

Existing data analysis in pediatric critical care research

Tellen D. Bennett^{1*}, Michael C. Spaeder², Renée I. Matos³, R. Scott Watson⁴, Katri V. Typpo⁵, Robinder G. Khemani⁶, Sheri Crow⁷, Brian D. Benneyworth⁸, Ravi R. Thiagarajan⁹, J. Michael Dean¹⁰, Barry P. Markovitz⁶ and for the Pediatric Acute Lung Injury and Sepsis Investigators (PALISI)

¹ Pediatric Critical Care, University of Colorado School of Medicine, Aurora, CO, USA

² Critical Care Medicine, Children's National Medical Center, Washington, DC, USA

³ Pediatric Critical Care Medicine, San Antonio Military Medical Center, United States Air Force, San Antonio, TX, USA

⁴ CRISMA Center and Pediatric Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

⁵ Pediatric Critical Care, University of Arizona College of Medicine, Tucson, AZ, USA

⁶ Anesthesiology and Critical Care Medicine, Children's Hospital Los Angeles, University of Southern California Keck School of Medicine, Los Angeles, CA, USA

⁷ Pediatric Critical Care Medicine, Mayo Clinic, Rochester, MN, USA

⁸ Pediatric Critical Care Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

⁹ Pediatric Critical Care Medicine, Boston Children's Hospital, Boston, MA, USA

¹⁰ Pediatric Critical Care, University of Utah School of Medicine, Salt Lake City, UT, USA

Edited by:

John McGuire, University of Washington School of Medicine, USA

Reviewed by:

John McGuire, University of Washington School of Medicine, USA
Paul Checchia, Texas Children's Hospital and Baylor College of Medicine, USA

*Correspondence:

Tellen D. Bennett, Pediatric Critical Care, University of Colorado School of Medicine, 13199 E Montview Blvd, Suite 300, Campus Mail F443, Aurora, CO 80045, USA
e-mail: tell.bennett@ucdenver.edu

Our objectives were to review and categorize the existing data sources that are important to pediatric critical care medicine (PCCM) investigators and the types of questions that have been or could be studied with each data source. We conducted a narrative review of the medical literature, categorized the data sources available to PCCM investigators, and created an online data source registry. We found that many data sources are available for research in PCCM. To date, PCCM investigators have most often relied on pediatric critical care registries and treatment- or disease-specific registries. The available data sources vary widely in the level of clinical detail and the types of questions they can reliably answer. Linkage of data sources can expand the types of questions that a data source can be used to study. Careful matching of the scientific question to the best available data source or linked data sources is necessary. In addition, rigorous application of the best available analysis techniques and reporting consistent with observational research standards will maximize the quality of research using existing data in PCCM.

Keywords: pediatrics, intensive care, outcomes research, epidemiology, health services

INTRODUCTION

“Big data” is the nickname in computer science, business, and public policy for the application of sophisticated analytic techniques to large and rapidly growing databases (1–3). Medicine has been lauded for its early adoption of data-driven “evidence-based” decision-making, but has been noted to be lagging behind other industries in leveraging the rich data available in

electronic health records, registries, and enriched administrative databases (1, 3, 4).

Secondary use of existing data is an attractive option for disease epidemiology, quality and safety questions, health services research, economic analyses, comparative effectiveness research, and implementation and dissemination science. Existing data often describe “real-world” care and may be used to define current practice variation, to analyze natural experiments such as policy changes, and to estimate available sample sizes for prospective studies. Existing data may be used to conduct studies that are not amenable to a randomized trial format (5), for example in areas with limited equipoise: published guidelines with incomplete evidence, persistent variation, or controversy. These potential benefits are balanced by the data quality limitations of many existing data sources and the “numerous examples of poorly designed studies utilizing datasets ill equipped to answer the research questions posed of them” (5, 6).

Relatively few children require critical care (7). Overall, each pediatric intensive care unit (PICU) cares for a small number of heterogeneous patients with relatively rare diseases. Care has improved such that mortality is rare, but the risk of significant morbidity is high (7, 8). This distribution of patients and outcomes has made clinical research in pediatric critical care logistically challenging and expensive because appropriately precise estimates

Abbreviations: AHA-GWTG-R, American Heart Association Get With The Guidelines Resuscitation; ANZPIC, Australia New Zealand Paediatric Intensive Care; CPCCRN, Collaborative Pediatric Critical Care Research Network; DoDTR, Department of Defense Trauma Registry; ECMO, extracorporeal membrane oxygenation; ELSO, Extracorporeal Life Support Organization; HCUP, Healthcare Cost and Utilization Project; KID, Kids' Inpatient Database; MAX, Medicaid Analytic eXtract; MIMIC, Multiparameter Intelligent Monitoring in Intensive Care; NEDS, Nationwide Emergency Department Sample; NIH, National Institutes of Health; NIS, National Inpatient Sample; PCCM, pediatric critical care medicine; PECARN, Pediatric Emergency Care Applied Research Network; PELOD, pediatric logistic organ dysfunction; PHIS, pediatric health information systems; PICANet, pediatric intensive care audit network; PICU, pediatric intensive care unit; PIM, pediatric index of mortality; PRISM, pediatric risk of mortality; REP, Rochester Epidemiology Project; SED, State Emergency Department; SID, State Inpatient Database; STROBE, Strengthening the Reporting of Observational Studies in Epidemiology; STS-CHSD, Society of Thoracic Surgeons Congenital Heart Surgery Database; VPS, Virtual PICU Systems, Inc.

of effect require data from many centers (9). Despite these challenges, clinicians and researchers in pediatric critical care medicine (PCCM) have the potential to decrease a lifetime of disease burden for their patients.

Pediatric critical care medicine research differs from adult critical care research in that no dominant claims database analogous to Medicare exists; pediatric patients are usually reimbursed via a mixture of private payers and state-based Medicaid systems that are not uniformly reported. Large, multi-center existing data sources and linkage of multiple data sources may provide solutions to both challenges in PCCM research: the small sample size of any one patient type at each institution and the lack of a dominant claims database.

The objectives of this paper are to review and categorize the existing data sources that are important to PCCM investigators and the types of questions that have been or could be studied with each data source. Our goal is to provide PCCM investigators with resources to assist them in matching a research question with the most appropriate available data.

