

ARTICLE OPEN



Developing a shared sepsis data infrastructure: a systematic review and concept map to FHIR

Emily B. Brant^{1,2,3,6}✉, Jason N. Kennedy^{1,2}, Andrew J. King¹, Lawrence D. Gerstley⁴, Pranita Mishra⁴, David Schlessinger⁴, James Shalaby⁵, Gabriel J. Escobar⁴, Derek C. Angus^{1,2}, Christopher W. Seymour^{1,2,3,7} and Vincent X. Liu^{4,7}

The development of a shared data infrastructure across health systems could improve research, clinical care, and health policy across a spectrum of diseases, including sepsis. Awareness of the potential value of such infrastructure has been heightened by COVID-19, as the lack of a real-time, interoperable data network impaired disease identification, mitigation, and eradication. The *Sepsis on FHIR* collaboration establishes a dynamic, federated, and interoperable system of sepsis data from 55 hospitals using 2 distinct inpatient electronic health record systems. Here we report on phase 1, a systematic review to identify clinical variables required to define sepsis and its subtypes to produce a concept mapping of elements onto Fast Healthcare Interoperability Resources (FHIR). Relevant papers described consensus sepsis definitions, provided criteria for sepsis, severe sepsis, septic shock, or detailed sepsis subtypes. Studies not written in English, published prior to 1970, or “grey” literature were prospectively excluded. We analyzed 55 manuscripts yielding 151 unique clinical variables. We then mapped variables to their corresponding US Core FHIR resources and specific code values. This work establishes the framework to develop a flexible infrastructure for sharing sepsis data, highlighting how FHIR could enable the extension of this approach to other important conditions relevant to public health.

npj Digital Medicine (2022)5:44; <https://doi.org/10.1038/s41746-022-00580-2>

INTRODUCTION

Sepsis is among the deadliest medical conditions recognized worldwide. It occurs when infection triggers a dysregulated systemic immune response and can rapidly lead to organ failure and death¹. Recent Global Burden of Disease estimates suggests that there were nearly 50 million cases in 2017 that resulted in 11 million deaths². In the US, sepsis is the single most costly cause of hospitalization and contributes to as many as half of all hospital deaths^{2,3}. Early diagnosis and treatment of infected patients with sepsis or high risk of progression to sepsis is imperative. Yet, sepsis diagnosis remains a major challenge due to several factors, including the absence of a gold standard test⁴, variability of patient presentation⁵, heterogeneity of illness progression⁶, among others^{7–10}.

There is an urgent need to improve diagnostic excellence in sepsis¹¹. One potential solution is the development of a flexible, scalable, and interoperable data infrastructure to screen and identify sepsis patients across health systems using real-time, granular clinical data available within electronic health records (EHRs)¹¹. Despite clinical data for millions of sepsis or pre-sepsis patients routinely collected in EHRs, the healthcare enterprise has accessed only a small fraction of these data to improve the clinical understanding of sepsis. The development of a sepsis data backbone could be extended to meet diverse needs, including sepsis translational research, clinical care delivery, machine learning/artificial intelligence deployment, disease surveillance, quality improvement, and health policy. Furthermore, once established in sepsis, a similar data infrastructure could be tailored to include other conditions. For example, awareness of the potential value of this data infrastructure has

been magnified by the COVID-19 pandemic—many COVID-19 inpatients met sepsis criteria with systemic inflammation, organ failure, and a high risk of mortality—highlighting the critical need for real-time data interoperability that could inform pandemic response, health system preparedness, translational research and clinical care^{12–14}.

The Fast Healthcare Interoperability Resource (FHIR) Health Level Seven International (HL7) standard has shown great potential for modernizing the interchange of data through standardized FHIR resources and application programming interfaces (APIs) to access and exchange these data^{15,16}. Indeed, FHIR has been endorsed by the Center for Medicare and Medicaid Services (CMS) and the Office of the National Coordinator for Health Information (ONC) as the preferred standard for EHR interoperability¹⁷.

In this study, we report on the development of *Sepsis on FHIR*, a framework to accelerate the interoperability of sepsis data across healthcare systems and EHRs. The work is a collaboration between Kaiser Permanente Northern California and the University of Pittsburgh Medical Center (UPMC) health systems to establish a dynamic, federated, and interoperable system of sepsis data from 55 diverse hospitals using two distinct inpatient EHR software systems. Here, we discuss phase 1 of *Sepsis on FHIR*. We first conducted a systematic review of the sepsis literature and identified a comprehensive set of relevant EHR clinical variables as a sharable set of features across sites. We then produced a concept mapping of clinical variables to FHIR resources. The resulting infrastructure can be adapted for the eventual expansion to other hospitals with different patient populations.

¹Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. ²Clinical Research, Investigation and Systems Modeling of Acute Illness (CRISMA) Center, Pittsburgh, PA, USA. ³Department of Emergency Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. ⁴Kaiser Permanente Division of Research, Oakland, CA, USA. ⁵Elimu Informatics, Richmond, CA, USA. ⁶Present address: Assistant Professor of Critical Care and Emergency Medicine, University of Pittsburgh School of Medicine, 200 Lothrop Street, #607, Pittsburgh, PA 15261, USA. ⁷These authors contributed equally: Christopher W. Seymour, Vincent X. Liu.

✉email: branteb@upmc.edu

RESULTS

Study characteristics

We conducted a systematic review that retrieved 4812 citations; removing duplicate, non-English, non-full-text citations yielded 1075 full-text manuscripts (Fig. 1). After manual review, we excluded 1022 manuscripts that did not meet inclusion criteria or reported previously cited sepsis definitions, leaving 55 unique articles for analysis (Supplementary Data set 1). These included 8 consensus definitions, 7 systematic reviews and meta-analyses, 9 prospective cohort studies, 26 retrospective cohort studies, and 5 narrative literature reviews (Supplementary Fig. 1). All studies were reported between 1989 and 2020 (Supplementary Fig. 2).

