Published in final edited form as:

Cardiol Young. 2016 August ; 26(6): 1090–1096. doi:10.1017/S1047951115001833.

Data integrity of the Pediatric Cardiac Critical Care Consortium (PC⁴) clinical registry

Michael Gaies¹, Janet E. Donohue², Gina M. Willis², Andrea T. Kennedy³, John Butcher⁴, Mark A. Scheurer⁵, Jeffrey A. Alten⁶, J. William Gaynor⁷, Jennifer J. Schuette⁸, David S. Cooper⁴, Jeffrey P. Jacobs⁹, Sara K. Pasquali¹, and Sarah Tabbutt¹⁰

¹Department of Pediatrics and Communicable Diseases, Division of Cardiology, C.S. Mott Children's Hospital and University of Michigan Medical School

²Michigan Congenital Heart Outcomes Research and Discovery (MCHORD) Unit, University of Michigan Congenital Heart Center, Ann Arbor, Minnesota, United States of America

³The Cardiac Center, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, United States of America

⁴Heart Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, United States of America

⁵Department of Pediatrics, Division of Cardiology, Medical University of South Carolina, Charleston, South Carolina, United States of America

⁶Department of Pediatrics, Division of Critical Care, University of Alabama at Birmingham, Birmingham, Alabama, United States of America

⁷Division of Pediatric Cardiac Surgery, Department of Surgery, The Cardiac Center, The Children's Hospital of Philadelphia, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, United States of America

⁸Department of Critical Care Medicine and Cardiology, Children's National Medical Center, Washington, District of Columbia, United States of America

⁹Department of Surgery, Division of Cardiac Surgery, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America

¹⁰Department of Pediatrics, Benioff Children's Hospital, University of California San Francisco School of Medicine, California, United States of America

Abstract

Correspondence to: M. Gaies, MD, MPH, Pediatric Cardiac Critical Care Consortium (PC⁴), C.S. Mott Children's Hospital, University of Michigan Congenital Heart Center, 1540 E. Hospital Drive, Ann Arbor, MI 48109-4204, United States of America. Tel: + 734 883 2986; Fax: + 734 936 9470; mgaies@med.umich.edu.

Conflicts of Interest

None.

Supplementary material

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/S1047951115001833>

Background—Clinical databases in congenital and paediatric cardiac care provide a foundation for quality improvement, research, policy evaluations and public reporting. Structured audits verifying data integrity allow database users to be confident in these endeavours. We report on the initial audit of the Pediatric Cardiac Critical Care Consortium (PC⁴) clinical registry.

Materials and methods—Participants reviewed the entire registry to determine key fields for audit, and defined major and minor discrepancies for the audited variables. In-person audits at the eight initial participating centres were conducted during a 12-month period. The data coordinating centre randomly selected intensive care encounters for review at each site. The audit consisted of source data verification and blinded chart abstraction, comparing findings by the auditors with those entered in the database. We also assessed completeness and timeliness of case submission. Quantitative evaluation of completeness, accuracy, and timeliness of case submission is reported.

Results—We audited 434 encounters and 29,476 data fields. The aggregate overall accuracy was 99.1%, and the major discrepancy rate was 0.62%. Across hospitals, the overall accuracy ranged from 96.3 to 99.5%, and the major discrepancy rate ranged from 0.3 to 0.9%; seven of the eight hospitals submitted >90% of cases within 1 month of hospital discharge. There was no evidence for selective case omission.

Conclusions—Based on a rigorous audit process, data submitted to the PC⁴ clinical registry appear complete, accurate, and timely. The collaborative will maintain ongoing efforts to verify the integrity of the data to promote science that advances quality improvement efforts.

Keywords

Cardiac; critical care; database; audit; paediatric

Clinical registries represent important sources of data on practice and outcomes across multiple clinical fields, but particularly so in pediatric and congenital cardiac care where administrative data often fails to accurately capture detailed anatomic and procedural information.^{1,2} Clinical databases provide an ideal foundation for quality improvement initiatives, research activities, policy decisions, and public reporting.^{1–3} In this context, it is important for these registries to maintain high standards of completeness and accuracy for the clinicians, researchers, and public who use the information.

