Published in final edited form as:

Prog Pediatr Cardiol. 2012 January ; 33(1): 21–24. doi:10.1016/j.ppedcard.2011.12.004.

Opportunities and challenges in linking information across databases in pediatric cardiovascular medicine

Sara K. Pasqual, M.D., M.H.S.¹, Jennifer S. Li, M.D., M.H.S.², Marshall L. Jacobs, M.D.³, Samir S. Shah, M.D., MSCE⁴, and Jeffrey P. Jacobs, M.D.⁵

¹Department of Pediatrics, Duke University School of Medicine, and the Duke Clinical Research Institute, Duke University Medical Center, Durham, NC

²Department of Pediatrics, Duke University School of Medicine, and the Duke Clinical Research Institute, Duke University Medical Center, Durham, NC

³Department of Pediatric and Congenital Heart Surgery, Cleveland Clinic, Cleveland, OH

⁴Divisions of Infectious Diseases and General and Community Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH

⁵Division of Thoracic and Cardiovascular Surgery, Congenital Heart Institute of Florida, All Children's Hospital and Children's Hospital of Tampa, University of South Florida College of Medicine, St. Petersburg and Tampa, FL

Abstract

Multicenter databases are increasingly utilized in pediatric cardiovascular research. In this review, we discuss the rationale for using these types of data sources, provide several examples of how large datasets have been utilized in clinical research, and describe different mechanisms for linking databases to enable studies not possible with individual datasets alone.

Keywords

congenital heart disease; outcomes research

Introduction

Multicenter clinical registries and administrative databases have become increasingly utilized in pediatric cardiovascular research. There are now multiple different databases in the field, and their development and utilization in a variety of analyses has helped to overcome some of the historic barriers to conducting pediatric research. In this review, we

© 2012 Elsevier Ireland Ltd. All rights reserved.

Corresponding Author: Sara K. Pasqual, M.D., M.H.S., Associate Professor of Pediatrics, Division of Cardiology, Department of Pediatrics, Duke University Medical Center, Duke Clinical Research Institute, PO Box 17969, Durham, NC 27715, Phone: 919-668-4686, Fax: 919-668-7058, sara.pasquali@duke.edu.

Disclosures:

Dr. Pasquali: Grant support (1K08HL103631-01), National Heart, Lung, and Blood Institute

Dr. Shah: Grant support National Institute of Allergy and Infectious Diseases (K01 AI73729), and Robert Wood Johnson Foundation Physician Faculty Scholar program.

Dr. J. Jacobs: Chair, Society of Thoracic Surgeons Congenital Heart Surgery Database Task Force

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

discuss the rationale for using these types of data sources, several examples of how different databases have been utilized in clinical research, as well as different mechanisms for linking databases to enable studies not possible with individual datasets alone.

Background

In pediatric cardiovascular medicine, evidence to guide optimal care is often limited. Historically, there have been many barriers to conducting clinical research including the relative rarity and heterogeneity of disease, limited research infrastructure, cost and ethical issues surrounding large trials, and difficulty in identifying valid endpoints (1,2). Instead of clinical trial data, most treatment decisions are based on small observational studies, extrapolation from adult data, or clinical experience. A recent survey evaluating over 1000 clinical decisions made by 10 pediatric cardiologists during a 1 week period illustrates this point. Investigators found that less than 3% of clinical decisions were based on a research study specific to the pertinent clinical question at hand (3).

There are several reasons why this approach to care may not be ideal. First, with limited evidence to guide the care of children, there are few best practice guidelines and significant variation in all aspects of care from center to center. Numerous analyses of multicenter databases and registries have shown wide variation extending to multiple aspects of in-hospital and follow-up care, including variation in models of peri-operative ICU care, variation in the use of various medications including peri-operative corticosteroids and antifibrinolytics, variation in the way procedures are performed including the use of delayed sternal closure, variation in the proportion of patients enrolled in home monitoring programs following hospital discharge, as well as variation in routine medications used post-operatively in the outpatient setting (4–10). In addition to variation in practice, multiple reports have also demonstrated that outcomes also vary widely from center to center (11–13). For example, Jacobs and colleagues evaluated 2342 infants from 69 centers undergoing the Norwood operation in the Society of Thoracic Surgeons (STS) Congenital Heart Surgery Database from 2005–2009 and found that institutional adjusted mortality rates varied from 7% to 42% (12). A second problem related to the limited pediatric evidence base is that off label use of drugs and devices without adequate pediatric efficacy and safety data is common (14,15). It has been shown that >75% of pediatric inpatients with cardiovascular disease receive one or more medications off-label and 31% receive 3 off-label medications (15). Inadequate information on drug efficacy and safety may deny children potential therapeutic benefits, and can also place them at risk for adverse events. Third, children are not “tiny adults” and adult data cannot simply be extrapolated to the pediatric population (16–18). For example, the optimal dose of clopidogrel in infants and children with a cardiac condition at risk for arterial thrombosis was found to be only one fifth of what would be given if extrapolating from adult data (16).