The systematic review confirmed that several definitions and criteria were used to identify sepsis patients. Studies published prior to international consensus conference definitions identified sepsis patients using standardized criteria for patient enrollment used in trials performed by Bone et al.^{18–21}. Organ dysfunction was defined using clinical markers of hypoperfusion including altered mentation, hypoxemia, elevated lactic acid, and oliguria¹⁸. Following the First International Consensus on Sepsis and Septic Shock, 16 studies analyzed defined sepsis patients using Systemic Inflammatory Response Syndrome (SIRS) criteria^{6,20,22–35}. After the Sepsis-3 task force in 2016, 14 studies analyzed defined sepsis patients using the Sequential Organ Failure Assessment (SOFA) score^{6,23–25,27,30–32,34–39}. Two studies defined organ dysfunction in sepsis patients using the logistic organ dysfunction score (LODS); two studies reported the multiple organ dysfunction syndrome (MODS)^{21,22,40,41}. The Acute Physiologic Assessment and Chronic Health Evaluation (APACHE) score defined sepsis-associated organ dysfunction in eight analyzed studies^{22,26,29,34,37,42–44}. Seven studies defined sepsis using claims-based definitions or International Classification of Diseases (ICD) standards^{28,33,34,44–48}.

Definitions used to establish suspected or proven infection varied considerably. Nearly half (46%) of studies incorporated blood culture data, including body fluid sampling for culture or culture positivity. Many, however, relied on clinical suspicion as defined by the treating physician^{18,34,35,41,43,49–55}. Two studies used clinical risk scores, including the clinical pulmonary infection score and infection probability score^{35,43,56}.

Data elements

We identified 788 clinical variables from the 55 manuscripts (Supplementary Table 3). Recorded data elements included clinical measurements (e.g., heart rate, respiratory rate, white blood cell count), infectious signs and symptoms (e.g., dysuria, abdominal pain), and individual ICD codes for diagnosis and procedures, current procedural terminology (CPT) and diagnosis-related groups (DRG) codes for sepsis, septicemia, and sepsis syndromes, organ dysfunction (e.g., acute kidney injury, hypotension) and

infection (e.g., sepsis due to anaerobes, candidal sepsis) (Supplementary Table 3).

After the removal of duplicates and unstructured variables, 151 unique clinical variables remained (Supplementary Table 4). Variables represented 7 broad domains including, patient characteristics, vital signs and laboratory tests, interventions data (e.g., medication administration, diagnostic tests, catheterization, surgical events), fluid balance, location information, healthcare use and outcomes, and administrative/billing codes. All major sepsis consensus definitions were represented by selected clinical variables, capturing the evolution of sepsis definitions over time (Fig. 2).

The most frequently reported patient demographic characteristics included age, sex, and measures of comorbidity (Fig. 3). Vital signs and laboratory tests were both used for sepsis definitions and to describe baseline characteristics. Infection was most often described by body fluid sampling for culture and white blood cell count. Organ dysfunction was most commonly described using respiratory rate, temperature, Glasgow Coma Scale (GCS) score, and urine output. Other variables, including troponin, albumin level, and erythrocyte sedimentation rate, were reported only once, but were considered critical for defining specific sepsis subtypes⁶.

FHIR mapping

We mapped the 151 clinical variables to their corresponding FHIR resources and, where appropriate, to specific code values that represented each element (Supplementary Table 5). Variables mapped most frequently to the FHIR DiagnosticReport resource (41%), followed by the observation (33%), patient (12%), encounter (9%), or procedure (5%) resources (Supplementary Table 5).

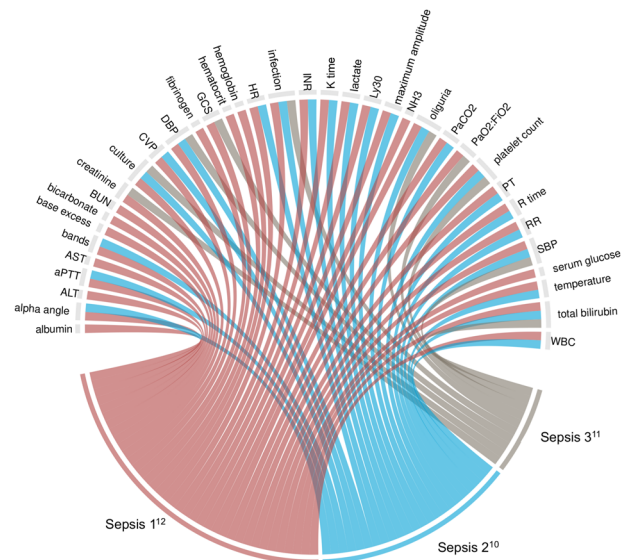


Fig. 2 Chord diagrams showing the representation of clinical elements across international consensus definitions for sepsis. Ribbons connect from each of three international consensus definitions for sepsis to data elements. Only those data elements encompassed by sepsis consensus definitions are displayed. **Abbreviations:** ALT alanine aminotransferase, aPTT activated partial thromboplastin time, AST aspartate aminotransferase, BUN blood urea nitrogen, CVP central venous pressure, DBP diastolic blood pressure, GCS Glasgow coma scale score, HR heart rate, INR international normalized ratio, Ly30 lysis in 30 minutes, NH3 ammonia level, PaCO₂ partial pressure of arterial carbon dioxide, FiO₂ fraction of inspired oxygen, PT prothrombin time, RR respiratory rate, SBP systolic blood pressure, WBC white blood cell count.

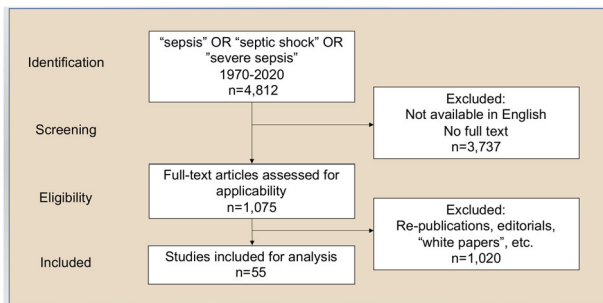


Fig. 1 Flowchart for study selection. Initial searches identified 4812 manuscripts. After removing duplicate, non-English, non-full-text citations, 1075 full-text manuscripts were reviewed against inclusion criteria. After exclusion, 55 unique studies were included in the analysis.