The Pediatric Cardiac Critical Care Consortium (PC⁴) is a quality improvement collaborative of hospitals from North America who voluntarily participate and submit clinical data on patients with critical cardiovascular disease cared for in cardiac ICUs.⁴ The mission of PC⁴ is to provide benchmarks of quality to participants through clinical outcome reporting, to identify variation in performance between hospitals, and to use the data as a catalyst to initiate quality improvement efforts that will lead to better outcomes for these critically ill patients.

Data collection officially began at eight institutions between October 15, 2013 and January 1, 2014. A detailed audit was performed at each of these institutions during the 1st year of data collection. In this manuscript, we describe the audit methodology and report the findings from the audit to provide evidence regarding the integrity of the data in the PC⁴ clinical registry.

Materials and methods

The clinical registry and data collection processes

An overview of the variables in the PC⁴ registry and the database features were described previously.⁴ Before data submission, the data coordinating centre staff perform on-site training with the local data collection team members over 2 days, including reviewing data definitions and data entry processes and assistance with the submission of multiple cases. All data collectors must pass a certification exam testing knowledge of the database variables and coding rules initially and after every software version change. PC⁴ conducts weekly teleconferences with all data collection teams and hospital clinical champions to review definitional questions and ensure consistency of interpretation across sites; the data coordinating centre circulates content from these calls to all sites. The collaborative also hosts all data collection teams at an annual in-person meeting to further refine these processes.

The registry contains information on all encounters where a patient with cardiovascular disease is admitted to any critical care unit under the care of the attending cardiac intensivist. The database is relational in design and includes patient, hospitalisation, and intensive care encounter variables. Detailed information is collected and date-stamped related to critical care interventions and outcomes specific to cardiac critical care, including complications. More than 90% of the fields in the registry are mandatory for case submission. A robust system of internal data checks was built into the data entry software to prevent missing, illogical, or implausible data to pass from the site to the data warehouse; all errors are flagged and must be resolved by the data collector before the case can be successfully submitted.

Several variables with common terminology and definitions are shared between the PC⁴ registry, the Society of Thoracic Surgeons Congenital Heart Surgery Database, the Joint Congenital Cardiac Anesthesia Society–Society of Thoracic Surgeons Congenital Cardiac Anesthesia Database, and the American College of Cardiology IMPACT Registry®; the nomenclature of the International Pediatric and Congenital Cardiac Code is used in all of these databases.⁵ The data entry software allows sharing of these common fields at the local level to avoid duplicate data collection and to minimise the burden of data entry. For the purpose of maintaining consistency across databases, PC⁴ data collection teams are instructed to use the fundamental diagnosis, procedural diagnosis, and define all procedures according to how they are coded in the Society of Thoracic Surgeons Congenital Heart Surgery Database.

Audit content

The Data Coordinating Centre analytic team and the PC⁴ Database and Executive Committees reviewed the entire registry to determine the key fields for audit and definitions of major and minor discrepancies for the audited variables. The list of these fields and definitions of major and minor discrepancies are shown in Appendix A. Shared fields with other registries are not included in the audit as these fields are already subject to review by other entities – for example, The Society of Thoracic Surgeons. We did evaluate whether the

dates of all cardiothoracic surgeries were coded correctly to ensure that surgical cases are not missing in the PC⁴ database and that they are appropriately timed in relation to therapies and complications.

Audit process

The overarching approach to the maintenance of data integrity focusses on two main domains: data accuracy and timely submission of completed cases. In-person audits began in January 2014; three members from the Data Coordinating Centre (M.G., J.E.D., and G.M.W.) audited seven sites. The clinical champion (S.T.) and data collector from the University of California-San Francisco performed the University of Michigan audit. The audit was carried out over 2 days. Auditors were granted access to the medical records by the participant institution for the purpose of quality improvement review. Before arrival, the data coordinating centre generated a randomly selected sample of intensive care encounters. A patient could have more than one encounter in the sample.