Leveraging Multi-center Databases in Pediatric Cardiovascular Research

While clinical trials are generally considered the gold standard for generating evidence and evaluating new therapies, there are also certain limitations to consider (17,18). Clinical trials are not always feasible for every research question. Trials can be expensive and time consuming, and can have limited generalizability. In addition, several trials that have been performed in the congenital heart disease population have been limited by baseline knowledge gaps and have failed scientifically (17,18). These knowledge gaps include natural history of disease, baseline event rates, baseline treatment rates, and available sample size. Thus, our group and others have hypothesized that analysis and linkage of existing large pediatric data sources can help to address many of the current knowledge gaps, and limitations of current methodology and infrastructure, and provide data to better plan future

studies (19). Leveraging multi-center databases and registries in pediatric cardiovascular research can provide adequate power to answer important clinical questions and overcome the relative rarity and heterogeneity of many conditions. These data sources also allow us to describe areas of significant practice variation between centers, and evaluate how this variation impacts outcome. This type of evaluation can help to identify areas of focus for future prospective study and may be a first step toward defining best practices. Many types of outcomes and comparative effectiveness studies can be conducted using existing pediatric data sources, and these analyses can also provide data on sample sizes, event rates, and natural history, which may aid in better planning future studies.

There are multiple different registries and administrative databases which collect information on pediatric cardiac patients and have become increasingly utilized in clinical research. Some of the largest databases include:

1. The STS Congenital Heart Surgery Database (collects data on pediatric cardiac surgery)
2. The Extracorporeal Life Support Organization (ELSO) Registry (collects data on extracorporeal life support)
3. The Virtual Pediatric Intensive Care Unit System (VPS) Database (collects intensive care unit data)
4. The American College of Cardiology Improving Pediatric and Adult Congenital Treatment (IMPACT) Database (collects cardiac catheterization data)
5. The Pediatric Heart Transplant Study Group Database (collects transplant data).

This list is not meant to be comprehensive and there are several additional clinical registries in the field (20–23). In addition to these registries, there are also several administrative databases (containing information derived from the hospital bill) that have been used in pediatric cardiovascular research. These include the Pediatric Health Information Systems (PHIS) database, and several Agency for Healthcare Research and Quality Healthcare Cost and Utilization Project (HCUP) databases, including the Kids' Inpatient Database (KID) and others (19,24). These clinical and administrative databases have been used to conduct a variety of different studies, including: evaluating center-level factors such as the relationship between center volume and outcome and assessing center variation in outcomes for benchmark operations, assessing outcome in certain patient populations such as adults with congenital heart disease, reporting outcomes following repair of rare lesions such as truncus arteriosus, evaluating the impact of pre-operative factors such as presence of various genetic syndromes factors on outcome, studying the association of peri-operative medications with outcome, as well as assessing the impact of postoperative complications on outcome, to name a few (7,8,11,12,25–32).

Thus, multi-center databases and registries can be useful in conducting a variety of clinical studies. However, there are also certain limitations to the use of these databases. There are relatively few existing datasets, each containing limited data, and the databases do not communicate with each other. Analyses of administrative databases can be limited as these data sources rely on International Classification of Diseases, Ninth Revision (ICD-9) diagnosis and procedure codes to identify patients undergoing congenital heart surgery which may have limited accuracy in some cases (33,34). For example, there is no ICD-9 code for the Norwood operation, such that a combination of other codes must be used, the validity of which is not known. Finally, there is no current mechanism for efficient real-time data collection to answer important clinical questions that may arise.

Rationale for Linking Databases

We have hypothesized that linkage of existing pediatric data sources can address many of the limitations associated with the use of individual datasets alone. Linking databases expands the pool of available data for analysis and capitalizes on the strengths of different types of data sources. Linkage allows analyses otherwise not possible with single center data or individual datasets alone. In the next section of the review, we will discuss a variety of different methodologies that can be used to link data.