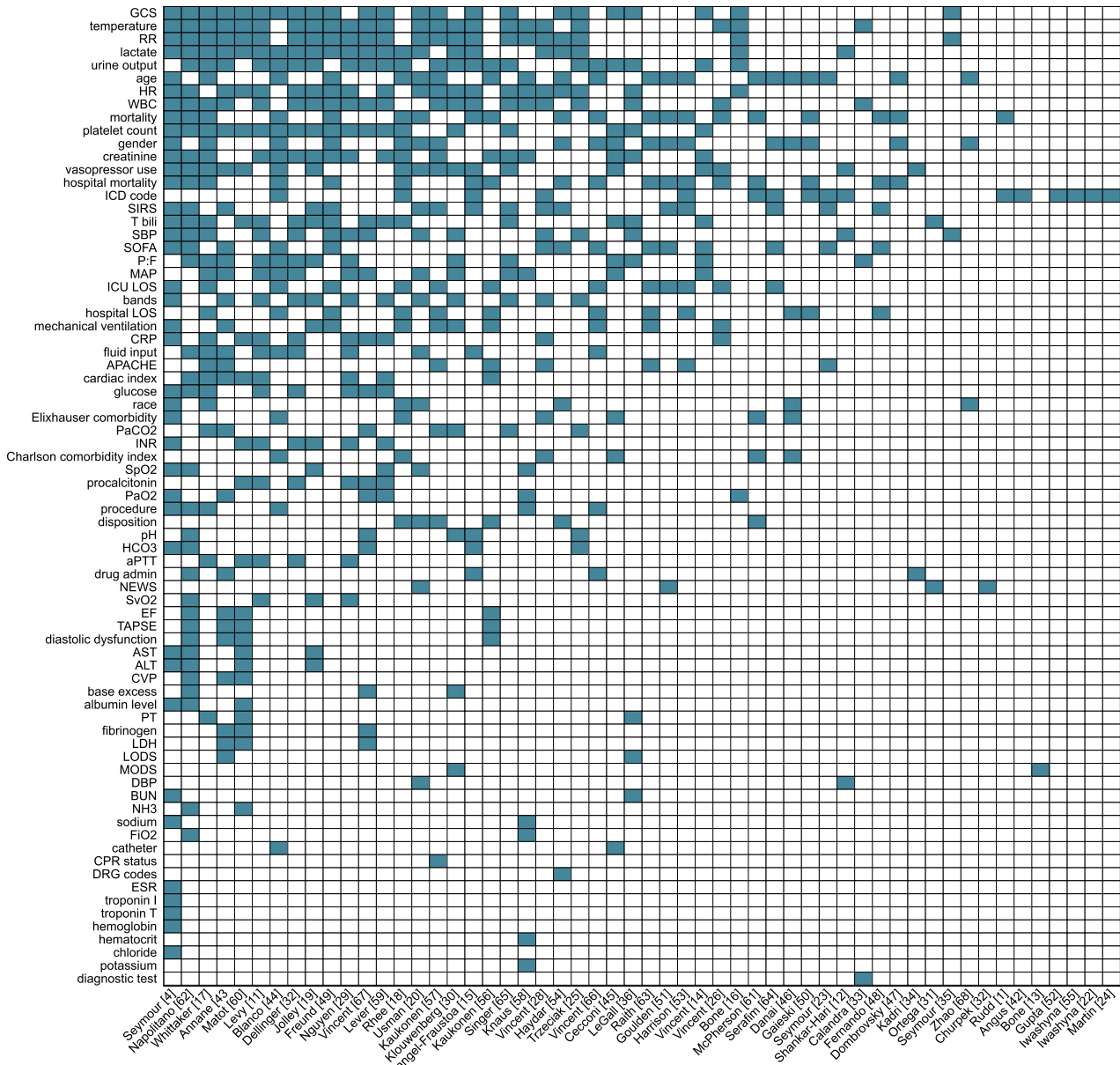


Fig. 3 Data elements reported by a manuscript in heatmap. Data elements are displayed on the y axis; manuscripts included in analysis displayed on x axis. Axes are sorted so variables most frequently cited are grouped together on the left, whereas manuscripts contributing infrequently reported variables are grouped on the right. Blue shading represents the data element reported by the manuscript, whereas data elements not reported by the manuscript are not shaded. RR respiratory rate, GCS Glasgow Coma Scale score, WBC white blood cell count, HR heart rate, T Bili, total bilirubin, SIRS systemic inflammatory response syndrome, SBP systolic blood pressure, P:F ratio of partial pressure of arterial oxygen to fraction of inspired oxygen, ICD international classification of diseases, SOFA sequential organ failure assessment score, ICU intensive care unit, LOS length of stay, MAP mean arterial pressure, CRP c-reactive protein, APACHE acute physiology and chronic health enquiry, INR international normalized ratio, PaCO₂ partial pressure of arterial carbon dioxide, SpO₂ oxygen saturation, aPTT activated partial thromboplastin time, HCO₃ bicarbonate, ALT alanine aminotransferase, AST aspartate aminotransferase, TAPSE tricuspid annular plane systolic excursion, EF ejection fraction, SvO₂ saturation of venous oxygen, NEWS national early warning system, LDH lactate dehydrogenase, PT prothrombin time, CVP central venous pressure, FiO₂ fraction of inspired oxygen, NH₃ ammonia level, BUN blood urea nitrogen, DBP diastolic blood pressure, MODS multiple organ dysfunction syndrome, LODS logistic organ dysfunction syndrome, ESR erythrocyte sedimentation rate, DRG diagnosis-related groups.

To promote flexibility, we linked variables to several FHIR resource types whenever possible. For example, the variable “serum sodium” is represented as a FHIR Observation resource where the Observation.code data element is mapped to the Logical Observation Identifiers Names and Codes (LOINC) code 2951-2 (referring to “Sodium [Moles/volume] in Serum or Plasma”). LOINC is the most commonly used international healthcare terminology standard which describes a reference set of health

data and codes for laboratory and clinical observations. FHIR mappings frequently link to more than one LOINC code. For example, blood urea nitrogen (BUN) mapped to both LOINC codes 3094-0 (BUN SerPl-mCnc) and 6299-2 (BUN Bld-mCnc). For variables linked to more than one code, we established a HL7 FHIR-based sepsis value set to represent a set of codes with the same clinical meaning. Most variables (99 out of 151) were FHIR resource observations with direct mappings to LOINC codes.

In contrast, the variable “Admit Time” mapped directly to an existing FHIR data element for healthcare encounter time stamps (Encounter.period.start).

We also identified some variables that require additional computation. For example, “Days of Vasopressors”—the duration of treatment with vasoactive agents to support blood pressure—requires calculation using variables mapped to other resources. The start and end dates of vasopressor therapy, which are represented as the FHIR MedicationAdministration resource, are required. In addition, these data may also be represented by RxNORM codes, the US reference terminology standard representing vasopressors (e.g., epinephrine, norepinephrine).