Data accuracy was assessed through source data verification at the participant site. The planned audit included review of 62 encounters at each site; this number was chosen because we predicted that this was the maximum number of cases that could be completed over a 1.5-day audit. This case number represents a 5–15% sample of yearly cases depending on the centre volume. The audit team aims to perform blind abstraction of the entire data set included in the audit for two encounters, and on the sample of key fields chosen by the Database Committee necessary for reporting outcomes and performing risk-adjustment on 20 encounters. In addition, the team performs source data verification on the same key fields with the data collector on up to 40 additional encounters. As the system was learnt over the first three audits, the team reviewed only 75% of the maximum number of planned encounters. Process improvements were enacted to increase efficiency, but specific circumstances prevented the team from reviewing some records in the sample at the remaining sites.

In addition, two auditors (M.G. and G.M.W.) completed the blind chart abstraction using the electronic medical record to obtain values for each data variable without knowledge of how those fields were originally coded by the participant data collection team. We then compared the results of this blind abstraction with the record in the data warehouse and identified discrepancies. Source data verification was performed one-on-one with the data collector at each site. The data collector opened the medical record for selected cases and reported a value for each field in the audit. A member of the audit team (J.E.D.) compared responses with the values in the data warehouse for that case. Discrepancies were recorded and ultimately resolved through a secondary review. The discrepancies were reported to the local data collection team, who subsequently verified whether there was a true discrepancy or an error in abstraction by the auditor. After completion of this adjudication process, a final quantitative report was released to the site.

PC⁴ assesses completeness and timeliness of case submission by comparing all the cases submitted to the data coordinating centre at the time of audit to a third-party source document. This is provided by the participant and includes all admissions to the cardiac ICU over the same time period as the audit cases. This list of admissions has to be from a source

other than that maintained by the data collection team – for example, hospital billing or census records. Discrepancies between these lists were reviewed during the in-person audit. The expectation is that all cases are submitted to the data coordinating centre within 1 month of the hospital discharge date.

Quantitative analysis

All audited fields were included in the analysis; however, if an error was made on a data field with a “parent–child relationship” to other fields, we removed the “child fields” from calculations in order to avoid double counting discrepancies – for example, if a central venous line was not recorded, we did not count errors for all the other fields that would have been related to that catheter had it been recorded: start/stop time, site of insertion, location of insertion, etc. In the extremely rare instance where a software data check allowed an implausible value to enter the data warehouse – for example, illogical dates – we did not count the discrepancy, and the denominator of adjudicated fields was adjusted to reflect these situations. If such an error was discovered, the software vendor immediately updated the data check and re-submitted cases. Finally, if the audit team discovered a systematic error in data entry, such as misinterpretation of a variable definition, this field was removed from the entire audit accuracy calculation and the cases were corrected and re-submitted with the correct data by the local data collection team.

We calculated two different accuracy metrics. Overall accuracy was calculated as follows:

$$\frac{\text{Fields without a major or minor discrepancy}}{\text{Total fields audited}} \times 100$$

We then calculated a major discrepancy rate; as not all fields could have an associated major discrepancy, we only included fields in the denominator where there was a potential major discrepancy:

$$\frac{\text{Fields with major discrepancy}}{\text{Fields audited with potential for a major discrepancy}} \times 100$$

In addition to calculating aggregate overall accuracy and major discrepancy rate for the entire group, we determined these rates for each hospital. We then calculated overall accuracy and major discrepancy rate within specific domains of the registry in aggregate and by site: patient demographics, hospital-level and intensive care encounter-level information, cardiac surgical data – number and dates of procedures – devices and therapies and complications.

We also calculated the timeliness of case submission by comparing the number of cases in the data warehouse with the document provided at the time of audit and reported the percentage of encounters submitted within 1 month of the hospital discharge date. We also performed a simple count of patients not appearing in the data warehouse whose hospitalisation ended in mortality.

Results

Across eight hospitals, we audited 434 encounters and 29,476 individual data fields. At the first three centres, 41, 45, and 44 encounters were reviewed, respectively. As the efficiency of the process improved, no less than 58 encounters were reviewed at each of the five subsequent sites.