Mechanisms for Linking Data

Linking on Indirect Identifiers

Direct, or unique, identifiers such as social security number have traditionally not been collected or readily available for analysis in many databases due to a variety of regulatory requirements and concerns (35). Thus, methodology has been developed to link databases through the use of “indirect” identifiers (36). These include date of birth, date of admission, date of discharge, sex, and center where hospitalized. It has been shown in the adult population that nearly all records at a given center can be uniquely identified using these indirect identifiers, and that a crosswalk can then be created between two datasets, linking patients on the values of center where hospitalized and the indirect identifiers (36) (Figure 1). This method has been used to successfully link adult cardiac databases, but had not been used in the pediatric population to date (36–38).

We set out to use this methodology for the first time in children to link the STS Congenital Heart Surgery Database with the PHIS Database (19). The STS Congenital Heart Surgery Database is the largest clinical pediatric heart surgery registry in the world. It currently contains data on >180,000 surgeries performed since 1998 at >100 North American centers, and represents more than three quarters of all US centers performing congenital heart surgery (39). This clinical registry contains pre-operative, operative, and outcomes data on all patients undergoing congenital heart surgery at participating centers. Diagnoses and procedures are coded using the International Pediatric and Congenital Cardiac Code, which was developed through an international collaborative effort of pediatric cardiologists and congenital heart surgeons (40). Data quality and reliability are assured through intrinsic verification of data as well as a formal process of site visits and data audits (41). The STS Database does not currently collect medication data or other resource utilization data. The PHIS Database is a large administrative database containing inpatient data from 41 free-standing not for profit childrens hospitals in the US affiliated with the Child Health Corporation of America (19). The database currently contains information from >4.6 million inpatient discharges. Data quality and reliability are assured through a joint effort between Child Health Corporation of America and participating hospitals. Data are collected from the hospital bill, including demographics, ICD-9 diagnoses and procedure codes, in-patient outcomes data, and resource utilization data including billing codes for medications, imaging, and laboratory studies, and hospital charges.

As displayed in Figure 2, linking these two datasets allows us to utilize the detailed diagnosis, procedure, and operative data from the STS Congenital Heart Surgery Database, and the resource utilization data from the PHIS Database, including data on medications and hospital charges. Both databases contain useful information regarding various patient and center factors and outcomes. All of these data can be pooled together for analysis, and enable studies not possible using either dataset alone.

From 2004–2008, 30 US centers submitted data to both the STS Congenital Heart Surgery Database and the PHIS Database, accounting for approximately 45,000 admissions and

encompassing a wide range of center surgical volume and geographic area of the country (19). Through using the methodology of linking on indirect identifiers, we successfully linked data on 90% of 45,830 eligible patients, and validated the link in a 10% sample (19). We have subsequently updated the link through 2010 such that linked data is now available for more than 60,000 patients from 33 US centers. The linked data are currently being used to assess practice variation and perform comparative analyses of various medications utilized in the congenital heart surgery population including peri-operative corticosteroids, anti-fibrinolytic medications, and antibiotic prophylaxis. Subsequent studies will involve cost analyses and evaluation of differences in case ascertainment and outcomes assessment in the two datasets.

Linking on Unique Identifiers

Some datasets do contain unique identifiers such as social security number and this information can facilitate further linkages beyond those performed through linking on indirect identifiers. For example, the STS Adult Cardiac Surgery Database recently began collecting unique identifiers including social security number which has enabled linking of this database to the Social Security Death Master File in order to evaluate long-term mortality (38,42). Unique identifiers were also recently incorporated into the STS Congenital Heart Surgery Database and this may facilitate similar linkages in the future to assess long-term mortality in this cohort also. This methodology has also been successfully used by other investigators to link outpatient records regarding pediatric cardiology visits for chest pain to the National Death Index and Social Security Death Master File to evaluate for subsequent mortality in this cohort (43).

Center-level Linkages

Linking registry data to other center level data through matching on center is also possible. For example, a survey regarding different center-level models of ICU care was successfully linked to the STS Congenital Heart Surgery Database. This linkage enabled evaluation of the association of the variables collected in the survey with outcomes data collected in the STS Database (44).

Modules

Linking data can also be accomplished through the development and use of a modular data collection system that enables collection of supplemental datapoints to the main registry. The modules are web-based and can be quickly created and deployed to allow “real-time” data collection to answer important clinical questions. They are more time and cost-efficient compared with traditional data collection methods that may duplicate data already being collected in the main registry. This methodology has been used in the STS Adult Cardiac Surgery Database to design a module to collect detailed information on atrial fibrillation, which is linked to the rest of the data collected in the main registry. Modules are currently being designed for use with the STS Congenital Heart Surgery Database to collect data to pilot test several proposed quality measures (45).