Taken together, clinical variables were successfully mapped to US Core FHIR elements, LOINC, and RxNorm codes to ensure and enhance generalizability and flexibility across health systems and EHR.

DISCUSSION

In phase 1 of *Sepsis on FHIR*, we conducted a systematic review of 55 studies, identifying 151 variables relevant to sepsis diagnosis, sepsis definitions and subtyping. We then mapped these variables to a data standard (FHIR) endorsed by national healthcare entities¹⁷. The goal of this work is not to define a final or complete set of sepsis diagnostic and clinical criteria, nor to develop a data interoperability resource limited only to two large US healthcare systems. Instead, we aim to establish a publicly available *Sepsis on FHIR* resource that will promote the use of a flexible, scalable, federated, and interoperable source of clinical data that can be contributed to and used by many stakeholders. Our work could be easily applied to other clinical conditions.

This work focused on sepsis as it is common, costly, and deadly, with significant opportunities to improve treatment and outcomes². Sepsis treatment can be challenging due to a number of factors, however, the key step to effective treatment is early diagnosis. Timely diagnosis is the critical factor determining patient outcomes, since delays in diagnosis (with concomitant delayed treatment with antibiotics, resuscitation, and organ system stabilization) are associated with increased mortality^{7,57}. Numerous leading entities, including the National Institute of General Medical Sciences and the National Academy of Medicine, have identified the lack of standardized and interoperable data resources as a critical limitation to improving sepsis diagnosis and treatment^{6,58}.

Despite the availability of millions of EHR records detailing patients’ presentation, organ failure, pathogen, treatment, and outcomes today, the lack of access to these data in aggregate stymie further progress⁵⁹. Multicenter data may be available in some instances, such as within networks of hospitals linked by a single EHR software vendor or within larger health systems or research consortia, for example, but there is a lack of a standardized resource for diverse stakeholders who wish to contribute to or access these data. Larger registries of sepsis data exist through state or national efforts, however, these comprise a relatively narrowly-defined set of criteria targeted to performance or quality improvement efforts⁶⁰. Further, the cost of maintaining these data has also raised alarm among entities subject to reporting requirements^{61,62}.

The development of an automated, flexible infrastructure for sharing sepsis data represents a key step toward achieving diagnostic excellence and improving sepsis outcomes for several reasons. First, compiling high-fidelity health data from millions of patient encounters creates a substantial opportunity to uncover sepsis subgroups⁶³. Recent studies have highlighted the heterogeneity in treatment effects among specific subgroups which demonstrates potentially adverse effects of current treatment approaches in some patient subtypes^{6,64}. Thus, establishing representative subtypes and identifying them in real-time will

be key for initiating clinical trials and targeting treatments. Second, while focused, small repositories of biospecimens exist in sepsis, they cannot be easily appended to detailed clinical data shared across sites to improve their utility and statistical power⁶⁵. Third, as described above, current reporting requirements for the CMS SEP-1 measure is highly resource-intensive and of uncertain benefit for patients⁶². Automated extraction of patient data can improve reporting while also helping to refine the use of policy-driven targets. Fourth, while there are many emerging machine learning and artificial intelligence algorithms designed to improve the prediction of sepsis onset or deterioration, they can be brittle (i.e., predictive performance degrades) when developed in one healthcare context and exported to an external context⁶⁶. A federated learning system built using data from many diverse environments can improve the generalizability, representativeness, and transparency of such algorithms for clinical care. Fifth, as evident in the COVID-19 pandemic, the lack of a national real-time system of interoperable data substantially hampers the identification, mitigation, and eradication of large-scale communicable diseases¹². Finally, the extension of this platform beyond the United States may allow for standardized international data exchange.

There are many steps that logically follow from this work. Using the Sepsis on FHIR foundation, our two health systems will instantiate FHIR repositories at each site while continuing work to overcome several barriers. Among these continuing barriers, the most important are the instability of the EHRs (in which codes are frequently changed/alterd), persistence of data elements that are unmapped, and difficulties in bulk extraction of FHIR-mapped data elements. For example, real-time extraction of FHIR data in sepsis can be costly and, thus, shadow or parallel data systems that offload computation to non-operational systems may be needed to ensure the availability of data relevant to clinical care or trials. In addition, data elements collected serially will need to be refreshed at a clinically relevant tempo (Supplementary Table 5).

Further, several key issues involving data governance, privacy, and security must be addressed prior to the aggregation of any patient data. To address data governance and privacy, the development of a federated data infrastructure (i.e., local data that never leave an individual system) reduces the risk associated with data transmission. The FHIR standard supports key interoperability enablers that can support a federated analytics platform⁶⁷. To ensure data security, FHIR also supports secure exchange of clinical data and the ability to populate powerful standards-based research analytics platforms such as PCORNET Common Data Model (CDM)⁶⁸. Deidentified FHIR-based data can be used to populate CDMs at each local site that can in turn support federated queries across multiple sites against the deidentified data in the CDM instances.

Scalability beyond the originating health systems can also be challenging given the need for additional resources for initiation and maintenance. Thus, to enhance standardization for data exchange, FHIR Sepsis Implementation Guides (IG) can be developed and balloted for general use⁶⁹. FHIR IGs constrain the standard to meet specific use cases for interoperability and are the cornerstone of FHIR interoperability. IGs contain profiles for each resource that specify which terminology value sets should be used for valid messaging of observational data between systems (e.g., which LOINC values are used to map clinical variables). Value sets leverage a change management and governance process that is a well-established part of the Value Set Authority Center (VSAC), a nationally available National Library of Medicine repository. This includes the submission, review and publishing process used to manage requests to expand or change value set content over time. The endorsed value sets can also be maintained and distributed widely through the VSAC as a national portal. Once established, FHIR profiles can easily reference VSAC value sets through a process known as terminology binding. Clinical decision

support and analytic tools can be standardized through these nationally cataloged VSACs.