Accuracy

Results of the audit are shown in aggregate in Table 1 and broken down according to the registry domains in Table 2. The aggregate overall accuracy was 99.1%, 255 total discrepancies. By centre, rates of overall accuracy ranged from 98.9 to 99.5%. The overall accuracy within each domain of the registry ranged from 97.5% to 99.7%.

The aggregate major discrepancy rate across hospitals was 0.62%, 121 major discrepancies, with 19,639 fields eligible for a major discrepancy. The major discrepancy rates within the individual registry domains ranged from 0.33% for complications to 1.1% for patient demographics. The major discrepancy rates for the individual hospitals ranged from 0.3 to 0.92%.

Timeliness and completeness of case submission

The percentage of cases submitted to the data warehouse within 1 month of hospital discharge ranged from 80.2 to 98.8% across centres. Of the eight hospitals, seven submitted >90% of their cases within 1 month of discharge. The one centre with 80% submission had recently lost their data manager; an analysis of missing cases revealed that all of them were in the local database but had not yet been submitted to the data warehouse. When analysing cases that did not appear in the data warehouse at each site, five records were those of patients whose hospital stays ended in mortality. These patients' records existed in the local database but had not yet been submitted due to incomplete data fields. Since the audit, the data coordinating centre has confirmed the submission of all cases that were outstanding at the time of the audit.

Discussion

Results from the 1st year audit at the initial eight hospitals submitting data to the PC⁴ clinical registry demonstrate excellent accuracy with no evidence of purposeful omission of cases. These data compare favourably with those reported by investigators from congenital cardiac registries and others.^{6–10} The integrity of the data evident in this report should instill confidence in those who would use the data for outcome reporting and benchmarking, clinical research, and future dissemination to the public.

Data accuracy is particularly important within PC⁴ due to the relatively unique process of data flow from and back to the participants. Unlike many other registries in congenital cardiac care where cases are submitted in batches, PC⁴ data collection teams are encouraged to submit cases individually as soon as the hospitalisation ends and all data fields are complete. This process exists to support the real-time nature of outcome reporting within the collaborative. Cases that are submitted become incorporated into the online reporting

platform within 24 hours, and participants can view their aggregate results and perform local drill downs to identify individual patients in the registry with the most current information. This process and the online platform were created to give participants the tools for near real-time, ongoing assessment of performance in order to stimulate local quality improvement initiatives.

Maintaining such a process requires a robust set of data integrity checks built in to the data entry and reporting platforms. In collaboration with the software partners supporting these solutions, PC⁴ has created a system where clean records can be submitted without missing or illogical data, and these records are ready for use immediately by clinicians and researchers. Records with missing data do not enter the data warehouse. In addition to these electronic data checks, PC⁴ has invested a significant amount of time to promote and maintain consistent data collection across sites that conforms to the rigorously defined data definitions manual. These include dedicated in-person installations and training, regular contact with data collection teams through scheduled and ad hoc meetings, and maintenance of certification examinations for data collectors. The results of the initial audit confirm that these efforts have been successful.

In addition, in April 2015, PC⁴ participants began sharing data internally in an unblinded manner; clinical champions at each site are able to see aggregate centre-level data and identify the other participants. Sites that have submitted data for at least 1 year and passed the initial audit become eligible for this function using an opt-in opt-out approach; seven of the eight initially eligible sites opted in. This degree of transparency demands the highest standard of data quality, and various processes maintained by PC⁴, including the audit method presented here, have given participants confidence to use the data in this manner.