Collaboration/Partnering Between Databases

Data can also be shared or linked through collaboration/partnering between different subspecialty organizations and databases. For example, beginning in 2010 the Society of Thoracic Surgeons partnered with the Congenital Cardiac Anesthesia Society to add a new anesthesia section to the STS data collection forms (22). Anesthesia data are now collected, harvested, reported, and analyzed along with surgical data for participating centers. This approach proved to be more time and cost efficient than creating a separate anesthesia

database in which many of the fields regarding patient characteristics and the operative procedure would have been duplicated between databases.

Conclusions

In summary, leveraging multi-center databases in pediatric cardiovascular research can provide a variety of useful data. Linking databases is also possible and can involve a variety of different methodologies. Linkage capitalizes on the strengths, and mitigates some of the weaknesses, of different types of data sources and can allow analyses not otherwise possible with single center data or individual datasets alone. Further development of this methodology may facilitate analyses not only of isolated outcomes associated with a particular operation, device, or medication, but enable longitudinal evaluation of the child diagnosed with congenital heart disease who may receive multiple different therapies during their lifetime.

References

1. Sanders SP. Conducting pediatric cardiovascular trials. *Am Heart J.* 2001; 142:218–223. [PubMed: 11479457]
2. Mahony L, Sleeper LA, Anderson PA, et al. The Pediatric Heart Network: a primer for the conduct of multicenter studies in children with congenital and acquired heart disease. *Pediatr Cardiol.* 2006; 27:191–8. [PubMed: 16261271]
3. Darst JR, Newburger JW, Resch S, Rathod RH, Lock JE. Deciding Without Data. *Conget Heart Dis.* 2010; 5:339–342.
4. Wernovsky G, Ghanayem N, Ohye RG, et al. Hypoplastic left heart syndrome: consensus and controversies in 2007. *Cardiol Young.* 2007; 17:S75–86.
5. Burstein D, Rossi A, Jacobs J, et al. Variation in models of care delivery for children undergoing congenital heart surgery in the United States. *World Journal for Pediatric and Congenital Heart Surgery.* 2010; 1:8–14.
6. Johnson JN, Ansong AA, Li JS, et al. Celiac artery flow pattern in infants with single right ventricle following the Norwood procedure with modified Blalock-Taussig or right ventricle to pulmonary artery shunt. *Pediatr Cardiol.* 2011; 32:479–483. [PubMed: 21331516]
7. Pasquali SK, Hall M, Li JS, et al. Corticosteroids and outcome in children undergoing congenital heart surgery: Analysis of the Pediatric Health Information Systems Database. *Circulation.* 2010; 122:2123–2130. [PubMed: 21060075]
8. Pasquali SK, Li JS, He X, et al. Comparative analysis of antifibrinolytic medications in pediatric heart surgery. *Journal of Thoracic and Cardiovascular Surgery.* In press.
9. Ohye RG, Pearson GD, Lu M, et al. for the Pediatric Heart Network Investigators. Variation in perioperative management of Norwood surgery in the Pediatric Heart Network Single Ventricle Reconstruction Trial. *J Am Coll Cardiol.* 2011; 57:E428. (abstract).
10. Anderson PA, Breitbart RE, McCrindle BW, et al. for the Pediatric Heart Network Investigators. The Fontan patient: inconsistencies in medication therapy across seven pediatric heart network centers. *Pediatr Cardiol.* 2010; 31:1219–1228. [PubMed: 20938655]
11. Welke KF, O'Brien SM, Peterson ED, et al. The complex relationship between pediatric cardiac surgical case volumes and mortality rates in a national clinical database. *J Thorac Cardiovasc Surg.* 2009; 137:1133–1140. [PubMed: 19379979]
12. Jacobs JP, O'Brien SM, Pasquali SK, et al. Richard E. Clark Paper: Variation in Outcomes for Benchmark Operations: An analysis of the STS Congenital Heart Surgery Database. *Annals of Thoracic Surgery.* In press.
13. Pasquali SK, Jacobs JP, He X, et al. The complex relationship between center volume and outcome in patients undergoing the Norwood operation. *Annals of Thoracic Surgery.* In press.
14. Shah SS, Hall M, Goodman DM, et al. Off-label drug use in hospitalized children. *Arch Pediatr Adolesc Med.* 2007; 161:282–290. [PubMed: 17339510]