DATA SOURCES FOR PEDIATRIC CRITICAL CARE RESEARCH

Choosing a data source for an analysis begins with carefully assessing the strengths and limitations of each data source. Investigators evaluating data source quality may benefit from using a tool that Black and Payne (10) developed and Cooke and Iwashyna (6) adapted for use with adult critical care data sources. That schema evaluates databases based on coverage (representativeness, completeness of recruitment, variables included, and amount of missing data) and accuracy (raw data collection, explicit variable definitions and rules, reliability of coding, independence of observations, and data validation).

Matching the level of clinical detail in the data source to the research question is also very important (Table 1). Evaluating causal relationships or conducting comparative effectiveness studies requires a high level of clinical detail to allow accurate adjustment for confounding by indication, severity of illness, and other factors (6). Identifying risk factors for an outcome may require only a moderate level of clinical detail, and descriptive epidemiologic studies or policy evaluations may require only a low level of clinical detail.

The authors of this manuscript met in March, 2013 to discuss this topic. At that time, we developed a preliminary list of data

sources considered important to PCCM that at least one author was familiar with. We also developed a preliminary method to categorize data sources. These were refined over the next several months. The data source types are arranged below from generally more clinical details to less, with example data sources for each type (Table 2). Examples of questions that have been answered or could be answered with each data source type are also provided. Because several of the authors are members or users of the data sources described, the initials of the primary authors for each section are noted.

PUBLIC USE DATASETS FROM FUNDED STUDIES AND NETWORKS (TELLEN D. BENNETT/J. MICHAEL DEAN)

Examples: public use datasets from the Collaborative Pediatric Critical Care Research Network (CPCCRN) and the Pediatric Emergency Care Applied Research Network (PECARN)

National Institutes of Health (NIH)-funded studies and networks are now required to produce and make available a public use dataset after study completion. The scope of these datasets may be narrow in keeping with the study population, but the level of clinical detail is often very high. Collaborative Pediatric Critical Care Research Network (CPCCRN) and Pediatric Emergency Care Applied Research Network (PECARN) each have several public use datasets. Datasets from individual funded studies and from other networks such as the Canadian Critical Care Trials Group may be available by directly contacting the primary investigators. Datasets that are known to be available include studies of prone positioning in acute lung injury (11), activated protein C (12), and restrictive transfusion thresholds (13). Farris et al., for example, recently published a study of functional outcomes in children with severe sepsis using a dataset from the activated protein C trial (14).

PEDIATRIC CRITICAL CARE REGISTRIES: UNITED STATES (ROBINDER G. KHEMANI/BARRY P. MARKOVITZ)

Example: Virtual PICU systems (VPS), LLC

Although the main aim of Virtual PICU systems (VPS) is to provide comparative data for benchmarking and quality improvement, it is structured as a PCCM registry. It includes nearly 120 pediatric and pediatric cardiac ICUs from 100 participating sites, including some outside North America. One particularly useful aspect of VPS is that it contains severity of illness scores including Pediatric Risk of Mortality (PRISM) III, Pediatric Index of Mortality (PIM) 2, Pediatric Logistic Organ Dysfunction (PELOD), and several cardiac intensive care unit complexity scores.

All institutions report a minimal dataset of required elements. Additional, non-mandatory data are available for a large proportion of cases; each institution decides whether to report each class of non-mandatory data. This allows inference about specific procedures or diagnoses at the patient level if analysis is performed accounting for the institutional profile. Investigators have used VPS to answer questions regarding quality improvement and severity of illness in PICUs (15–17), and to analyze risk factors for outcomes in several specific cohorts of patients (18–20).

Table 1 | Level of clinical detail in existing data sources.

Level of clinical detail	Data source contents
High	Many clinical variables such as vital signs, physiologic data, laboratory results, or severity of illness scores
Moderate	Some clinical variables or utilization data (medications, imaging studies, etc.) or charge/cost information
Low	Data limited to standard administrative fields such as demographics, length of stay, disposition, and diagnosis and procedure codes

Table 2 | Data source types in pediatric critical care research.

Type of data	Clinical detail	Example data sources	Represented population	Accessibility/cost	Notes
Public use datasets from funded studies and networks	High	BioLINCC, CPCCRN, PECARN	Study-specific or network-specific	High/free	Often available online (e.g., www.pecarn.org and www.cpccrn.org), requires data use agreement and IRB approval. May have limited data support
Pediatric critical care registries	High	Virtual PICU systems (VPS)	119 participating hospitals (fee) with ~600,000 PICU cases	Moderate/free to VPS members	Use by non-members requires partnership with a member hospital investigator. Requires review by the Research Committee, which is primarily intended to ensure that multiple investigators are not attempting to answer the same question
	High	ANZPIC registry	24 PICUs in Australia and New Zealand	Moderate/free	
	High	PICANet registry	32 PICUs in the United Kingdom and Ireland	High/free	May be merged with continental European PICU registries in the future
Therapy-specific registries	High	Society of thoracic surgeons – STS congenital database	Children who have undergone cardiac surgery at participating centers	Moderate/low	Cardiac anesthesia-specific data was introduced in 2010. Linked to PHIS
	Moderate	ECMO registry of the Extracorporeal Life Support Organization	230 voluntarily contributing centers. Internationally representative sample of ECMO utilization	Moderate/free	Available to member centers, special requests may be made to the ECMO Registry steering committee. Minimal ability to risk adjust, but plans in place to improve this in 2013. Also contains neonatal and adult ECMO runs. Interpretation of outcome and complication data should be done with care
Disease-specific registries	High	Department of Defense Trauma Registry	U.S. military, coalition soldiers, and civilian trauma patients	Moderate/free	Department of Defense only. Use requires partnership with a military investigator
	Moderate	American Heart Association Get With the Guidelines – Resuscitation	>400 voluntarily contributing hospitals (hospitals pay a fee to participate as a quality improvement initiative)	Moderate/free	Use requires approval of research request by AHA-GWTG-R Research Task Force
	Moderate	National Trauma Data Bank	> 700 voluntarily contributing hospitals and > 100,000 pediatric trauma admissions each year	High/\$300 per year	Obtained from the American College of Surgeons. A nationally representative sample of adults treated at Levels I and II facilities is also available for purchase
Population-based registries	High	Rochester Epidemiology Project (REP)	All residents of Olmsted County, MN from January 1, 1966 to the present, with ~500,000 individuals and ~1.2 million records	Moderate/free	Unique population-based resource. Use requires permission from the REP obtained through online application
Quality improvement or benchmarking databases	Moderate	Pediatric Health Information Systems (PHIS)	44 free-standing children's hospitals, >7 million inpatient cases and 20 million Emergency Department encounters	High/free to member hospitals	No physiologic variables. Resource utilization and charge data are detailed, but results of tests and studies are not currently widely available. Several linkages completed or planned