In addition to these challenges, we recognize several limitations in this study. First, we only identified structured clinical variables used in prior studies and already available in the EHR to define sepsis and sepsis subgroups. Other data that are likely to be valuable to sepsis diagnosis and subgrouping such as clinician unstructured documentation, radiographic results, biomarkers or other -omic data, or pathogenic variables were not identified and included. The value of a flexible *Sepsis on FHIR* approach, however, is that such data can be added incrementally to the existing resources as they become available or central to sepsis management. Second, we recognize several limitations in using FHIR for data interoperability: (1) clinical variables were mapped to best-fit FHIR resources, (2) not all clinical data variables were mapped to a FHIR resource, and (3) not all EHR vendors are implementing the same FHIR APIs, potentially limiting expansion to other health systems. Finally, though literature summarized in the systematic review is from across the globe and selected variables mapped to FHIR standards are widely applicable, we focus on a U.S. EHR landscape. Future work to expand Sepsis on FHIR to international sites is needed.

In conclusion, we identified 151 clinical variables from 55 manuscripts needed to define sepsis and sepsis subgroups. We mapped these variables to established FHIR resource standards. This work represents the first step towards a real-time, federated, and flexible *Sepsis on FHIR* clinical data resource which can form the foundation needed for interoperability of data across healthcare systems to improve sepsis diagnostic excellence and outcomes.

METHODS

The Institutional Review Boards of Kaiser Permanente Northern California (1533936) and the University of Pittsburgh approved the study (STUDY 20020141).

Study overview

The study objective was to conduct a systematic literature review to identify a core set of variables needed to define sepsis and its subtypes, then, map these variables onto existing FHIR standards. We sought to first identify and map a core set of variables that have been historically used to define sepsis and its subtypes. However, we anticipate expansion of this flexible, foundational framework in response to the future evolution of sepsis care.

We conducted the literature search in three steps. First, we identified studies describing a consensus sepsis definition. Second, we identified studies in which a sepsis, severe sepsis, or septic shock cohort was defined, cataloging variables needed to identify these patient groups based on the multiple sepsis definitions in common use today. Finally, we identified clinical variables needed to perform detailed subtyping of sepsis patients based on currently published studies describing sepsis subtypes. When reviewing the studies, we also selected variables known to be associated with adverse outcomes for sepsis, including the source of infection (e.g., pneumonia vs urinary tract infection), the severity of organ dysfunction (e.g., presence of hypotension or altered mental status), need for organ support treatments (e.g., mechanical ventilation or dialysis) and other treatment modalities (e.g., corticosteroids). Once variables were identified, we assigned each to an existing FHIR resource, the common building blocks used to define and exchange specific EHR data. This review followed guidance published by the Cochrane collaboration and conforms to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards^{70,71}. The study protocol was not prospectively registered.

This phase represents the key first steps in the development of *Sepsis on FHIR*, a framework to accelerate the interoperability of a dynamic, and federated system of sepsis data across healthcare systems and EHRs (Supplementary Fig. 3).

Systematic review

We sought to identify relevant studies that (1) describe consensus sepsis definitions, (2) provide criteria for sepsis, severe sepsis, septic shock or a sepsis subgroup, or (3) detail sepsis subtyping. Search methods were designed following guidance published by the Cochrane collaboration⁷⁰. This systematic review conforms all these criteria.

To develop the list of search terms, we collated applicable MeSH terms provided for each major Sepsis consensus definition (i.e., Sepsis-1⁴⁰, Sepsis-2⁷², and Sepsis-3³²). These included “sepsis”, “septic shock”, “severe sepsis”. Natural language search terms included “definition”, “decision rule”, “diagnosis”, “characteristics”, “variables”, “criteria”, and “epidemiology”. Boolean operators OR and AND for combining search term(s) were used to streamline the procedure. Next, we employed the PICOS (population, intervention, comparator, outcome, and study design) search strategy, a widely cited tool for synthesizing scientific research⁷³. By using the PICOS structure, we sought to identify studies reporting on sepsis patient Populations to elucidate sepsis cohort definitions used.

We sought original research published in peer-reviewed journals reporting on adult patients with sepsis, severe sepsis, and/or septic shock. Selected study designs included consensus statements, systematic reviews, meta-analyses, randomized clinical trials, case-control studies, and cohort studies. Studies not written in English, published prior to 1970, reporting pediatric populations or animals, or “gray” literature were prospectively excluded.

We searched English language studies published from January 1, 1970 through January 31, 2020. We systematically searched three databases: PubMed, MEDLINE, and the Cochrane Library. Search strategies were developed for each database, starting with PubMed then adapted for each subsequent search (Supplementary Tables 1 and 2). We screened all titles and abstracts returned from searches against inclusion criteria and removed duplicate studies. Reference lists from selected manuscripts were scrutinized by two authors (E.B. and J.K.) to identify additional studies meeting selection criteria. Manuscripts were reviewed and ranked by two independent reviewers (E.B. and J.K.). Full-text versions were then obtained and independently reviewed by co-authors. Discrepancies throughout the review process were resolved through discussion with study senior authors (C.S. and V.L.) Initial searches were conducted June 2019 with a second search completed at the end of January 2020.

Variable selection

We collected specific study characteristics (e.g., title, authors, year of publication, journal, study design); population characteristics (e.g., demographics, admission diagnosis, level of care); clinical variables (e.g., vital signs, laboratory measurements); organ support (e.g., provision of mechanical ventilation or dialysis); treatments (e.g., corticosteroids), and outcomes (e.g., 28-day mortality, intensive care unit (ICU) admission, hospital and/or ICU length of stay) using a standardized data extraction form developed in Microsoft Word (Supplementary Table 3). After removing duplicates, the final list of variables was independently reviewed by members of the study team (E.B., J.K., V.L., and C.S.). We also convened a stakeholder group with expertise in sepsis clinical care, quality improvement, patient and family experience, EHR-based research, and healthcare interoperability to adjudicate the final list of clinical variables for completeness and applicability (Supplementary Note 1).

FHIR mapping

After compiling the list of variables, we then assigned each a FHIR resource data element to each (Supplementary Fig. 4). A resource is the building block in the FHIR standard, providing a common way to define and represent all exchangeable data⁷⁴. A FHIR resource is composed of data elements that define the content of that resource. For example, a clinical observation resource is defined by data elements such as the date of the observation, the type of observation (e.g., a vital sign, laboratory measurement), the value and units of measure. US FHIR resources are constrained to which data elements are required to exchange patient data in the US; these constraints are called US Core FHIR profiles. Example US Core FHIR profiles include vital signs, laboratory measurements, conditions or problems, medications, and other key data element types.