PC⁴ committed to auditing every hospital within the 1st year of data submission in order to ensure that every data collection team was given the opportunity to identify potential barriers to accurate and timely data submission as soon as possible. The audits were conducted in a completely non-punitive manner. Assessing and improving the accuracy and efficiency of data collection processes at the local level were the primary goals of the initial audit. The auditing team searches for evidence of systematic issues with data collection and case submission and provides education and support to the local teams to remedy any potential issues – for example, if an error in definition interpretation is uncovered, the audit team helps identify all cases – that is, even those cases that are not part of the audit sample – that may contain an error, and the participant data collection team edits and re-submits those cases as appropriate. At the conclusion of the audit analysis, the data coordinating centre personnel review the findings with the participants and cases are updated to reflect the most accurate data. Most data collection teams expressed great satisfaction with the process, were receptive to feedback on their performance, and embraced opportunities to improve their work. The data coordinating centre personnel learnt a great deal by observing data collection *in situ* at each hospital, and this knowledge continues to foster our ability to disseminate successful strategies for data collection across the collaborative and to new participants. The result of this shared experience is a path of continuous quality improvement in data collection methods.

The PC⁴ Executive Committee defined a passing audit score as a major discrepancy rate of <1.5%, >97% overall accuracy, and >90% of cases submitted within 1 month of hospital discharge. Participants achieving these standards receive and maintain privileges of data use within the collaborative, including access to unblinded data. Failure to meet standards of accuracy and timeliness trigger a remediation programme including re-education and close monitoring. The site would be re-audited within 1 year.

We found no significant differences in audit performance across sites, although the participating programmes varied in case volume and available resources for data collection. Based on the 1st year of experience, most hospitals dedicate approximately one full-time equivalent of data collector time per 800 CICU admissions; however, hospitals with fewer resources available for data collection have created systems to enter cases accurately and on time, as demonstrated in this report. Hospital case volume varied among the eight participants from 388 to 977 CICU admissions per year, and therefore a 62-case sample during the audit represents 6–16% of a hospital's yearly volume. As described, the sample size for audit was chosen based on the maximum number of cases that can be completed in a one and a half-day on-site visit. The data coordinating centre is considering the feasibility of adjusting sample sizes based on hospital case volume in future audits.

Going forward, PC⁴ will continue to perform audits at each new participant site within the 1st year of data collection initiation. All hospitals will be audited at least once every 2 years thereafter. There will likely be changes to the audit process in both content and methods, as the registry evolves and the number of participating sites grows. As new data variables are incorporated into the database, some or all of these will need to be verified in future audits – for example, in August 2014, participants began collecting an additional set of variables that will be used for case-mix adjustment across hospitals in order to evaluate differences in mortality rates and other clinical outcomes. The next template of key fields used in the audit must include those that are included in the final risk model to verify accurate collection of data necessary for complexity-adjusted outcome reporting. Conversely, some fields were identified in the initial audit where errors are exceedingly rare, and/or external auditors are unlikely to abstract the medical record more accurately than the local collection teams. These fields will be dropped in future audits, allowing the auditing team to focus on higher-impact aspects of the database.

Although the Data Coordinating Centre personnel have conducted all but the University of Michigan audit, this may be an unsustainable model over time. Future iterations of the audit process might involve other PC⁴ data collection teams in a peer-to-peer audit, a third-party organisation, or some combination of these options. Each approach presents numerous potential advantages and disadvantages. Although peer-to-peer auditing could be advantageous by allowing bidirectional exchange of knowledge, this approach could be perceived as less rigorous with a potential concern that the auditors may “let mistakes slide”; however, given the degree to which hospitals in the collaborative share data transparently with one another, and the commitment to using the data for quality improvement purposes among participants, it seems highly unlikely that peers will allow inaccurate data collection to occur. Although a third-party group may mitigate some of these potential concerns, there

can be limitations regarding content knowledge necessary to perform a detailed audit to the level currently required by PC⁴, and the cost could also be prohibitive.

Conclusion

Results from the 1st year audit of the PC⁴ registry demonstrate accurate, timely, and complete submission of data to a database designed to promote collaborative quality improvement for children with critical medical and surgical cardiovascular diseases. Regular source verification of data is the gold standard approach among congenital cardiac surgical databases such as the Society for Thoracic Surgeons Congenital Heart Surgery Database and the European Association for Cardio-Thoracic Surgery Congenital Heart Surgery Database, and it should be the expectation that critical-care databases for this population maintain the same standards. PC⁴ remains committed to the ongoing efforts to verify and report on the data integrity of the registry.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors recognise the work of all data collection team members at the participating Consortium sites whose efforts contributed to the data quality.