(Continued)

Table 2 | Continued

Type of data	Clinical detail	Example data sources	Represented population	Accessibility/cost	Notes
Claims databases	Moderate	State medicaid files	Data available from 1999 to present for all 50 states and D.C.	High/~\$1,000–1,500 per year, per state	Limited use to date in PCCM research
Government administrative databases	Low	Healthcare Cost and Utilization Project (HCUP) databases (details below)			Prices for HCUP products frequently discounted for students
		Kids inpatient database (KID)	Every 3 years: 1997, 2000, 2003, 2006, 2009	High/\$200–350 per year	Allows national-level estimates of pediatric conditions. Sample weighting requires analytic adjustment. Two to three million hospital discharges in each file
		National Inpatient Sample (NIS)	Annual ~20% stratified sample of hospital discharges, ~1,000 hospitals per year	High/\$160–350 per year	Sample weighting requires analytic adjustment
		National emergency department sample (NEDS)	Annual ~20% stratified sample of ED visits in 28 states, 2006–2010	High/\$500 per year	Linked to state inpatient databases to determine ED outcomes. Sample weighting requires analytic adjustment
		State inpatient databases (SID)	All inpatient discharge abstracts in participating states	High/~\$35–3,000 per year, per state	Component files of the NIS and KID
State emergency department databases (SEDD) files	All ED visits that do not result in admission, for each participating state	High/~\$35–3,200 per year, per state	Component files of the NEDS. Information about patients seen in an ED and admitted is found in the corresponding SID		

Level of clinical detail: high = includes many clinical variables such as vital signs, laboratory results, or severity of illness scores. Moderate = includes some clinical variables or utilization data (medications, imaging studies, etc.) or charge/cost information. Low = data limited to standard administrative content such as demographics, length of stay, disposition, and diagnosis and procedure codes.

PEDIATRIC CRITICAL CARE REGISTRIES OUTSIDE THE UNITED STATES (MICHAEL C. SPAEDER)

Examples: the Australia New Zealand Paediatric Intensive Care (ANZPIC) Registry and the Paediatric Intensive Care Audit Network (PICANet)

A number of national and multi-national pediatric critical care databases and registries are maintained worldwide. Similar to data sources in the United States like VPS, the primary purpose is benchmarking among institutions. Recently, these data sources have been increasingly used for clinical research. The Australia New Zealand Paediatric Intensive Care (ANZPIC) Registry includes data from 24 PICUs in Australia and New Zealand. Publications include investigations of outcomes related to acute lung injury (21) and hyperglycemia (22), as well as inter-unit practice variation in duration of respiratory support (23) and length of stay (24).

Similar in content to ANZPIC, Paediatric Intensive Care Audit Network (PICANet) maintains data from 32 PICUs in the United Kingdom and Ireland. Investigations employing PICANet include studies of physiologic associations in the post-cardiac arrest population (25), utilization of palliative care services following PICU

discharge (26), acute disseminated encephalomyelitis (27), and diabetes (28). A collaboration to create a standardized European pediatric critical care dataset is currently underway between PICANet and the owners of databases in the Netherlands, Italy, and Portugal.

THERAPY-SPECIFIC REGISTRIES (SHERI CROW: STS-CHSD AND KATRI V. TYPPO/RAVI R. THIAGARAJAN: ELSO)

Examples: the Society of Thoracic Surgeons Congenital Heart Surgery Database (STS-CHSD) and the Extracorporeal Membrane Oxygenation (ECMO) Registry of the Extracorporeal Life Support Organization (ELSO)

The Society of Thoracic Surgeons Congenital Heart Surgery Database (STS-CHSD) was developed in 1994 (29) and now contains data from 108 U.S. centers, representing 86% of the 125 U.S. pediatric cardiac surgical programs, and three of the eight centers in Canada (30) [personal communication to Sheri Crow from Jeff Jacobs, and Marshall Jacobs]. Participating centers submit data about congenital heart surgery procedures including patient risk factors, surgical complexity scoring, operative techniques, care processes, and clinical outcomes. The data are available

to participating hospitals, physicians, and the healthcare industry for benchmarking, quality improvement, and research.

The STS-CHSD data specifications are upgraded every 3 years. The 2010 upgrade included new fields facilitating improved long-term outcome assessment and linkage to other databases for pediatric cardiology and critical care. The STS-CHSD now includes data from 36 centers regarding the anesthetic techniques (31) used for congenital cardiac surgical procedures. The STS-CHSD has been used to study delayed sternal closure (32), perioperative corticosteroid use (33), pediatric cardiac surgical case volume (34), and gender and race effects on surgical outcomes (35).

The Extracorporeal Membrane Oxygenation (ECMO) registry of Extracorporeal Life Support Organization (ELSO) captures information on ECMO use for any indication. The registry was started in 1984 but contains data from patients supported with ECMO since 1976. Approximately 230 U.S. and international ELSO members submit data. Data elements include diagnosis and procedural information, pre-ECMO level of illness and support, indication for ECMO [pulmonary, cardiac, or to support cardiopulmonary resuscitation (ECPR)], ECMO equipment used, ECMO flow, duration of ECMO, and adverse events during ECMO. Outcome information collected includes survival to discharge, discharge disposition, and reason for death for non-survivors.

Each submission pertains to an individual ECMO run, with multiple ECMO runs possible for a given patient. Each patient has a unique identifier that can support patient-level analyses of outcomes, although a variable for ECMO center is not available. The database is currently being redesigned to improve data quality and to expand severity of illness information (36). Examples of analyses of ELSO data include studies of the use of ECMO in children with respiratory failure (37, 38), the use of ECPR in children (39, 40), and the use of different pump technologies to support children on ECMO (41).