To share clinical data across health systems, variables from the EHR must be mapped onto existing FHIR resources to the data element level. This was done in two steps. First, each clinical variable was reviewed against FHIR documentation for a corresponding resource data element^{75–89}.

Next, each variable was fit within an existing FHIR resource deemed best-fit for the source concept.

To ensure generalizability across health systems and EHR, we first attempted to find US Core FHIR Profile mappings and if they were unavailable, we mapped data elements to the FHIR base resources.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

Aggregate data supporting the findings of this article are cataloged in Supplementary Table 5.

Received: 16 September 2021; Accepted: 24 February 2022;

Published online: 04 April 2022

REFERENCES

- Seymour, C. W. & Angus, D. C. Sepsis and septic shock. In *Harrison's Principles of Internal Medicine, 20e* (eds. Ameson, J. et al.) (McGraw-Hill, 2018).
- Rudd, K. E. et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. *Lancet* **395**, 200–211 (2020).
- Rhodes, A. et al. Surviving sepsis campaign: International guidelines for management of sepsis and septic shock: 2016. *Intensive Care Med.* **43**, 304–377 (2017).
- Angus, D. C. et al. A framework for the development and interpretation of different sepsis definitions and clinical criteria. *Crit. Care Med.* **44**, e113–e121 (2016).
- Cohen, J. V. J. et al. Sepsis: a roadmap for future research. *Lancet Infect. Dis.* **15**, 581–614 (2015).
- Seymour, C. W. et al. Derivation, validation and potential treatment implications of novel clinical phenotypes for sepsis. *J. Am. Med. Assoc.* **321**, 2003–2007 (2019).
- Seymour, C. W. et al. Time to treatment and mortality during mandated emergency care for sepsis. *N. Engl. J. Med.* **376**, 2235–2244 (2017).
- Prescott, H. C. & Iwashyna, T. J. Improving sepsis treatment by embracing diagnostic uncertainty. *Ann. Am. Thorac. Soc.* **16**, 426–429 (2019).
- Swenson, K. E. & Winslow, D. L. Impact of sepsis mandates on sepsis care: unintended consequences. *J. Infect Dis* **222**, S166–S173 (2020).
- Barbash, I. J. et al. Treatment patterns and clinical outcomes after the introduction of the medicare sepsis performance measures (SEP-1). *Ann. Intern Med.* <https://doi.org/10.7326/M20-5043> (2021).
- National Academies of Sciences, Engineering, and Medicine. Achieving Excellence in Sepsis Diagnosis: Proceedings of a Workshop—in Brief. Washington, DC: The National Academies Press. <https://doi.org/10.17226/26034> (2020).
- O'Reilly-Shah V. N. et al. The COVID-19 pandemic highlights shortcomings in US health care informatics infrastructure: a call to action. *Anesth Analg.* <https://doi.org/10.1213/ANE.0000000000004945> (2020).
- Madhavan, S. et al. Use of electronic health records to support a public health response to the COVID-19 pandemic in the United States: a perspective from 15 academic medical centers. *J. Am. Med. Inf. Assoc.* **28**, 393–401 (2021).
- Angus, D. C., Gordon, A. C. & Bauchner, H. Emerging lessons from COVID-19 for the US clinical research enterprise. *J. Am. Med. Assoc.* **325**, 1159–1161 (2021).
- Saripalle, R., Runyan, C. & Russell, M. Using HL7 FHIR to achieve interoperability in patient health record. *J. Bioinform* **94** <https://doi.org/10.1016/j.bi.2019.103188> (2019).
- Mandel, J. C., Kreda, D., Mandl, K. D., Kohane, I. S. & Ramoni, R. B. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. *J. Am. Med. Inf. Assoc.* **23**, 899–908 (2016).
- U.S. Centers for Medicare & Medicaid Services. "Policies and technology for interoperability and burden reduction". Health Informatics and Interoperability Group. CMS.gov. Retrieved from <https://www.cms.gov/Regulations-and-Guidance/Guidance/Interoperability/index#FHIR> (2021).
- Bone, R. C. et al. Sepsis syndrome: a valid clinical entity. *Crit. Care Med.* **17**, 389–393 (1989).
- Knaus, W. A., Sun, X., Nystrom, P. & Wagner, D. P. Evaluation of definitions for sepsis. *Chest* **101**, 1656–1662 (1992).
- Rangel-Frausto, M. S. et al. The natural history of the systemic inflammatory response syndrome (SIRS). *J. Am. Med. Assoc.* **273**, 117–123 (1995).
- Le Gall, J. R. et al. The logistic organ dysfunction system: a new way to assess organ dysfunction in the intensive care unit. *J. Am. Med. Assoc.* **276**, 802–810 (1996).
- Annane, D., Bellissant, E. & Cavillon, J. M. Septic shock. *Lancet* **365**, 63–78 (2005).
- Fernando, S. M. et al. Prognostic accuracy of the quick sequential organ failure assessment for mortality in patients with suspected infection. *Ann. Intern Med.* **168**, 266–275 (2018).
- Freund, Y. et al. Prognostic accuracy of sepsis-3 criteria for in-hospital mortality among patients with suspected infection presenting to the emergency department. *J. Am. Med. Assoc.* **317**, 301–308 (2017).
- Goulden, R. et al. qSOFA, SIRS, and NEWS for predicting in-hospital mortality and ICU admission in emergency admissions treated as sepsis. *Emerg. Med. J.* **35**, 345–349 (2018).
- Harrison, D. A., Welch, C. & Eddleston, J. M. The epidemiology of severe sepsis in England, Wales and Northern Ireland, 1996 to 2004: secondary analysis of a high quality clinical database, the ICNARC Case Mix Programme. *Crit. Care* **10**, 1–10 (2006).
- Haydar, S., Spanier, M., Weems, P., Wood, S. & Strout, T. Comparison of QSOFA score and SIRS criteria as screening mechanisms for emergency department sepsis. *Am. J. Emerg. Med.* **35**, 1730–1733 (2017).
- Jolley, R. J. et al. Validation and optimisation of an ICD-10 coded case definition for sepsis using administrative health data. *BMJ Open* **5**, e009487 (2015).
- Kaukonen, K. et al. Systemic inflammatory response syndrome criteria in defining severe sepsis. *N. Engl. J. Med.* **2015**, 1629–1638 (2015).
- Napolitano, L. M. Sepsis 2018: Definitions and guideline changes. *Surg. Infect* **19**, 117–125 (2018).
- Serafim, R. et al. A comparison of the quick-SOFA and systemic inflammatory response syndrome criteria for the diagnosis of sepsis and prediction of mortality. *Chest* **153**, 646–655 (2018).
- Singer, M. et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). *J. Am. Med. Assoc.* **315**, 801–810 (2016).
- Usman, O. A., Usman, A. & Ward, M. A. Comparison of SIRS, qSOFA, and NEWS for the early identification of sepsis in the emergency department. *Am. J. Emerg. Med.* **37**, 1490–1497 (2019).
- Seymour, C. W. et al. Application of a framework to assess the usefulness of alternative sepsis criteria. *Crit. Care Med.* **44**, e122–e130 (2016).
- Vincent, J. L. Definition of Sepsis and non-infectious SIRS. In *Sepsis and Non-infectious Systemic Inflammation* (eds. Cavillon, J. M. & Adrie, C.), 3–12 (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2009).
- Blanco, J. et al. Incidence, organ dysfunction and mortality in severe sepsis: a Spanish multicentre study. *Crit. Care* **12**, 1–14 (2008).
- Raith, E. P. et al. Prognostic accuracy of the SOFA score, SIRS criteria and qSOFA score for in-hospital mortality among adults with suspected infection admitted to the intensive care unit. *J. Am. Med. Assoc.* **317**, 290–300 (2017).
- Vincent, J. L. et al. The SOFA (sepsis-related organ failure assessment) score to describe organ dysfunction/failure: on behalf of the working group on sepsis-related problems of the European Society of Intensive Care Medicine. *Intensive Care Med.* **22**, 707–710 (1996).
- JL., V. Clinical sepsis and septic shock—definition, diagnosis and management principles. *Langenbecks Arch. Surg.* **393**, 817–824 (2008).
- Bone, R. C., Balk, R. & Cerra, F. B. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Chest* **101**, 1644–1655 (1992).
- Klouwenberg, P. M. C. K., Ong, D., Bonten, M. J. M. & Cremer, O. L. Classification of sepsis, severe sepsis and septic shock: the impact of minor variations in data capture and definition of SIRS criteria. *Intensive Care Med.* **38**, 811–819 (2012).
- Kaukonen, K. et al. Mortality related to severe sepsis and septic shock among critically ill patients in Australia and New Zealand, 2000–2012. *J. Am. Med. Assoc.* **311**, 1308–1316 (2014).
- Vincent, J. L., Mira, J. & Antonelli, M. Sepsis: older and newer concepts. *Lancet* **4**, 237–240 (2016).
- Whittaker, S. et al. Severe sepsis cohorts derived from claims-based strategies appear to be biased towards a more severely ill patient population. *Crit Care Med.* **41**, 945–953 (2013).
- Rhee, C. et al. Incidence and trends of sepsis in US hospitals using clinical vs claims data, 2009–2014. *J. Am. Med. Assoc.* **318**, 1241–1249 (2017).
- Fleischmann, C. et al. Assessment of global incidence and mortality of hospital-treated sepsis. Current Estimates and Limitations. *Am. J. Respir. Crit. Care Med.* **193**, 259–272 (2016).
- Iwashyna, T. 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* **52**, e39–e43 (2014).
- Martin, G. S., Mannino, D., Eaton, S. & Moss, M. The epidemiology of sepsis in the United States from 1979–200. *N. Engl. J. Med.* **348**, 1546–1554 (2008).
- Trzeciak, S. et al. Inclusion criteria for clinical trials in sepsis. *Chest* **127**, 242–245 (2005).
- Levy, M. M. et al. 2001 SCCM/ESICM/ACCP/ATS/SOS international sepsis definitions conference. *Intensive Care Med.* **29**, 530–538 (2003).