Financial Support

Dr Michael Gaies received support from the National Heart, Lung, and Blood Institute (K08HL116639, PI: Gaies) that indirectly supported this research.

References

1. Pasquali SK, He X, Jacobs JP, et al. Measuring hospital performance in congenital heart surgery: administrative versus clinical registry data. *Ann Thorac Surg*. 2015; 99:932–938. [PubMed: 25624057]
2. Jantzen DW, He X, Jacobs JP, et al. The impact of differential case ascertainment in clinical registry versus administrative data on assessment of resource utilization in pediatric heart surgery. *World J Pediatr Congenit Heart Surg*. 2014; 5:398–405. [PubMed: 24958042]
3. Pasquali SK, Jacobs JP. The role of databases in improving the quality of care for congenital heart disease. *World J Pediatr Congenit Heart Surg*. 2013; 4:139–141. [PubMed: 23799726]
4. 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; 25:951–957. [PubMed: 25167212]
5. 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]
6. Clarke DR, Breen LS, Jacobs ML, et al. Verification of data in congenital cardiac surgery. *Cardiol Young*. 2008; 18(Suppl 2):177–187. [PubMed: 19063789]
7. Paugh TA, Dickinson TA, Theurer PF, et al. Validation of a perfusion registry: methodological approach and initial findings. *J Extra Corpor Technol*. 2012; 44:104–115. [PubMed: 23198390]
8. Magee MJ, Wright CD, McDonald D, Fernandez FG, Kozower BD. External validation of the society of thoracic surgeons general thoracic surgery database. *Ann Thorac Surg*. 2013; 96:1734–1739. discussion 8–9. [PubMed: 23998406]

9. Maruszewski B, Lacour-Gayet F, Monro JL, Keogh BE, Tobota Z, Kansy A. An attempt at data verification in the EACTS congenital database. *Eur J Cardiothorac Surg.* 2005; 28:400–404. discussion 5–6. [PubMed: 16054826]
10. Jacobs ML. Editorial comment. *Eur J Cardio-Thorac Surg.* 2005; 28:405–406.

Table 1

Aggregate scores (n = 8 sites).

Individual sites									
	Overall	A	B	C	D	E	F	G	H
Overall accuracy (%)	99.13	99.38	98.91	98.90	99.12	99.40	98.88	99.04	99.52
Major discrepancy rate (%)	0.62	0.47	0.92	0.63	0.80	0.30	0.71	0.80	0.33

Table 2

Accuracy scores by domain (n = 8 sites, %).

	Individual sites							
	Overall	A	B	C	D	E	F	H
Patient info								
Overall accuracy	98.37	98.48	99.50	98.17	98.36	98.85	98.39	96.27 99.50
Major discrepancy rate	1.11	1.52	0.50	1.47	1.09	0.76	0.81	2.24 0.00
Hospitalisation/intensive care encounter								
Overall accuracy	98.22	98.08	97.18	97.05	97.96	99.33	98.90	97.98 99.21
Major discrepancy rate	0.92	0.73	2.76	0.54	2.17	0.00	0.55	1.15 0.25
Cardiothoracic surgery								
Overall accuracy	99.36	98.94	0.00	0.00	98.73	0.00	98.92	99.28 99.00
Major discrepancy rate	0.64	1.06	0.00	0.00	1.27	0.00	1.08	0.72 1.00
Therapies								
Overall accuracy	98.64	99.26	97.85	98.52	98.71	99.00	96.90	99.25 99.54
Major discrepancy rate	0.72	0.49	1.20	0.89	0.50	0.64	1.19	0.41 0.39
Complications								
Overall accuracy	99.67	99.95	99.76	99.69	99.74	99.56	99.48	99.60 99.64
Major discrepancy rate	0.33	0.08	0.11	0.40	0.24	0.24	0.55	0.58 0.33