DISEASE-SPECIFIC REGISTRIES (RENÉE I. MATOS)

Examples: the Department of Defense Trauma Registry (DoDTR) and the American Heart Association Get With the Guidelines Resuscitation (AHA-GWTG-R) database

The Department of Defense Trauma Registry (DoDTR) (formerly the Joint Theater Trauma Registry) includes U.S. military, non-U.S. coalition soldiers, and local civilian trauma patients (42). Data contained in the registry includes resuscitation information, injuries, procedures, Abbreviated Injury Scores (AIS), complications, burn details, blood products and fluids administered, medications, laboratory results, and disposition. An overview of the DoDTR has been published, and the registry has expanded with modifications to data collection, uniform definitions, and improvements in standardization (43). Although quality improvement is the primary mission of the DoDTR, several pediatric studies have used this data source (44, 45). It is unique because it contains a higher incidence of penetrating, blast, and burn injuries than is seen in most U.S. civilian hospitals.

The Association Get With the Guidelines Resuscitation (AHA-GWTG-R) (formerly the National Registry of Cardiopulmonary

Resuscitation) began in 1999 and is a prospective multi-center registry of consecutive patients with in-hospital cardiac arrests (46). The registry contains facility, patient demographics, pre-event, event, patient outcome (specifically, return of spontaneous circulation, neurologic outcome, and survival to hospital discharge), and quality improvement data. Limitations of the AHA-GWTG-R include the lack of physiologic variables, such as laboratory results or vital signs, and the convenience sampling frame. Despite its limitations, it has been the source for several important analyses of critically ill children. Examples of research using the AHA-GWTG-R database include studies of the effects of patient age (46), heart rhythm (47), and cardiopulmonary resuscitation duration (48) on outcome after cardiac arrest.

POPULATION-BASED REGISTRIES (SHERI CROW)

Example: the Rochester Epidemiology Project (REP)

The Rochester Epidemiology Project (REP) (49) is a research infrastructure system supported by the NIH that collects, archives, links, and indexes the medical records of virtually all individuals who have resided in Olmsted County, MN since 1966 (50). The REP supports near complete assessment of health care utilization by a population of children, allowing estimation of the true incidence and prevalence of pediatric critical illness within a geographically defined area. Furthermore, medical record linkage for county residents throughout their lifetime facilitates long-term follow-up (51). A 2011 census identified 41,332 children <20 years of age who were current residents of Olmsted County (49).

The REP has been used for epidemiologic studies on a wide range of medical conditions. Most recently, REP data linkage with Olmsted County school records was utilized to identify a higher rate of attention deficit disorder and learning disability in children with two or more anesthetic exposures prior to age 4 (52).

QUALITY IMPROVEMENT OR BENCHMARKING DATABASES

(TELLEN D. BENNETT)

Example: the Pediatric Health Information Systems (PHIS) database

The Pediatric Health Information Systems (PHIS) database was originally created by the Children's Hospital Association (CHA), a business alliance of 44 free-standing children's hospitals, as a collective purchasing database (53). PHIS is now more often used for quality improvement, benchmarking, and research purposes. Data are readily available since January 1, 2001 (for some hospitals, back to 1992). The resource utilization data are the most unique feature of PHIS; specific codes are used for each medication, imaging study, laboratory test, nursing or respiratory therapy, hospital room, or supplied material. Although identifiers are masked within the database, patients have one medical record number at a given hospital, making longitudinal studies of hospitalizations possible. The PHIS+ project at six member hospitals includes laboratory, microbiology, and imaging results data (53).

The PHIS database has most often been used to identify practice variation (54), to track trends in utilization (55), and to analyze readmissions (56). In addition, some comparative effectiveness studies have been reported (57). Other work relevant to PCCM includes analyses of cardiac surgery (33), complicated pneumonia (58), asthma (59), and traumatic brain injury (55).

CLAIMS DATABASES (R. SCOTT WATSON)

Example: state medicaid databases

Medicaid claims data have been used in policy-related research in neonatal critical care (60), but they have not been well studied in pediatric critical care. Major potential advantages of Medicaid data are that they include longitudinal data on large populations of children, including use of inpatient, outpatient, and long-term care services; prescription medications; and durable medical equipment. Thus, they can provide information about events preceding and following a hospitalization for critical illness, including information regarding geographic patterns of care and regionalization. The challenges to working with Medicaid data are related to differences in Medicaid eligibility for families between states and over time. In addition, covered services for which data are available vary by state.

Medicaid data have become easier to obtain and analyze. Person-level data are provided as part of the Medicaid Analytic eExtract (MAX) files (61). Applications to obtain data are submitted through the Research Data Assistance Center.

GOVERNMENT ADMINISTRATIVE DATABASES (BRIAN D. BENNEYWORTH)

Example: Healthcare Cost and Utilization Project (HCUP) databases

The Agency for Healthcare Research and Quality (AHRQ) developed the Healthcare Cost and Utilization Project (HCUP) databases to provide a source of nationally representative inpatient discharges. The National Inpatient Sample (NIS) and the Kids' Inpatient Database (KID) contain inpatient data. Both are derived from the individual states' State Inpatient Database (SID) files. The Nationwide Emergency Department Sample (NEDS), a compilation of State Emergency Department (SED) files is also available (62).

Most pediatric studies have used the KID, which has been aggregated every 3 years from 1997 to 2009. The 2009 KID contains data on hospital discharges for patients ≤ 20 years old from 4,121 hospitals in 44 states. The KID stratifies hospitals by geographic region, hospital control, urban/rural location, teaching status, bed size, and hospital type to obtain a nationally representative sample. Within each hospital, routine normal newborn birth-related discharges are then sampled at 10% while all other pediatric hospitalizations are sampled at 80%. This allows for excellent statistical power to detect rare diseases.

Healthcare Cost and Utilization Project datasets have been used in pediatric populations to study central line infections (63), chronic mechanical ventilation (64), complex chronic conditions (65), sepsis (66), and trauma (67).

DATABASE LINKAGE

Record linkage techniques can be used to connect the records of the same patient in two or more different data sources. Linkage can enrich data sources by bringing in complementary and/or longitudinal patient variables from different phases of care and expand the array of questions a single data source might be used to study (68). Database linkage is generally accomplished by deterministic ("direct") linkage when databases share a unique identifier or by probabilistic ("indirect") linkage when they do not. The PHIS database has been a center of recent linkage activity, including links

to the STS database (69), the Children's Oncology Group database (70), and the electronic medical records of six children's hospitals (53). Other linkages of the PHIS database, including a linkage of VPS and PHIS, are in progress or have been proposed (Matthew Hall, PhD, personal communication to Tellen D. Bennett).