15. Pasquali SK, Hall M, Slonim AD, et al. Off-label use of cardiovascular medications in children hospitalized with congenital and acquired heart disease. *Circ Cardiovasc Qual Outcomes*. 2008; 1:74–83. [PubMed: 20031793]
16. Li JS, Yow E, Berezny KY, Bokesch PM, et al. PICOLO Investigators. Dosing of clopidogrel for platelet inhibition in infants and young children: primary results of the Platelet Inhibition in Children On cLOpidogrel (PICOLO) trial. *Circulation*. 2008; 117:553–559. [PubMed: 18195173]
17. Li JS, Cohen-Wolkowicz M, Pasquali SK. Pediatric clinical trials: Lessons learned. *J Cardiovasc Pharmacol*. 2011 Jan 14. [Epub ahead of print].
18. Li JS, Colan SD, Sleeper LA, et al. Lessons learned from a pediatric clinical trial: The Pediatric Heart Network Angiotensin-Converting Enzyme Inhibition in Mitral Regurgitation Study. *Am Heart J*. 2011; 161:233–240. [PubMed: 21315203]
19. 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. *American Heart Journal*. 2010; 160:1099–1104. [PubMed: 21146664]
20. Jacobs ML, Jacobs JP, Franklin RCG, et al. Databases for assessing the outcomes of the treatment of patients with congenital and paediatric cardiac disease – the perspective of cardiac surgery. *Cardiol Young*. 2008; 18(Suppl 2):101–115. [PubMed: 19063780]
21. Jenkins KJ, Beekman RH III, Bergersen LJ, et al. Databases for assessing the outcomes of the treatment of patients with congenital and paediatric cardiac disease – the perspective of cardiology. *Cardiol Young*. 2008; 18(Suppl 2):116–123. [PubMed: 19063781]
22. Vener DF, Jacobs JP, Schindler E, Maruszewski B, Andropoulos D. Databases for assessing the outcomes of the treatment of patients with congenital and paediatric cardiac disease – the perspective of anaesthesia. *Cardiol Young*. 2008; 18(Suppl 2):124–129. [PubMed: 19063782]
23. LaRovere JM, Jeffries HE, Sachdeva RC, 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):130–136. [PubMed: 19063783]
24. Healthcare Cost and Utilization Project. [Accessed October 17, 2011] Available at: <http://www.hcup-us.ahrq.gov/>
25. Welke KF, Diggs BS, Karamlou T, Ungerleider RM. The relationship between hospital surgical case volumes and mortality rates in pediatric cardiac surgery: A national sample, 1988–2005. *Ann Thorac Surg*. 2008; 86:889–896. [PubMed: 18721578]
26. Karamlou T, Diggs BS, Person T, Ungerleider RM, Welke KF. National practice patterns for management of adult congenital heart disease: operation by pediatric heart surgeons decreased in-hospital death. *Circulation*. 2008; 118:2345–2352. [PubMed: 18997167]
27. Mascio CE, Pasquali SK, Jacobs JP, Jacobs ML, Austin EH. Outcomes in adult congenital heart surgery: Analysis of the Society of Thoracic Surgeons Database. *Journal of Thoracic and Cardiovascular Surgery*. 2011 Sep 10. [Epub ahead of print].
28. Russell H, Pasquali SK, Jacobs JP, et al. Outcomes of repair of common arterial trunk with truncal valve surgery: A review of the STS Congenital Heart Surgery Database. *Annals of Thoracic Surgery*. In press.
29. Fudge JC, Li S, Jaggars J, et al. Outcomes in patients with Down syndrome undergoing congenital heart surgery: Analysis of a national clinical database. *Pediatrics*. 2010; 126:315–322. [PubMed: 20624800]
30. Patel A, Hickey E, Mavroudis C, et al. Impact of noncardiac congenital and genetic abnormalities on outcomes in hypoplastic left heart syndrome. *Ann Thorac Surg*. 2010; 89:1805–1813. [PubMed: 20494032]
31. Hornik CP, He X, Jacobs JP, et al. Complications after the Norwood operation: An analysis of the STS Congenital Heart Surgery Database. *Annals of Thoracic Surgery*. 2011 Sep 19. [Epub ahead of print].
32. Pasquali SK, Li JS, Burstein DB, et al. The association of center volume with mortality and complications in pediatric heart surgery. *Pediatrics*. In press.
33. Strickland MJ, Riehle-Colarusso TJ, Jacobs JP, et al. The importance of nomenclature for congenital cardiac disease: implications for research and evaluation. *Cardiol Young*. 2008; 18(Suppl 2):92–100. [PubMed: 19063779]

