and Septic Shock: Accuracy of Diagnosis Codes. *J Pediatr* 2015;167(6):1295–1300 e1294. [PubMed: 26470685]
6. Richesson RL, Sun J, Pathak J, et al. Clinical phenotyping in selected national networks: demonstrating the need for high-throughput, portable, and computational methods. *Artif Intell Med* 2016;71:57–61. [PubMed: 27506131]
7. 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(7):1303–1310. [PubMed: 11445675]
8. 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(16):1546–1554. [PubMed: 12700374]
9. 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* 2014;52(6):e39–43. [PubMed: 23001437]
10. Seymour CW, Coopersmith CM, Deutschman CS, et al. Application of a Framework to Assess the Usefulness of Alternative Sepsis Criteria. *Crit Care Med* 2016;44(3):e122–130. [PubMed: 26901560]
11. Goldstein B, Giroir B, Randolph A, et al. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatr Crit Care Med* 2005;6(1):2–8. [PubMed: 15636651]
12. Henry KE, Hager DN, Pronovost PJ, et al. A targeted real-time early warning score (TREWScore) for septic shock. *Sci Transl Med* 2015;7(299):299ra122.
13. Lindell RB, Nishisaki A, Weiss SL, et al. Comparison of Methods for Identification of Pediatric Severe Sepsis and Septic Shock in the Virtual Pediatric Systems Database. *Crit Care Med* 2019; in press

14. Johnson AEW, Aboab J, Raffa JD, et al. A Comparative Analysis of Sepsis Identification Methods in an Electronic Database. *Crit Care Med* 2018;46(4):494–499. [PubMed: 29303796]
15. Matics TJ, Sanchez-Pinto LN. Adaptation and Validation of a Pediatric Sequential Organ Failure Assessment Score and Evaluation of the Sepsis-3 Definitions in Critically Ill Children. *JAMA Pediatr* 2017;171(10):e172352. [PubMed: 28783810]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript