



Leveraging big data to advance knowledge in pediatric heart failure and heart transplantation

Justin Godown¹, Michael Gaies², James D. Wilkinson³

¹Division of Pediatric Cardiology, Monroe Carell Jr. Children's Hospital, Nashville, TN, USA; ²Division of Cardiology, Department of Pediatrics, C.S. Mott Children's Hospital, University of Michigan Medical School, Ann Arbor, MI, USA; ³Department of Pediatrics, Monroe Carell Jr. Children's Hospital, Nashville, TN, USA

Correspondence to: Justin Godown, MD. Division of Pediatric Cardiology, Monroe Carell Jr. Children's Hospital, Vanderbilt University, 2200 Children's Way, Suite 5230 DOT, Nashville, TN 37232, USA. Email: justin.godown@vumc.org.

Submitted May 10, 2019. Accepted for publication Jul 11, 2019.

doi: 10.21037/tp.2019.07.09

View this article at: <http://dx.doi.org/10.21037/tp.2019.07.09>

Introduction

Despite many therapeutic advances for children with heart failure and those who undergo heart transplantation, robust evidence to guide clinical practice remains limited. The approach to these patients has been largely based upon available evidence in the adult population. Each pediatric heart center has a relatively small number of heart failure and heart transplant patients, making single-center studies largely underpowered. While multi-center studies have resulted in significant advances in the field, challenges in enrolling sufficient children to achieve appropriate statistical power in a reasonable time period remains a challenge (1-4). This fact has increasingly led to the use of large multi-center databases to facilitate clinical research in pediatric cardiology, and the number of publications utilizing this approach has increased exponentially over time (*Figure 1*). The use of large databases is cost-effective, results in acceptable statistical power, and facilitates timely dissemination of results. However, understanding the inherent limitations and challenges of leveraging these large datasets is critical to successful research. This review will highlight the advantages and limitations of using large databases and clinical registries to conduct research in the field of pediatric heart failure and heart transplantation, review existing data sources, discuss novel approaches to data linkage, and explore future directions.

Existing datasets

There are numerous large datasets with potential utility

for researchers addressing questions pertaining to pediatric heart failure and heart transplantation. These include a variety of clinical registries, research databases, administrative databases, and quality improvement initiatives (5). Each data source has specific advantages and limitations as well as different requirements for researchers requesting access.

Heart Transplant-specific databases

The Scientific Registry of Transplant Recipients and the Organ Procurement and Transplantation Network (<https://www.srtr.org/> and <https://optn.transplant.hrsa.gov/>)

The Scientific Registry of Transplant Recipients (SRTR) and the Organ Procurement and Transplantation Network (OPTN) databases contain information pertaining to all donors, waitlisted candidates, and transplant recipients in the United States beginning in late 1987. These data are publicly available upon request and therefore represent a popular data source for transplantation research. Members of the United Network for Organ Sharing are required to contribute data, ensuring that these datasets encompass the entire spectrum of transplant medicine in the United States. Data are collected at discrete intervals including at: (I) listing, (II) transplant, (III) discharge following transplantation, and (IV) annually post-transplant. Despite mandatory data submission, missing data can be problematic. Additionally, changes to data collection requirements over time present a challenge for clinical research as incorporation of multiple different variables may

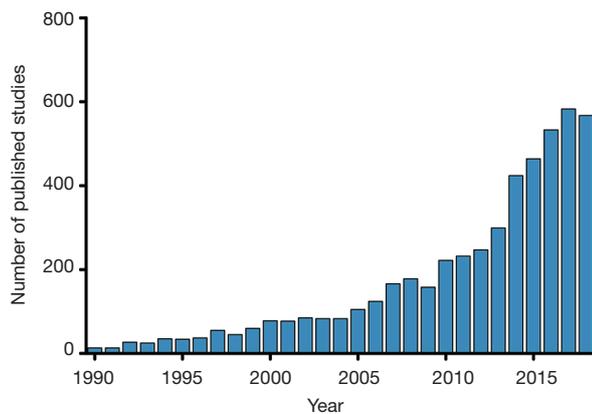


Figure 1 Number of published pediatric cardiology studies by year with the keywords “database” or “registry” or naming a specific data source (i.e., PHTS, PCMR, SRTR). Data from PubMed (accessed 5/2019).

be necessary to accomplish analyses. Lastly, assessing the timing of post-transplant events is impaired by the fact that data are collected at specified time points and are not event-driven (5). Researchers should be aware of these concerns and data quality should be carefully assessed for all analyses performed.

The International Society for Heart and Lung Transplantation (<https://isblt.org/registries/ttx-registry>)

The International Society for Heart and Lung Transplantation (ISHLT) registry contains data derived from SRTR/OPTN and therefore contains similar data pertaining to transplant recipients in the United States. However, these data are supplemented by data from international centers, which facilitates more robust analyses across international sites. Given that much of the data contained in this registry are derived from SRTR/OPTN, many of the same limitations apply. These data are available for research upon request and there are grant mechanisms in place to support research using these data.

The Pediatric Heart Transplant Society (<https://www.uab.edu/medicine/phts/>)

The Pediatric Heart Transplant Society (PHTS) is a large multi-center event-driven database housed at the University of Alabama at Birmingham that was established in 1993. Currently, 56 centers contribute data, accounting for approximately 85% to 90% of the pediatric heart

transplant volume annually in the United States (personal communication). This dataset was designed specifically to answer questions pertaining to pediatric heart transplantation, and therefore has distinct advantages over the previously mentioned data sources. Data collection instruments have changed over time to evolve with changing patterns in clinical practice and to improve data collection and quality. While this may limit some analyses, changes to data collection fields over time are well-documented and can be navigated.

In contrast to other data sources, PHTS has a centralized statistical core. Research proposals are submitted by participating centers and reviewed by a scientific committee. This approach improves the quality and consistency of statistical analyses. However, it also limits the number of projects that can be pursued at any given time. Researchers may also request limited datasets for analysis. Given that the PHTS database contains >25 years of data collection, it currently represents one of the most robust resources for retrospective research in the field of pediatric heart transplantation.

Pediatric heart failure databases

The Pediatric Cardiomyopathy Registry (<https://dev.childrenscardiomyopathy.org/Pediatric-Cardiomyopathy-Registry-71-315>)

The Pediatric Cardiomyopathy Registry (PCMR) was started in 1994 as a large multi-center observational study of cardiomyopathies in children (6). While the initial focus of the PCMR was to assess the epidemiology of pediatric cardiomyopathy from a retrospective cohort of patients, the study also included a prospective cohort. Since survival analyses did not differ significantly, the cohorts have subsequently been combined into a single database (7). The PCMR has evolved into a multicenter collaborative effort with a single data coordinating center focusing on hypothesis-driven prospective research in pediatric cardiomyopathy, beyond its original registry design. Limited de-identified datasets of the merged retrospective and prospective cohorts are available upon request and have been used for secondary analyses (8).

The International Pediatric Heart Failure Registry (<https://isblt.org/registries/ipbf-registry>)

The International Pediatric Heart Failure Registry was

established in 2016 to prospectively collect data regarding pediatric heart failure from 16 large U.S. and international centers. The registry aimed to, “*better understand the natural history and response to current treatment regimes.*” (9). Unfortunately, enrollment has recently been closed and the registry is no longer collecting data. However, data collected to date will likely be available to researchers for subsequent analyses in the future.

Mechanical circulatory support databases

The Interagency Registry for Mechanically Assisted Circulatory Support and The Pediatric Interagency Registry for Mechanically Assisted Circulatory Support (<https://www.uab.edu/medicine/intermacs/>)

The Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS) was established in 2006 to collect data pertaining to the implantation of durable mechanical circulatory support devices in adult patients. The Pediatric Interagency Registry for Mechanically Assisted Circulatory Support (PediMACS) was modeled after INTERMACS and collects data pertaining to ventricular assist device support in children and adolescents. As of April 2019, 46 centers participate in PediMACS and 766 patients have been enrolled (10). Investigators may submit research proposals to access PediMACS data, which are processed through the Society of Thoracic Surgeons (STS) Research Center.

Extracorporeal Life Support Organization (<https://www.else.org/>)

The Extracorporeal Life Support Organization (ELSO) is a large international collaboration that maintains a registry of extracorporeal membrane oxygenation (ECMO) use. The registry was established in 1990 and to date contains information pertaining to >100,000 ECMO runs across all ages. De-identified data are available upon request from participating centers for research, internal benchmarking, as well as quality assurance (11).

Congenital Cardiac Surgery Databases

The Society of Thoracic Surgeons Congenital Heart Surgery Database (<https://www.sts.org/>)

The Society of Thoracic Surgeons Congenital Heart Surgery Database (STS-CHS) represents the largest data source

regarding the surgical repair of congenital heart disease, containing data pertaining to >475,000 procedures (12). The majority of hospitals performing congenital cardiac surgery in the United States contribute data. The STS-CHS has distinct advantages over other data sources to assess surgical outcomes in patients with congenital heart disease as STS-CHS utilizes a robust classification system for congenital heart diagnoses, surgical procedures, and post-operative complications. The database also collects detailed syndromic and genetic information, which increases data granularity and expands potential analyses. Data requests from STS member institutions are reviewed twice yearly on a rolling cycle by a committee with both physician and biostatistics representation (12).

Quality Improvement Data Sources

Advanced Cardiac Therapies Improving Outcomes Network (<https://www.actionlearningnetwork.org/>)

The Advanced Cardiac Therapies Improving Outcomes Network (ACTION) is a learning network focused on quality improvement for pediatric patients requiring ventricular assist device support. ACTION aims to improve outcomes for ventricular assist device support in pediatric patients through the development of a learning network of pediatric heart failure providers, enabling faster improvements in care compared to traditional research. While the primary focus of this effort is quality improvement, there are opportunities for research and the network aims to become a platform for device trials in the future (13).

National Pediatric Cardiology Quality Improvement Collaborative (<https://npcqic.org/>)

The National Pediatric Cardiology Quality Improvement Collaborative (NPC-QIC) is a multicenter network developed with the goal to “reduce mortality and improve the quality of life of infants with hypoplastic left heart syndrome during the inter-stage period.” (14). Currently, there are 60 pediatric cardiac centers that participate in this quality improvement network which has resulted in the generation of a national registry containing detailed clinical and outcome data. These data are available to researchers from network institutions, pending approval of the research proposal by a dedicated research and publications committee.

The Pediatric Cardiac Critical Care Consortium (<http://pc4quality.org/>)

The Pediatric Cardiac Critical Care Consortium (PC4) is a quality improvement collaborative that collects data on all pediatric patients admitted to cardiac intensive care units in participating hospitals (15). The registry uses standardized data definitions shared with the International Pediatric and Congenital Cardiac Code (<http://ipccc.net>), STS-CHS, and American College of Cardiology Improving Pediatric and Adult Congenital Treatment Registry. Patients with acute decompensated heart failure, those status-post transplant, those with VADs, and other patients with heart failure who require critical care are included in this cohort. The PC4 registry has already been used to study the epidemiology and outcomes of patients with acute decompensated heart failure (16). PC4 has partnered with several other organizations (including NPC-QIC and PAC3) to form an integrated, comprehensive data infrastructure known as Cardiac Networks United (17).

The Pediatric Acute Care Cardiology Collaborative (<https://pac3quality.org/>)

The Pediatric Acute Care Cardiology Collaborative (PAC3) was founded in 2014 with an emphasis on improving outcomes of pediatric cardiology patients in hospital-based inpatient non-intensive care units (18). PAC3 aims to deliver higher quality and greater value care by facilitating the sharing of ideas and building alignment among its member institutions. The clinical registry for PAC3 was launched in February 2019 at eleven hospitals. The data are submitted and integrated with data obtained in the cardiac intensive care unit for the PC4, and similarly utilizes shared definitions with other major registries in CHS clinical care.

Administrative databases

The Pediatric Health Information System (<https://www.childrenshospitals.org/>)

The Pediatric Health Information System (PHIS) is an administrative database that collects clinical and resource utilization data for hospital encounters from >50 large U.S. children's hospitals. Data from inpatient hospitalizations, observation, ambulatory surgery, and emergency department encounters are collected. This database records diagnosis and procedural ICD codes which are supplemented with detailed billing data.

Providers and staff from member institutions can request access to PHIS data. Training is required prior to accessing PHIS data. The volume of data collected in PHIS provides a robust resource, but analyses are limited by the inherent challenges regarding the use of ICD codes, and those limitations should be considered when planning all analyses.

Kids' Inpatient Database (<https://www.hcup-us.abrq.gov/kidoverview.jsp>)

The Kids' Inpatient Database (KID) was developed as part of the Healthcare Cost Utilization Project. The KID is a publicly available administrative database reporting inpatient hospitalization data for children in the United States. Data from >4,000 hospitals in 47 states are included, making it the largest publicly available database of pediatric inpatient hospitalizations (19). Similar to the PHIS database, the KID collects clinical information through the use of ICD codes and also collects hospital charge data. Starting in 1997, the KID is produced every 3 years.

Expanding analytic possibilities through data linkage

Each large database has inherent advantages and disadvantages. Database linkage is a strategy that is increasingly utilized to expand analytic possibilities. Multiple existing linkages have been developed that can be utilized to study pediatric heart failure and heart transplantation (*Figure 2*). The advantages of linking large databases are clear. Establishing the ability to draw data from either source increases data availability, expands analytic possibilities, and allows cross-verification of data between data sources. However, there are potential pitfalls as well. While data linkage using indirect identifiers has been shown to be generally reliable (20-23), it is not perfect and there are inevitably some patients who are not successfully matched between datasets. Assessment of patients who failed to link successfully is important to ensure that resulting analyses are not biased. Additionally, it is important to perform internal validation to assess the quality and accuracy of established linkages.

Future directions

Leveraging large databases will remain a prominent strategy to advance research in the field of pediatric heart failure and heart transplantation in the future. However, optimizing