DATA SOURCE REGISTRY

This manuscript was necessarily selective in choosing databases to discuss, and it is likely that we are unaware of other rich data sources. To facilitate efficient and productive use of existing data, we have created a registry (<http://vpicu.info/pedal/>), where known data sources, their contents, and their availability are listed. A form on the site can be used to enter information about other data sources.

CONCLUSION

Many sources of existing data are available to PCCM investigators. We have categorized a number of frequently used data sources, identified research questions for which they may be appropriate, and created an online registry of data sources. Two recent manuscripts have achieved some of these goals for researchers studying adult patients (6, 71). The growth of quality improvement and safety research, improvements in personal computer and statistical package capability, and the increasing number of investigators with skills in data analysis have contributed to the growth of research using existing data (6). Efforts to improve the quality of studies using existing data have resulted in several reporting standards for observational research, including the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (72).

Overall, PCCM would benefit from increased linkage and integration of data sources to improve their granularity and level of clinical detail. Patient identifiers that would facilitate database linkage could dramatically expand the range of questions that could be answered with existing data. The privacy risks inherent in such a strategy are substantial and may require novel technological solutions. Similar risks are involved with inclusion of provider identifiers in large databases, but that will be necessary to accurately analyze within- and between-hospital variation. More clinical detail including the contents of electronic medical records systems, severity of illness measures, and longitudinal follow-up variables would improve the ability of investigators to fully understand the health of the population and the long-term impact of a PICU stay. Individual databases might accelerate their improvement processes by developing strategies to incorporate user feedback.

Efficiency is a major potential benefit of existing data analysis; a multi-center dataset may contain enough patients of a particular type to perform an adequately powered study at much lower cost and in a shorter period of time than a prospective study (73). The NIH is likely to be increasingly attentive to return on their investment in the future, and a large prospective study may not be funded if the question can be adequately answered using existing data (74).

Few of the datasets used by PCCM investigators are "big data" compared to those used in computer science, business, and public policy. However, many of them are "quirky and messy" (e.g.,

informative missingness, dependent observations, lack of a unique identifier, and evolving data standards) in ways that provide challenges to their use (75). Expertise in the analysis of existing data will be beneficial to investigators using these sources. New data types such as genomic data and signal data [e.g., the adult patient-focused Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC II) database] often are “big data” and will present new challenges. Collection and analysis of waveform data (heart rate, arterial blood pressure, end-tidal carbon dioxide, etc.) from the PICU patients at many centers should be a goal for the future.

In conclusion, many existing data sources are potentially useful for PCCM investigators, and analyses of existing data are likely to have a growing impact on the field of pediatric critical care. Careful matching of the scientific question and the best available data source or linked data sources is necessary. In addition, rigorous application of the best available analysis techniques and reporting consistent with observational research standards will maximize the impact of research using existing data in PCCM.

AUTHOR CONTRIBUTIONS

Tellen D. Bennett, Michael C. Spaeder, R. Scott Watson, Katri V. Typpo, Robinder G. Khemani, Sheri Crow, Brian D. Benneyworth, J. Michael Dean, and Barry P. Markovitz designed the study, each author drafted at least one section and contributed to the tables, Tellen D. Bennett wrote the first draft of the manuscript, and all authors contributed to its revision. All authors have seen and approved this final version of the manuscript and agree to be accountable for all aspects of the work.

ACKNOWLEDGMENTS

We are indebted to Matthew Hall, Ph.D., Jeff Jacobs, M.D., and Marshall Jacobs, M.D., for the information they provided. Tellen D. Bennett was supported by NICHD 1K23HD074620. This work was performed while Dr. Bennett was at the University of Utah School of Medicine.

