Published in final edited form as:

Crit Care Med. 2019 January ; 47(1): 135–137. doi:10.1097/CCM.0000000000003485.

Exploring the pathways revealed by international sepsis benchmarking

Allan J. Walkey, MD, MSc

¹Department of Medicine, Division of Pulmonary, Allergy and Critical Care, Center for Implementation and Improvement Sciences, Boston University School of Medicine, Boston, Massachusetts, USA.

²Department of Health Law, Policy and Management, Boston University School of Public Health, Boston, Massachusetts, USA.

Benchmarking clinical practice involves comparison of local outcomes to a reference standard. In order to accurately benchmark, groups need to: 1) harmonize data, and 2) produce unbiased comparisons of processes and/or outcomes of interest.(1,2) The overarching purpose of most benchmarking is to identify high and low performers, discover the characteristics responsible for high and low performance, and make changes targeting modifiable characteristics intended to improve performance. Thus, benchmarking is an essential part of a learning healthcare system of continuous improvement and innovation.(3)

Recently sepsis has been thrust onto the world stage by a recent WHO resolution intended to improve its prevention, diagnosis, and management.(4) As part of the resolution, the 70th annual World Health Assembly adopted multiple recommendations relevant to sepsis benchmarking, including urging member states to “*apply and improve the use of International Classification of Diseases system to establish the prevalence and profile of sepsis...*”, to “*develop and implement monitoring and evaluation tools in order to focus attention on and monitor progress towards improving outcomes from sepsis, including the development and fostering of specific epidemiologic surveillance systems.*” and to “*develop and implement standard and optimal care...*”. However, prior efforts to benchmark international sepsis incidence and outcomes have been inadequate. A systematic review of studies investigating the global incidence and outcomes of sepsis found that population-based estimates for sepsis outcomes were generally lacking in low- and middle income countries, and exhibited wide variability due to differences in ICD-9 coding, infectious causes, and geographic factors leading to different admission criteria of patients into the hospital or ICU.(5)

The study by Ranzani and colleagues in this issue of *Critical Care Medicine* further explores the complexities of international sepsis benchmarking, seeking to compare outcomes of patients with sepsis hospitalized in intensive care units in Brazil and England.(6) In line with

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.

Conflicts of Interest: None.

Copyright form disclosure: Dr. Walkey has disclosed that he does not have any potential conflicts of interest.

step 1 of the benchmarking process, the authors harmonized data across two nationwide ICU databases. Although others have used administrative data,(7) surveys(8) or prospective cross-sectional studies(9, 10) to make international comparisons in critical care delivery across different conditions, harmonization of pre-existing health record data generated from the routine delivery of sepsis care is relatively unique. Data from Brazil was extracted from a prior, retrospective, multicenter ORCHESTRA study of critical care organization in 78 ICUs across 11 Brazilian states, whereas data from England came from the ongoing Intensive Care National Audit & Research Centre Case Mix Programme (ICNARC) Database of 164 ICUs. The harmonization process involved identifying variables of interest, achieving consensus on operational definitions of target variables in each dataset, anonymization of the individual datasets, and finally, merging of the harmonized data.

After the considerable work of data harmonization, hospital mortality rates were compared between England and Brazil in staged analyses adjusting for increasing number of demographic and clinical variables. Ranzani and colleagues identified that multiple factors differed between Brazilian and English ICU patients with sepsis, including age, comorbidity burdens, source of ICU admission (e.g., hospital ward or emergency department), site of infection, and type and number of acute organ dysfunctions. While unadjusted hospital mortality rates did not differ between Brazil and England, adjustment for large, baseline differences in patient characteristics resulted in similarly large changes in comparative mortality estimates. For example, adjustment for demographic and pre-existing comorbidity differences showed lower mortality rates in Brazil, whereas adjustment for markers of acute organ dysfunction and site of infection showed lower mortality in England.

The findings of Ranzani et al show the multiple technical and analytical difficulties of benchmarking international sepsis outcomes. First, clinical databases with the level of detail contained in ICNARC and ORCHESTRA are unlikely to be widely available, especially in low or middle income countries. Second, the substantial work required to accurately harmonize definitions across “only” 2 different data sets may not be realistic across dozens of countries. Harmonizing sepsis definitions is particularly prone to bias: prior studies have shown that small changes in sepsis definitions (measurement bias) cause large swings in risk-adjusted sepsis mortality estimates across healthcare systems.(11) Thus, centralized registries that prospectively collect, harmonize, and process representative samples of patients with sepsis across member countries are likely to be necessary for gathering reliable data needed for international benchmarking.

However, data collection and harmonization is not the biggest challenge to benchmarking sepsis outcomes. The large swings in national mortality estimates observed by Ranzani et al. across different statistical models raises a difficult question: which model was right? To start to answer this question, we need to understand that the process of benchmarking is a type of causal inference analysis subject to the potential biases of any observational research study. (12) In benchmarking, we are asking: what is the causal association between country and sepsis outcome? Once we understand the pathways between country, healthcare characteristics, and outcomes, we can then seek to identify modifiable and unmodifiable differences in patient characteristics, structures, and/or processes of care that best explain

international differences in sepsis outcomes in order to target modifiable explanatory factors for improvement efforts.