51. Dellinger, R. P. et al. Surviving sepsis campaign: International guidelines for management of severe sepsis and septic shock: 2012. *Crit. Care Med.* **41**, 580–637 (2013).
52. Shankar-Hari, M. et al. Developing a new definition and assessing new clinical criteria for septic shock: for the third international consensus definitions for sepsis and septic shock (Sepsis-3). *J. Am. Med. Assoc.* **315**, 775–787 (2016).
53. Nguyen, H. B. & Smith, D. Sepsis in the 21st century: recent definitions and therapeutic advances. *Am. J. Emerg. Med.* **25**, 564–571 (2007).
54. Orgeta, R. N., Rosin, C., Bingisser, R. & Nickel, C. H. Clinical scores and formal triage for screening of sepsis and adverse outcomes on arrival in an emergency department all-comer cohort. *J. Emerg. Med.* **57**, 460 (2019).
55. Churpek, M. M. et al. Quick sepsis-related organ failure assessment, systemic inflammatory response syndrome, and early warning scores for detecting clinical deterioration in infected patients outside the intensive care unit. *Am. J. Respir. Crit. Care Med.* **195**, 906–911 (2016).
56. Calandra, T. & Cohen, J. The international sepsis forum consensus conference on definitions of infection in the intensive care unit. *Crit. Care Med.* **33**, 1538–1548 (2005).
57. Levy, M. M., Evans, L. E. & Rhodes, A. The surviving sepsis campaign bundle: 2018 update. *Intensive Care Med.* **44**, 925–928 (2018).
58. NAGMSC Working Group on Sepsis. *Final Report*, <https://www.nigms.nih.gov/News/reports/Documents/nagmsc-working-group-on-sepsis-final-report.pdf> (2019).
59. The Office of the National Coordinator for Health Information Technology. “Health IT Quick Stats”. Retrieved from <https://dashboard.healthit.gov/quickstats/quickstats.php> (2021).
60. Barbash, I. J., Kahn, J. & Thompson, B. T. Opening the debate on the new sepsis definition. Medicare’s sepsis reporting program: two steps forward, one step back. *Am. J. Respir. Crit. Care Med.* **194**, 139–141 (2016).
61. Afshar, M. et al. Patient outcomes and cost-effectiveness of a sepsis care quality improvement program in a health system. *Crit. Care Med.* **47**, 1371–1379 (2019).
62. Barbash, I. J. et al. Hospital perceptions of Medicare’s sepsis quality reporting initiative. *J. Hosp. Med.* **12**, 963–968 (2017).
63. Prescott, H. C., Calfee, C., Thompson, B. T., Angus, D. C. & Liu, V. X. Toward smarter lumping and smarter splitting: rethinking strategies for sepsis and acute respiratory distress syndrome clinical trial design. *Am. J. Respir. Crit. Care Med.* **194**, 147–155 (2016).
64. Ma, P. et al. Individualized resuscitation strategy for septic shock formalized by finite mixture modeling and dynamic treatment regimen. *Crit. Care* **25**, 1–16 (2021).
65. Demerle, K. M. et al. Feasibility of embedding a scalable, virtually enabled bio-repository in the electronic health record for precision medicine. *JAMA Netw Open* **4**, e2037739 (2021).
66. Gianfresco, M. A., Tamang, S., Yazdany, J. & Schmajuk, G. Potential biases in machine learning algorithms using electronic health record. *JAMA Intern Med.* **178**, 1544–1547 (2018).
67. U.S. Department of Health and Human Services. Public Health Data Modernization: Listening session on real-world testing of 21st century cures act requirements, Virtual Meeting. (2020).
68. *Common Data Model (CDM) Specification, Version 6.0*, https://pcornt.org/wp-content/uploads/2020/12/PCORnet-Common-Data-Model-v60-2020_10_221.pdf (2021).
69. HL7 International. “Resource Implementation Guide”. Conformance. Retrieved from <http://hl7.org/fhir/implementationguide.html> (2011).
70. Higgins, J. P. T. et al. (eds). *Cochrane Handbook for Systematic Reviews of Interventions* 2nd edn. (John Wiley & Sons, Chichester (UK), 2019).
71. Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. G., The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* **6**, e1000097 (2009).
72. Levy, M. M. et al. 2001 SCCM/ESICM/ACCP/ATS/SIS international sepsis definitions conference. *Intensive Care Med.* **29**, 530–538 (2003).
73. Schardt, C., Adams, M. D., Owens, T., Keitz, S. & Fontelo, P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med. Inform. Decis Mak.* **7**, 1–6 (2007).
74. Lehne, M., Luijten, S., Vom Felde Genannt Imbusch, P. & Thun, S. The use of FHIR in digital health—a review of the scientific literature. *Stud. Health Technol. Inf.* **3**, 52–58 (2019).
75. International, H. L. “Welcome to FHIR”. HL7 FHIR Release 4. <https://www.hl7.org/fhir/>. Accessed October, 2020 (2016).
76. Angus, D. C. et al. Epidemiology of severe sepsis in the United States. *Crit. Care Med.* **29**, 1303–1310 (2001).
77. Cecconi, M., Evans, L., Levy, M. & Rhodes, A. Sepsis and septic shock. *Lancet* **392**, 75–87 (2018).
78. Danai, P. A. et al. Seasonal variation in the epidemiology of sepsis. *Crit. Care Med.* **35**, 410–415 (2007).
79. Dombrovsky, V. Y. et al. Rapid increase in hospitalization and mortality rates for severe sepsis in the United States: a trend analysis from 1993 to 2003. *Crit. Care Med.* **35**, 1244–1250 (2007).
80. Gaieski, D. F. et al. The relationship between hospital volume and mortality in severe sepsis. *Amer. J. Respir. Crit. Care Med.* **190**, 665–674 (2014).
81. Gupta, S. et al. Culture-negative severe sepsis: nationwide trends and outcomes. *Chest* **150**, 1251–1259 (2016).
82. Iwashyna, T. J. et al. Long-term cognitive impairment and functional disability among survivors of severe sepsis. *J. Am. Med. Assoc.* **304**, 1787–1794 (2010).
83. Kadri, S. S. et al. Estimating ten-year trends in septic shock incidence and mortality in United States academic medical centers using clinical data. *Chest* **151**, 278–285 (2017).
84. Lever, A. & Mackenzie, I. Sepsis: definition, epidemiology and diagnosis. *BMJ* **335**, 879–883 (2007).
85. Matot, I. & Sprung, C. Definition of sepsis. *Intensive Care Med.* **27**, S3–S9 (2001).
86. McPherson, D. et al. Sepsis-associated mortality in England: an analysis of multiple cause of death data from 2001 to 2010. *BMJ Open* **3**, e002586 (2013).
87. Seymour, C. W. et al. Assessment of clinical criteria for sepsis for the Third International Consensus definitions for sepsis and septic shock (Sepsis-3). *J. Am. Med. Assoc.* **315**, 762–774 (2016).
88. Vincent, J. L. et al. Sepsis in European intensive care units: results of the SOAP study. *Crit. Care Med.* **34**, 344–353 (2006).
89. Zhao, H. et al. An evaluation of the diagnostic accuracy of the 1991 American College of Chest Physicians/Society of Critical Care Medicine and the 2001 Society of Critical Care Medicine/European Society of Intensive Care Medicine/American College of Chest Physicians/American Thoracic Society/Surgical Infection Society sepsis definition. *Crit. Care Med.* **40**, 1700–1706 (2012).