REFERENCES

- Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA* (2013) **309**(13):1351–2. doi:10.1001/jama.2013.393
- Lohr S. Sizing up big data. *NY Times* (2013) **20**:2013.
- Hunt J, Chang AC. Big data in pediatric cardiac care: is it time? *Pediatr Crit Care Med* (2013) **14**(5):548–9. doi:10.1097/PCC.0b013e318291736b
- Safran C, Bloomrosen M, Hammond WE, Labkoff S, Markel-Fox S, Tang PC, et al. Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper. *J Am Med Inform Assoc* (2007) **14**(1):1–9. doi:10.1197/jamia.M2273
- Dreyer NA. Making observational studies count: shaping the future of comparative effectiveness research. *Epidemiology* (2011) **22**(3):295–7. doi:10.1097/EDE.0b013e3182126569
- Cooke CR, Iwashyna TJ. Using existing data to address important clinical questions in critical care. *Crit Care Med* (2013) **41**(3):886–96. doi:10.1097/CCM.0b013e31827bfc3c
- Namachivayam P, Shann F, Shekerdemian L, Taylor A, van Sloten I, Delzoppo C, et al. Three decades of pediatric intensive care: who was admitted, what happened in intensive care, and what happened afterward. *Pediatr Crit Care Med* (2010) **11**(5):549–55. doi:10.1097/PCC.0b013e3181ce7427
- Slonim AD, Khandelwal S, He J, Hall M, Stockwell DC, Turenne WM, et al. Characteristics associated with pediatric inpatient death. *Pediatrics* (2010) **125**(6):1208–16. doi:10.1542/peds.2009-1451
- Willson DF, Dean JM, Newth C, Pollack M, Anand KJ, Meert K, et al. Collaborative Pediatric Critical Care Research Network (CPCCRN). *Pediatr Crit Care Med* (2006) **7**(4):301–7. doi:10.1097/01.PCC.0000227106.66902.4F
- Black N, Payne M. Directory of clinical databases: improving and promoting their use. *Qual Saf Health Care* (2003) **12**(5):348–52. doi:10.1136/qhc.12.5.348
- Curley MA, Hibberd PL, Fineman LD, Wypij D, Shih MC, Thompson JE, et al. Effect of prone positioning on clinical outcomes in children with acute lung injury: a randomized controlled trial. *JAMA* (2005) **294**(2):229–37. doi:10.1001/jama.294.2.229
- Nadel S, Goldstein B, Williams M, Dalton H, Peters M, Macias W, et al. Drotrecogin alfa (activated) in children with severe sepsis: a multicentre phase III randomised controlled trial. *Lancet* (2007) **369**(9564):836–43. doi:10.1016/S0140-6736(07)60411-5
- Lacroix J, Hébert P, Hutchison J, Hume H, Tucci M, Ducruet T, et al. Transfusion strategies for patients in pediatric intensive care units. *N Engl J Med* (2007) **356**(16):1609–19. doi:10.1056/NEJMoa066240
- Farris RW, Weiss NS, Zimmerman JJ. Functional outcomes in pediatric severe sepsis: further analysis of the researching severe sepsis and organ dysfunction in children: a global perspective trial. *Pediatr Crit Care Med* (2013) **14**(9):835–42. doi:10.1097/PCC.0b013e3182a551c8
- Czaja AS, Scanlon MC, Kuhn EM, Jeffries HE. Performance of the pediatric index of mortality 2 for pediatric cardiac surgery patients. *Pediatr Crit Care Med* (2011) **12**(2):184–9. doi:10.1097/PCC.0b013e3181e89694
- Wetzel RC, Sachdeva R, Rice TB. Are all ICUs the same? *Paediatr Anaesth* (2011) **21**(7):787–93. doi:10.1111/j.1460-9592.2011.03595.x
- LaRovere JM, Jeffries HE, Sachdeva RC, Rice TB, Wetzel RC, Cooper DS, et al. Databases for assessing the outcomes of the treatment of patients with congenital and paediatric cardiac disease – the perspective of critical care. *Cardiol Young* (2008) **18**(Suppl 2):130–6. doi:10.1017/S1047951108002886
- Typpo KV, Petersen NJ, Petersen LA, Mariscalco MM. Children with chronic illness return to their baseline functional status after organ dysfunction on the first day of admission in the pediatric intensive care unit. *J Pediatr* (2010) **157**(1):108.e–13.e. doi:10.1016/j.jpeds.2009.12.029
- Kurachek SC, Newth CJ, Quasney MW, Rice T, Sachdeva RC, Patel NR, et al. Extubation failure in pediatric intensive care: a multiple-center study of risk factors and outcomes. *Crit Care Med* (2003) **31**(11):2657–64. [Erratum appears in *Crit Care Med*. **32**(7):1632–3. Note: Scanlon Mathew (corrected to Scanlon Matthew)]. doi:10.1097/01.CCM.0000094228.90557.85
- Ghuman AK, Newth CJ, Khemani RG. Impact of gender on sepsis mortality and severity of illness for prepubertal and postpubertal children. *J Pediatr* (2013) **17**(13):435–6. doi:10.1016/j.jpeds.2013.04.018
- Erickson S, Schibler A, Numa A, Nuthall G, Yung M, Pascoe E, et al. Acute lung injury in pediatric intensive care in Australia and New Zealand: a prospective, multicenter, observational study. *Pediatr Crit Care Med* (2007) **8**(4):317–23. doi:10.1097/01.PCC.0000269408.64179.FF
- Yung M, Wilkins B, Norton L, Slater A. Glucose control, organ failure, and mortality in pediatric intensive care. *Pediatr Crit Care Med* (2008) **9**(2):147–52. doi:10.1097/PCC.0b013e3181668c22
- Straney L, Clements A, Alexander J, Slater A. Variation in duration of respiratory support among Australian and New Zealand pediatric intensive care units. *Pediatr Crit Care Med* (2011) **12**(1):9–13. doi:10.1097/PCC.0b013e3181d8e90a
- Straney L, Clements A, Alexander J, Slater A. Quantifying variation of paediatric length of stay among intensive care units in Australia and New Zealand. *Qual Saf Health Care* (2010) **19**(6):e5. doi:10.1136/qshc.2008.028811
- Ferguson LP, Durward A, Tibby SM. Relationship between arterial partial oxygen pressure after resuscitation from cardiac arrest and mortality in children. *Circulation* (2012) **126**(3):335–42. doi:10.1161/CIRCULATIONAHA.111.085100
- Fraser LK, Fleming T, Miller M, Draper ES, McKinney PA, Parslow RC. Palliative care discharge from paediatric intensive care units in Great Britain. *Palliat Med* (2010) **24**(6):608–15. doi:10.1177/0269216310364200
- Absoud M, Parslow RC, Wassmer E, Hemingway C, Duncan HP, Cummins C, et al. Severe acute disseminated encephalomyelitis: a paediatric intensive care population-based study. *Mult Scler* (2011) **17**(10):1258–61. doi:10.1177/1352458510382554
- Burns MR, Bodansky HJ, Parslow RC. Paediatric intensive care admissions for acute diabetes complications. *Diabet Med* (2010) **27**(6):705–8. doi:10.1111/j.1464-5491.2010.02939.x
- Mavroudis C, Bove EL, Cameron DE, Drinkwater DC Jr, Edwards FH, Hammon JW Jr, et al. The society of thoracic surgeons' National Congenital Heart Surgery Database. *Ann Thorac Surg* (1995) **59**(2):554–6. doi:10.1016/0003-4975(94)00947-6