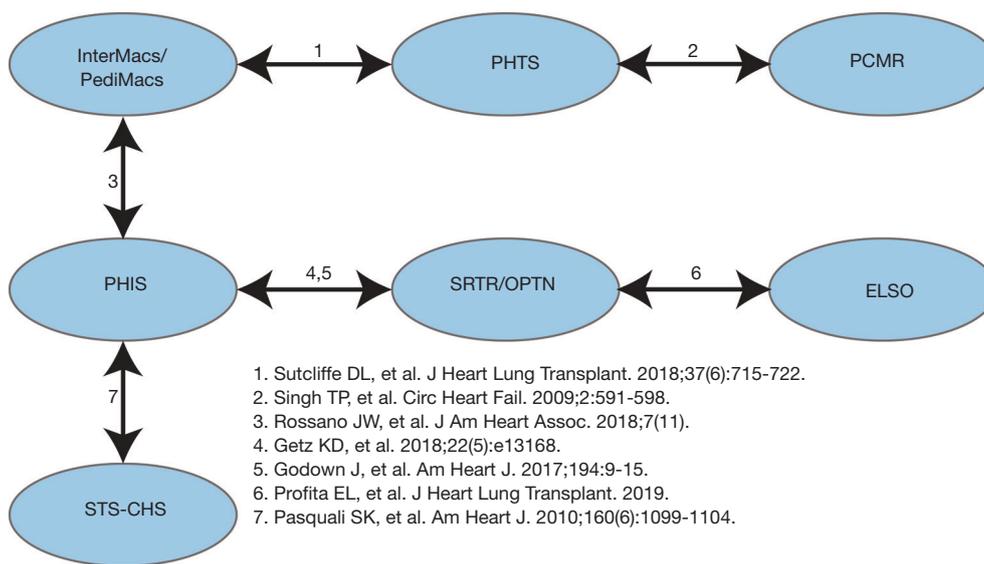


Figure 2 Existing data linkages in pediatric heart failure and heart transplantation. ELSO, Extracorporeal Life Support Organization; InterMacs, Interagency Registry for Mechanically Assisted Circulatory Support; OPTN, Organ Procurement and Transplantation Network; PCMR, Pediatric Cardiomyopathy Registry; Pedimacs, Pediatric Registry for Mechanically Assisted Circulatory Support; PHIS, Pediatric Health Information System; PHTS, Pediatric Heart Transplant Society; STS-CHS, Society of Thoracic Surgeons Congenital Heart Surgery Database.

data collection and integration across platforms is crucial to facilitate future work. Many existing databases collect the same patient information and therefore streamlining the process of data entry and collaboration between databases provides clear benefit. This would minimize the data entry burden at individual centers, improve data reliability, and ensure that data collected are consistent across datasets. This also provides an opportunity to assign individual subjects a standardized global unique identifier (GUID) to facilitate linkage of patient data across data sources without the need for linkage using indirect identifiers.

As referenced above, Cardiac Networks United (17) (<http://cardiacnetworksunited.org/>) is an organization aiming to integrate data from multiple sources in order to accelerate research and quality improvement efforts across congenital and pediatric cardiac care. Five initial networks formed Cardiac Networks United—ACTION, PC4, PAC3, Cardiac Neurodevelopmental Outcomes Collaborative (CNOC), and NPC-QIC—pledging to collaborate and share data and expertise. Several research and quality improvement initiatives are underway that leverage this infrastructure, and the organization will provide opportunities to link data between existing heart failure data repositories and the phenotypic and outcome data captured

by the member organizations in Cardiac Networks United.

Clinical registries and research databases are also being increasingly utilized as a platform for clinical trials (24,25). This represents a cost-effective and efficient strategy to perform clinical research by leveraging the existing infrastructure and data collection of registries. Given the numerous challenges associated with performing randomized clinical trials in pediatric heart failure and heart transplantation, this approach may be uniquely tailored to this area of study.

Conclusions

Large databases are commonly utilized for pediatric research, enabling multi-center analyses with increased statistical power. There are numerous available databases with potential utility in the field of pediatric heart failure and heart transplantation and database linkage is an increasingly common strategy to expand analytic possibilities. Researchers should recognize the limitations that exist surrounding the use of databases for clinical research. Efforts such as incorporating a GUID or incorporating clinical trials within the framework of a registry represent promising strategies for future advances.