ACKNOWLEDGEMENTS

This work was supported in part by the Gordon and Betty Moore Foundation. Additional funding support through the National Institutes of Health was provided for Dr. Brant (2-T32-HL007820-21), Dr. Seymour (K23GM104022), Dr. Liu (R35GM128672), and Mr. Kennedy (R35GM119519). We thank the significant contribution of the patients, families, researchers, clinical staff, and sponsors to the data included herein.

AUTHOR CONTRIBUTIONS

E.B. performed the literature review, data analysis, and drafted/edited the manuscript. J.K. assisted with the literature review, data analysis, drafted figures, and edited the manuscript. A.K., L.G., P.M., and D.S. assisted with FHIR mapping and edited the manuscript. J.S. performed FHIR mapping and assisted with manuscript drafting/editing. G.J.E., D.A., C.S., and V.L. conceived and planned the literature review and analysis, and assisted with manuscript drafting/editing.

COMPETING INTERESTS

Dr. Angus reports receiving personal fees from and serving as a consultant to Ferring Pharmaceuticals, Bristol-Myers Squibb, Bayer AG, and Beckman Coulter Inc; owning stock in Alung Technologies; and having patent applications pending for selepressin (compounds, compositions, and methods for treating sepsis) and proteomic biomarkers of sepsis in elderly patients. Dr. Seymour reports receiving personal fees from Edwards Inc. and Beckman Coulter Inc. No other conflicts of interest were reported.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-022-00580-2>.

Correspondence and requests for materials should be addressed to Emily B. Brant.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022