30. Jacobs JP, Jacobs ML, Mavroudis C, Backer CL, Lacour-Gayet FG, Tcher-venkov CI, et al. Nomenclature and databases for the surgical treatment of congenital cardiac disease – an updated primer and an analysis of opportunities for improvement. *Cardiol Young* (2008) **18**(Suppl 2):38–62. doi:10.1017/S1047951108003028
31. Vener DF, Guzzetta N, Jacobs JP, Williams GD. Development and implementation of a new data registry in congenital cardiac anesthesia. *Ann Thorac Surg* (2012) **94**(6):2159–65. doi:10.1016/j.athoracsur.2012.06.070
32. Johnson JN, Jagers J, Li S, O'Brien SM, Li JS, Jacobs JP, et al. Center variation and outcomes associated with delayed sternal closure after stage 1 palliation for hypoplastic left heart syndrome. *J Thorac Cardiovasc Surg* (2010) **139**(5):1205–10. doi:10.1016/j.jtcvs.2009.11.029
33. Pasquali SK, Li JS, He X, Jacobs ML, O'Brien SM, Hall M, et al. Perioperative methylprednisolone and outcome in neonates undergoing heart surgery. *Pediatrics* (2012) **129**(2):e385–91. doi:10.1542/peds.2011-2034
34. Welke KF, O'Brien SM, Peterson ED, Ungerleider RM, Jacobs ML, Jacobs JP. The complex relationship between pediatric cardiac surgical case volumes and mortality rates in a national clinical database. *J Thorac Cardiovasc Surg* (2009) **137**(5):1133–40. doi:10.1016/j.jtcvs.2008.12.012
35. DiBardino DJ, Pasquali SK, Hirsch JC, Benjamin DK, Kleeman KC, Salazar JD, et al. Effect of sex and race on outcome in patients undergoing congenital heart surgery: an analysis of the society of thoracic surgeons congenital heart surgery database. *Ann Thorac Surg* (2012) **94**(6):2054–9; discussion 9–60. doi:10.1016/j.athoracsur.2012.05.124
36. Paden ML, Conrad SA, Rycus PT, Thiagarajan RR. Extracorporeal Life Support Organization Registry Report 2012. *ASAIO J* (2013) **59**(3):202–10. doi:10.1097/MAT.0b013e3182904a52
37. Zabrocki LA, Brogan TV, Statler KD, Poss WB, Rollins MD, Bratton SL. Extracorporeal membrane oxygenation for pediatric respiratory failure: survival and predictors of mortality. *Crit Care Med* (2011) **39**(2):364–70. doi:10.1097/CCM.0b013e3181fb7b35
38. Brogan TV, Zabrocki L, Thiagarajan RR, Rycus PT, Bratton SL. Prolonged extracorporeal membrane oxygenation for children with respiratory failure. *Pediatr Crit Care Med* (2012) **13**(4):e249–54. doi:10.1097/PCC.0b013e31824176f4
39. Thiagarajan RR, Laussen PC, Rycus PT, Bartlett RH, Bratton SL. Extracorporeal membrane oxygenation to aid cardiopulmonary resuscitation in infants and children. *Circulation* (2007) **116**(15):1693–700. doi:10.1161/CIRCULATIONAHA.106.680678
40. Barrett CS, Bratton SL, Salvin JW, Laussen PC, Rycus PT, Thiagarajan RR. Neurological injury after extracorporeal membrane oxygenation use to aid pediatric cardiopulmonary resuscitation. *Pediatr Crit Care Med* (2009) **10**(4):445–51. doi:10.1097/PCC.0b013e318198bd85
41. Barrett CS, Thiagarajan RR. Centrifugal pump circuits for neonatal extracorporeal membrane oxygenation. *Pediatr Crit Care Med* (2012) **13**(4):492–3. doi:10.1097/PCC.0b013e31824174c8
42. United States Army Institute of Surgical Research. Joint trauma system. (2013). Available from: http://usaisr.amedd.army.mil/joint_trauma_system.html.
43. Eastridge BJ, Jenkins D, Flaherty S, Schiller H, Holcomb JB. Trauma system development in a theater of war: experiences from Operation Iraqi Freedom and Operation Enduring Freedom. *J Trauma* (2006) **61**(6):1366–72; discussion 72–3. doi:10.1097/01.ta.0000245894.78941.90
44. Burnett MW, Spinella PC, Azarow KS, Callahan CW. Pediatric care as part of the US Army medical mission in the global war on terrorism in Afghanistan and Iraq, December 2001 to December 2004. *Pediatrics* (2008) **121**(2):261–5. doi:10.1542/peds.2006-3666
45. Borgman M, Matos RI, Blackburn LH, Spinella PC. Ten years of military pediatric care in Afghanistan and Iraq. *J Trauma Acute Care Surg* (2012) **73**(6 Suppl 5):S509–13. doi:10.1097/TA.0b013e318275477c
46. Meaney PA, Nadkarni VM, Cook EF, Testa M, Helfaer M, Kaye W, et al. Higher survival rates among younger patients after pediatric intensive care unit cardiac arrests. *Pediatrics* (2006) **118**(6):2424–33. doi:10.1542/peds.2006-1724
47. Samson RA, Nadkarni VM, Meaney PA, Carey SM, Berg MD, Berg RA. Outcomes of in-hospital ventricular fibrillation in children. *N Engl J Med* (2006) **354**(22):2328–39. doi:10.1056/NEJMoa052917
48. Matos RI, Watson RS, Nadkarni VM, Huang HH, Berg RA, Meaney PA, et al. Duration of cardiopulmonary resuscitation and illness category impact survival and neurologic outcomes for in-hospital pediatric cardiac arrests. *Circulation* (2013) **127**(4):442–51. doi:10.1161/CIRCULATIONAHA.112.125625
49. Rochester Epidemiology Project. (2013). Available from: <http://www.rochesterproject.org>.
50. Rocca WA, Yawn BP, St Sauver JL, Grossardt BR, Melton LJ III. History of the Rochester Epidemiology Project: half a century of medical records linkage in a US population. *Mayo Clin Proc* (2012) **87**(12):1202–13. doi:10.1016/j.mayocp.2012.08.012
51. St Sauver JL, Grossardt BR, Yawn BP, Melton LJ III, Pankratz JJ, Brue SM, et al. Data resource profile: the Rochester Epidemiology Project (REP) medical records-linkage system. *Int J Epidemiol* (2012) **41**(6):1614–24. doi:10.1093/ije/dys195
52. Flick RP, Katusic SK, Colligan RC, Wilder RT, Voigt RG, Olson MD, et al. Cognitive and behavioral outcomes after early exposure to anesthesia and surgery. *Pediatrics* (2011) **128**(5):e1053–61. doi:10.1542/peds.2011-0351
53. Narus SP, Srivastava R, Gouripeddi R, Livne OE, Mo P, Bickel JP, et al. Federating clinical data from six pediatric hospitals: process and initial results from the PHIS+ Consortium. *AMIA Annu Symp Proc* (2011) **2011**:994–1003.
54. Rice-Townsend S, Hall M, Barnes JN, Baxter JK, Rangel SJ. Hospital readmission after management of appendicitis at freestanding children's hospitals: contemporary trends and financial implications. *J Pediatr Surg* (2012) **47**(6):1170–6. doi:10.1016/j.jpedsurg.2012.03.025
55. Bennett TD, Statler KD, Korgenski EK, Bratton SL. Osmolar therapy in pediatric traumatic brain injury. *Crit Care Med* (2012) **40**(1):208–15. doi:10.1097/CCM.0b013e31822e9d31
56. Berry JG, Hall DE, Kuo DZ, Cohen E, Agrawal R, Feudtner C, et al. Hospital utilization and characteristics of patients experiencing recurrent readmissions within children's hospitals. *JAMA* (2011) **305**(7):682–90. doi:10.1001/jama.2011.122
57. Coffin SE, Leckerman K, Keren R, Hall M, Localio R, Zaoutis TE. Oseltamivir shortens hospital stays of critically ill children hospitalized with seasonal influenza: a retrospective cohort study. *Pediatr Infect Dis J* (2011) **30**(11):962–6. doi:10.1097/INF.0b013e318232ede9
58. Shah SS, Hall M, Newland JG, Brogan TV, Farris RW, Williams DJ, et al. Comparative effectiveness of pleural drainage procedures for the treatment of complicated pneumonia in childhood. *J Hosp Med* (2011) **6**(5):256–63. doi:10.1002/jhm.872
59. Morse RB, Hall M, Fieldston ES, McGwire G, Anspacher M, Sills MR, et al. Hospital-level compliance with asthma care quality measures at children's hospitals and subsequent asthma-related outcomes. *JAMA* (2011) **306**(13):1454–60. doi:10.1001/jama.2011.1385
60. Rogowski J. Cost-effectiveness of care for very low birth weight infants. *Pediatrics* (1998) **102**(1 Pt 1):35–43. doi:10.1542/peds.102.1.35
61. Centers for Medicare and Medicaid Services. Medicaid Analytic eXtract (MAX) General Information. (2013). Available from: <http://www.medicare.gov/Medicaid-CHIP-Program-Information/By-Topics/Data-and-Systems/MAX/MAX-General-Information.html>.
62. Agency for Healthcare Research and Quality. HCUP.net. Healthcare Cost and Utilization Project (HCUP) 2000–2009. Rockville, MD (2013). Available from: <http://hcupnet.ahrq.gov>.
63. Patrick SW, Davis MM, Sedman AB, Meddings JA, Hieber S, Lee GM, et al. Accuracy of hospital administrative data in reporting central line-associated bloodstream infections in newborns. *Pediatrics* (2013) **131**(Suppl 1):S75–80. doi:10.1542/peds.2012-1427i
64. Benneyworth BD, Gebremariam A, Clark SJ, Shanley TP, Davis MM. Inpatient health care utilization for children dependent on long-term mechanical ventilation. *Pediatrics* (2011) **127**(6):e1533–41. doi:10.1542/peds.2010-2026
65. Simon TD, Berry J, Feudtner C, Stone BL, Sheng X, Bratton SL, et al. Children with complex chronic conditions in inpatient hospital settings in the United States. *Pediatrics* (2010) **126**(4):647–55. doi:10.1542/peds.2009-3266
66. Odetola FO, Gebremariam A, Freed GL. Patient and hospital correlates of clinical outcomes and resource utilization in severe pediatric sepsis. *Pediatrics* (2007) **119**(3):487–94. doi:10.1542/peds.2006-2353
67. Tilford JM, Aitken ME, Anand KJ, Green JW, Goodman AC, Parker JG, et al. Hospitalizations for critically ill children with traumatic brain injuries: a longitudinal analysis. *Crit Care Med* (2005) **33**(9):2074–81. doi:10.1097/01.CCM.0000171839.65687.F5
68. Navathe AS, Clancy C, Glied S. Advancing research data infrastructure for patient-centered outcomes research. *JAMA* (2011) **306**(11):1254–5. doi:10.1001/jama.2011.1341