The causal pathways between country and sepsis outcome likely do not involve confounders in the traditional sense, but rather, involve mediators (Figure 1) between country and outcome. Rather than adjusting away biased associations due to confounding, when we adjust for variables in international benchmarking studies we are more likely exploring the extent of direct and indirect mediation of causal pathways(13) between country and outcome. If large mortality differences between well-performing and poorly-performing countries are ‘explained away’ by differences in variables such as comorbidity burden, infection sources, number of ICU beds, or pre-hospital paramedic services, then targeting these factors may be the optimal approach to improve parity in sepsis outcomes.

One additional aspect of the study by Ranzani and colleagues warrants particular attention: the inclusion of only patients admitted to ICUs. ICU admission criteria and bed availability differs markedly between countries.(7) The selection of the cohort with sepsis from a likely mediator of international outcome differences introduces additional complexities to analysis (Figure 1). First, selecting only ICU patients does not allow exploration of ICU admission criteria or ICU structural characteristics as potential modifiable mediators of international outcome differences. Second, selecting a cohort that is also a mediator introduces additional confounding based upon ‘collider’ effects between unmeasured ICU admission criteria and mortality, further complicating modeling.(14) Thus, future efforts to benchmark sepsis mortality should refrain from using ICU-only cohorts.

How should we move forward with international benchmarking of sepsis outcomes? Ranzani et al have illustrated many of the potential difficulties in sepsis benchmarking. They have shown that rigorous data harmonization and organized analysis of different models are necessary steps to extract actionable information from international sepsis outcome benchmarking. Unadjusted models in international benchmarking will be important to identify outcome differences between countries that can then serve as opportunities to identify actionable improvements through rigorous mediation analyses. While international sepsis benchmarking will entail considerable work, it also represents a pathway towards novel discoveries to improve the outcomes of patients with sepsis.

Acknowledgments

AJW received support from NIH/NHLBI R01HL139751 and R01 HL136660.

References

1. Benson HR. An introduction to benchmarking in healthcare. *Radiol. Manage* 1994;16:35–39.
2. Mosel D, Gift B. Collaborative benchmarking in health care. *Jt. Comm. J. Qual. Improv* 1994;20:239–249. [PubMed: 8044219]
3. Anon. Engineering a Learning Healthcare System: A Look at the Future. Workshop Summary. 3, Healthcare System Complexities, Impediments, and Failures. *Inst. Med. Natl. Acad. Eng. Roundtable Value Sci. Heal. Care*. 2011.
4. Global Sepsis Alliance. WHA adopts resolution on sepsis. 2017.

5. Fleischmann C, Scherag A, Adhikari NKJ, et al. Assessment of Global Incidence and Mortality of Hospital-treated Sepsis. Current Estimates and Limitations. *Am. J. Respir. Crit. Care Med* 2016;193:259–272. [PubMed: 26414292]
6. Ranzani OT; Shankar-Hari M; Harrison DA; Rabello LS; Salluh JIF; Rowan KM; Soares M A comparison of mortality from sepsis (Sepsis-3) in Brazil and England: the impact of heterogeneity in general and sepsis-specific patient characteristics. *Crit. Care Med* 2018.
7. Wunsch H, Angus DC, Harrison DA, et al. Variation in critical care services across North America and Western Europe. *Crit. Care Med* 2008;36:2787–93, e1–9. [PubMed: 18766102]
8. Burns KEA, Raptis S, Nisenbaum R, et al. International Practice Variation in Weaning Critically Ill Adults from Invasive Mechanical Ventilation. *Ann. Am. Thorac. Soc* 2018;15:494–502. [PubMed: 29509509]
9. Laffey JG, Madotto F, Bellani G, et al. Geo-economic variations in epidemiology, patterns of care, and outcomes in patients with acute respiratory distress syndrome: insights from the LUNG SAFE prospective cohort study. *Lancet. Respir. Med* 2017;5:627–638. [PubMed: 28624388]
10. Vincent JL, Sakr Y, Sprung CL, et al. Sepsis in European intensive care units: results of the SOAP study. *Crit. Care Med* 2006;34:344–353. [PubMed: 16424713]
11. Walkey AJ, Shieh M-S, Liu VX, Lindenauer PK. Mortality Measures to Profile Hospital Performance for Patients With Septic Shock. *Crit. Care Med* 2018;46:1247–1254. [PubMed: 29727371]
12. Gershon AS, Jafarzadeh SR, Wilson KC, Walkey AJ. Clinical Knowledge from Observational Studies: Everything You Wanted to Know but Were Afraid to Ask. *Am. J. Respir. Crit. Care Med* 2018.
13. VanderWeele TJ. Policy-relevant proportions for direct effects. *Epidemiology* 2013;24:175–176. [PubMed: 23232624]
14. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. *Int. J. Epidemiol* 2013;42:1511–1519. [PubMed: 24019424]

Figure 1.

Directed acyclic graph demonstrating the potential relationships between country and sepsis mortality, including both observed (black lines and circles) and unobserved (grey lines and circles) mediating variables. Introducing different potential mediators, such as specific patient or healthcare system characteristics, can help assess modifiable factors to improve sepsis outcome performance. However, selection of the cohort within a category of a potential mediator (e.g., ICU admission) has the potential to introduce further confounding by unmeasured variables. Figure was constructed with dagitty.net, see Johannes Textor, Juliane Hardt, and Sven Knuppel. Dagitty: A graphical tool for analyzing causal diagrams. *Epidemiology*, 22(5):745, 2011.