34. Cronk CE, Malloy ME, Pelech AN, et al. Completeness of state administrative databases for surveillance of congenital heart disease. *Birth Defects Res A Clin Mol Teratol*. 2003; 67:597–603. [PubMed: 14703780]
35. Dokholyan RS, Muhlbaier LH, Falletta J, et al. Regulatory and Ethical Considerations for Linking Clinical and Administrative Databases. *Am Heart J*. 2009; 157(6):971–982. [PubMed: 19464406]
36. Hammill BG, Hernandez AF, Peterson ED, Fonarow GC, Schulman KA, Curtis LH. Linking inpatient clinical registry data to Medicare claims data using indirect identifiers. *Am Heart J*. 2009; 157:995–1000. [PubMed: 19464409]
37. Jacobs JP, Edwards FH, Shahian DM, et al. Successful Linking of The Society of Thoracic Surgeons Adult Cardiac Surgery Database to Centers for Medicare and Medicaid Services Medicare Data. *Ann Thorac Surg*. 2010; 90:1150–1157. [PubMed: 20868806]
38. Jacobs JP, Edwards FH, Shahian DM, et al. Successful Linking of the STS Database to Social Security Data to Examine Survival after Cardiac Surgery. *Ann Thorac Surg*. 2011; 92:32–37. [PubMed: 21718828]
39. Jacobs ML, Daniel M, Mavroudis C, et al. Report of the 2010 Society of Thoracic Surgeons Congenital Heart Surgery Practice and Manpower Survey. *Ann Thorac Surg*. 2011; 92:762–768. [PubMed: 21865107]
40. Franklin RC, Jacobs JP, Krogmann ON, et al. Nomenclature for congenital and paediatric cardiac disease: historical perspectives and The International Pediatric and Congenital Cardiac Code. *Cardiol Young*. 2008; 18(Suppl 2):70–80. [PubMed: 19063777]
41. Clarke DR, Breen LS, Jacobs ML, et al. Verification of data in congenital cardiac surgery. *Cardiol Young*. 2008; 18 (Suppl2):177–187. [PubMed: 19063789]
42. Jacobs JP, Haan CK, Edwards FH, et al. The rationale for incorporation of HIPAA compliant unique patient, surgeon, and hospital identifier fields in the STS Database. *Ann Thorac Surg*. 2008; 86:695–698. [PubMed: 18721549]
43. Saleeb SF, Li WYV, Warren SZ, Lock JE. Effectiveness of screening for life-threatening chest pain in children. *Pediatrics*. 10.1542/peds.2011-0408
44. Burstein DS, Jacobs JP, Sheng S, et al. Care models in congenital heart surgery and associated outcomes. *Pediatrics*. 2011; 127(6):e1482–e1489. [PubMed: 21576309]
45. Jacobs JP, Jacobs ML, Austin EH, et al. Quality Measures for Congenital and Pediatric Cardiac Surgery. *World Journal for Pediatric and Congenital Heart Surgery*. In press.

Figure 1. Linking on Indirect Identifiers

This figure describes how information on patient “X” entered into Database A and B may be linked through the use of indirect identifiers. This patient may have data on medications received during a hospitalization, for example, along with their indirect identifiers entered into Database A. The same patient may also have data on their cardiac surgery, for example, in addition to their indirect identifiers, entered into Database B. All of the data fields in Database A are tied together into one record, represented by the gray links. This occurs similarly in Database B. It is for this reason that linking the data on indirect identifiers on patient X in Database A to those in Database B (dashed arrows) results in linkage of their entire data record from Database A to Database B (solid arrow), such that all of their data from both datasets can be pooled and analyzed together.

Figure 2. STS-CHS and PHIS Database Link

Through linking the STS-CHS (Society of Thoracic Surgeons Congenital Heart Surgery) Database and PHIS (Pediatric Health Information Systems) Database, we are able to link the detailed diagnosis and operative data from the STS-CHS Database, with the resource utilization data from the PHIS Database (including medication administration and hospital charges). Both databases contain useful information regarding other patient and center factors, and outcomes. All these data may be pooled together for analysis, which enables studies not possible using either dataset alone.