Acknowledgments

None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Everitt MD, Wilkinson JD, Shi L, et al. Cardiac biomarkers in pediatric cardiomyopathy: Study design and recruitment results from the Pediatric Cardiomyopathy Registry. *Prog Pediatr Cardiol* 2019;53:1-10.
2. Pasquali SK, Lam WK, Chiswell K, et al. Status of the pediatric clinical trials enterprise: an analysis of the US ClinicalTrials.gov registry. *Pediatrics* 2012;130:e1269-77.
3. Zuckerman WA, Zeevi A, Mason KL, et al. Study rationale, design, and pretransplantation alloantibody status: A first report of Clinical Trials in Organ Transplantation in Children-04 (CTOTC-04) in pediatric heart transplantation. *Am J Transplant* 2018;18:2135-47.
4. Shaddy RE, Boucek MM, Hsu DT, et al. Carvedilol for children and adolescents with heart failure: a randomized controlled trial. *JAMA* 2007;298:1171-9.
5. Pasquali SK, Schumacher KR, Davies RR. Can linking databases answer questions about paediatric heart failure? *Cardiol Young* 2015;25 Suppl 2:160-6.
6. Wilkinson JD, Sleeper LA, Alvarez JA, et al. The Pediatric Cardiomyopathy Registry: 1995-2007. *Prog Pediatr Cardiol* 2008;25:31-6.
7. Towbin JA, Lowe AM, Colan SD, et al. Incidence, causes, and outcomes of dilated cardiomyopathy in children. *JAMA* 2006;296:1867-76.
8. Castleberry CD, Jefferies JL, Shi L, et al. No Obesity Paradox in Pediatric Patients With Dilated Cardiomyopathy. *JACC Heart Fail* 2018;6:222-30.
9. International Pediatric Heart Failure (IPHF) Registry. Available online: <https://ishlt.org/registries/iphf-registry>. Accessed April 29 2019.
10. Interagency Registry for Mechanically Assisted Circulatory Support. Available online: <http://www.uab.edu/intermacs>. Accessed April 29th 2019.
11. The Extracorporeal Life Support Organization (ELSO). Available online: <https://www.elso.org/>. Accessed April 29th 2019.
12. The Society of Thoracic Surgeons Congenital Heart Database (STS-CHD). Available online: <https://www.sts.org/registries-research-center/sts-national-database/sts-congenital-heart-surgery-database>. Accessed April 29th 2019.
13. Advanced Cardiac Therapies Improving Outcomes Network (ACTION). Available online: <https://www.actionlearningnetwork.org/>. Accessed April 29th 2019.
14. Kugler JD, Beekman Iii RH, Rosenthal GL, et al. Development of a pediatric cardiology quality improvement collaborative: from inception to implementation. From the Joint Council on Congenital Heart Disease Quality Improvement Task Force. *Congenit Heart Dis* 2009;4:318-28.
15. Gaies M, Cooper DS, Tabbutt S, et al. Collaborative quality improvement in the cardiac intensive care unit: development of the Paediatric Cardiac Critical Care Consortium (PC4). *Cardiol Young* 2015;25:951-7.
16. Lasa JJ, Gaies M, Rossano J, et al. Acute Decompensated Heart Failure in the Pediatric Population: A Report from the Pediatric Cardiac Critical Care Consortium (PC4). *J Heart Lung Transplant* 2018;37:S157-8.
17. Gaies M, Anderson J, Kipps A, et al. Cardiac Networks United: an integrated paediatric and congenital cardiovascular research and improvement network. *Cardiol Young* 2018. [Epub ahead of print].
18. Kipps AK, Cassidy SC, Strohacker CM, et al. Collective quality improvement in the paediatric cardiology acute care unit: establishment of the Pediatric Acute Care Cardiology Collaborative (PAC3). *Cardiol Young* 2018;28:1019-23.
19. Overview of the Kids' Inpatient Database (KID). Available online: <https://www.hcup-us.ahrq.gov/kidoverview.jsp>. Accessed April 29th 2019.
20. Godown J, Thurm C, Dodd DA, et al. A unique linkage of administrative and clinical registry databases to expand analytic possibilities in pediatric heart transplantation research. *Am Heart J* 2017;194:9-15.
21. Pasquali SK, Jacobs JP, Shook GJ, 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;160:1099-104.
22. Godown J, Hall M, Thompson B, et al. Expanding analytic possibilities in pediatric solid organ transplantation through linkage of administrative and clinical registry

- databases. *Pediatr Transplant* 2019;23:e13379.
23. Getz KD, He C, Li Y, et al. Successful merging of data from the United Network for Organ Sharing and the Pediatric Health Information System databases. *Pediatr Transplant* 2018;22:e13168.
 24. Li G, Sajobi TT, Menon BK, et al. Registry-based randomized controlled trials- what are the advantages, challenges, and areas for future research? *J Clin Epidemiol* 2016;80:16-24.
 25. James S, Rao SV, Granger CB. Registry-based randomized clinical trials--a new clinical trial paradigm. *Nat Rev Cardiol* 2015;12:312-6.

Cite this article as: Godown J, Gaies M, Wilkinson JD. Leveraging big data to advance knowledge in pediatric heart failure and heart transplantation. *Transl Pediatr* 2019;8(4):342-348. doi: 10.21037/tp.2019.07.09