69. Pasquali SK, Jacobs JP, Shook GJ, O'Brien SM, Hall M, Jacobs ML, et al. Linking clinical registry data with administrative data using indirect identifiers: implementation and validation in the congenital heart surgery population. *Am Heart J* (2010) **160**(6):1099–104. doi:10.1016/j.ahj.2010.08.010
70. Aplenc R, Fisher BT, Huang YS, Li Y, Alonzo TA, Gerbing RB, et al. Merging of the National Cancer Institute-funded cooperative oncology group data with an administrative data source to develop a more effective platform for clinical trial analysis and comparative effectiveness research: a report from the Children's Oncology Group. *Pharmacoepidemiol Drug Saf* (2012) **21**(Suppl 2):37–43. doi:10.1002/pds.3241
71. Smith AK, Ayanian JZ, Covinsky KE, Landon BE, McCarthy EP, Wee CC, et al. Conducting high-value secondary dataset analysis: an introductory guide and resources. *J Gen Intern Med* (2011) **26**(8):920–9. doi:10.1007/s11606-010-1621-5
72. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* (2007) **370**(9596):1453–7. doi:10.1016/S0140-6736(07)61602-X
73. Maas AI, Menon DK, Lingsma HF, Pineda JA, Sandel ME, Manley GT. Reorientation of clinical research in traumatic brain injury: report of an international workshop on comparative effectiveness research. *J Neurotrauma* (2012) **29**(1):32–46. doi:10.1089/neu.2010.1599
74. Emanuel EJ. The future of biomedical research. *JAMA* (2013) **309**(15):1589–90. doi:10.1001/jama.2013.2096
75. Leek J. Simply statistics blog, 2013 May 29: What statistics should do about big data: problem forward and not solution backward [Internet]. [cited 2013 July 16]. Available from: www.simplystatistics.org

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 June 2014; paper pending published: 25 June 2014; accepted: 14 July 2014; published online: 29 July 2014.

Citation: Bennett TD, Spaeder MC, Matos RI, Watson RS, Typpo KV, Khemani RG, Crow S, Benneyworth BD, Thiagarajan RR, Dean JM, Markovitz BP and for the Pediatric Acute Lung Injury and Sepsis Investigators (PALISI) (2014) Existing data analysis in pediatric critical care research. *Front. Pediatr.* 2:79. doi: 10.3389/fped.2014.00079
This article was submitted to *Pediatric Critical Care*, a section of the journal *Frontiers in Pediatrics*.

Copyright © 2014 Bennett, Spaeder, Matos, Watson, Typpo, Khemani, Crow, Benneyworth, Thiagarajan, Dean, Markovitz and for the Pediatric Acute Lung Injury and Sepsis Investigators (PALISI). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.