Original Article

Author for correspondence:
Jeremy M. Archer, MD, MS, University of Florida Congenital Heart Center, P.O. Box 100297, Gainesville, FL 32610-0297, USA.
Tel: +1 352 273 7770; Fax: +1 352 273 5927.
E-mail: jmarcher@ufl.edu

CAMBRIDGE UNIVERSITY PRESS

Lessons learned from linking two complementary databases: the Society of Thoracic Surgeons Congenital Heart Surgery Database and The Vermont Oxford Network Expanded Database

Jeremy M. Archer[1,2] , Connie S. Nixon[1,2], Livia Sura[1], Dan Neal[3] and Jeffrey P. Jacobs[1,2,3]

[1]Department of Pediatrics, University of Florida College of Medicine, Gainesville, FL, USA; [2]Congenital Heart Center, University of Florida, Gainesville, FL, USA and [3]Department of Surgery, University of Florida College of Medicine, Gainesville, FL, USA

Abstract

The Society of Thoracic Surgeons Congenital Heart Surgery Database and the Vermont Oxford Network Expanded Database are both large, international, well-established quality and outcomes databases with high penetration in their respective fields of congenital heart surgery and neonatology. Previous studies have shown the value of combining large databases for research purposes. Our aim was to examine the feasibility and value of combining these databases on a local level.

We included patients from both databases, cared for at our centre and born from 2015–2020, who had cardiac surgery as neonates or during the birth hospitalisation. We examined the number of patients from each database and overlap between the two. We compared cardiac diagnoses, surgeries performed, pre-operative factors, mortality, and length of stay between databases.

Of the 255 patients meeting criteria, 209 (81.9%) had records in both databases. The most common diagnoses in both were hypoplastic left heart syndrome, coarctation, and transposition of the great arteries. Surgical data were incompletely recorded in Vermont Oxford. Gestational age, birth weight, multiple gestation, mortality, and length of stay did not differ significantly between the databases, while the percentage of patients with an extracardiac malformation or genetic syndrome recorded was higher in the Society for Thoracic Surgeons group.

Larger-scale matching and comparison studies using these databases are feasible and desirable; for some variables, a record with data from both databases may be more complete. Specific attention should be given to inclusion criteria, reconciling different schema of diagnoses, and formulating questions relying on each database's relative strengths.

The Society of Thoracic Surgeons Congenital Heart Surgery Database is a voluntary database founded in 1994 and now representing 113 of 125 paediatric cardiac surgical centres in the United States. Because most high-volume centres participate, nearly all congenital heart surgeries in the country are captured.[1] The database contains detailed patient demographics, pre-operative risk factors and comorbidities, diagnoses, procedure(s) performed, and surgical complications and outcomes, as well as extremely granular data about the surgical course and anaesthetic administration. Aggregate data are published twice yearly, and risk-adjusted centre-specific outcomes are publicly reported on a voluntary basis. This database has brought transparency to the forefront with attention to the wide variation in patient outcomes and has established robust risk-adjustment models.[2–6]

Similarly, the Vermont Oxford Network was established in 1990 for neonatal ICUs to collect and report standardised benchmarking and quality improvement data, for use at scales from the single centre to worldwide.[7,8] While the original database contains records for very low birth weight infants (<1500 gm) from more than 1400 centres worldwide, the Expanded Database eliminates the birth weight restriction and contains records for all infants admitted to the neonatal ICUs in a smaller, but still robust, subset of centres (526 as of 2020).[9] Participating centres report neonatal process and outcome measures including delivery room management, respiratory support, neonatal care, and common conditions of pre-maturity including bronchopulmonary dysplasia, retinopathy of pre-maturity, patent ductus arteriosus, and necrotising enterocolitis.[9] The Vermont Oxford database also contains coded diagnoses and procedures including CHD and cardiac surgical procedures.

Previous work has shown the value linking databases in general [10,11] and of using the Vermont Oxford database to study CHD in pre-mature infants in particular [12,13]. With the Expanded Database including all admitted infants, regardless of gestational age or birth weight, our aim was to explore the feasibility of linking the Society of Thoracic Surgeons Congenital Heart Surgery Database with the Vermont Oxford Network Expanded Database to study questions in the field of neonatal CHD that could be better answered with a linked database than with either database alone.

Materials and methods

Data sources, inclusion criteria, and database queries

We extracted the records of patients born between January 1, 2015, and December 31, 2020, and cared for at Shands Children's Hospital at University of Florida, from (a) the Society of Thoracic Surgeons Congenital Heart Surgery Database from Shands Children's Hospital at University of Florida or (b) the Vermont Oxford Network Expanded Database. From the Society of Thoracic Surgeons Congenital Heart Surgery Database, we included all patients born between 2015 and 2020 who underwent heart surgery during the initial (birth) hospitalisation or at 28 days of age or less regardless of when they were admitted to the hospital. From the Vermont Oxford Network Expanded Database, we included all infants born from the same period who had a cardiac surgery recorded in their record or who had a corresponding record in the Society for Thoracic Surgeons database.

Because the Society of Thoracic Surgeons database contains a record for each operation, rather than for each patient, we included a single record per patient, the one for the "index operation". An Index Operation is the first Cardiovascular Surgical Operation of a given episode of care. This approach preserved all patient- and hospitalisation-specific data, allowing direct comparison and matching of patient-level records between databases. To ensure data completeness, we performed a primary data query on both databases and then a secondary query on the Vermont Oxford database specifically for patients who were identified in the Society of Thoracic Surgeons database query but were not identified in the initial Vermont Oxford database query. The purpose of this secondary query was to identify infants in the Vermont Oxford database whose cardiac surgery was omitted from their record; a similar secondary query for the Society of Thoracic Surgeons data was not applicable because inclusion from that database was based on date of surgery relative to date of birth, not the presence of a cardiac surgery in the record.

Because of the exploratory nature of the study, all variables were extracted with the exception of the patient's social security number, and the mother's name and social security number. Identifying information such as patient name, date of birth, and medical record number was used for matching between datasets and then deleted.

Data verification

Each Vermont Oxford site creates and maintains a plan for patient identification and data security. Data are collected and entered at each site by a dedicated data collection team, assisted by a Vermont Oxford Network account manager, and supervised at the site level by personnel dedicated to clinical, data, reporting, and financial oversight. The data collection team members undergo training via operation manuals, video tutorials, and webinars. The Vermont Oxford Network provides electronic data collection and submission software to ensure data consistency and accuracy, performs error checking after data submission, and provides technical support to member centres. [14]

Similarly, each Society of Thoracic Surgeons Congenital Heart Surgery Database site undergoes a data verification process including intrinsic data verification of all submitted data; random site audits that include chart review, operating room case log comparisons, and a mortality review; and remediation of data outliers. The audit process is particularly robust with regard to outcome variables such as mortality. [15] Data entry at each site is overseen by a dedicated data manager.

Data analysis

To determine the overall feasibility of patient-level matching, we determined the number of patients with a record in each database, in both databases, and in one but not the other database.

We then evaluated and qualitatively compared the distribution of the primary cardiac diagnosis and the primary cardiac surgical procedure performed in the two patient groups. Both databases require users to input diagnoses and surgical procedures from a pre-determined list, although the Vermont Oxford database allows free-text entries as well. In the Society for Thoracic Surgeons-derived dataset, there were 48 unique primary diagnoses and 44 unique primary cardiac procedures reported. Because the Vermont Oxford diagnoses and surgeries were not prioritised as they were in the Society for Thoracic Surgeons data, and because some were entered as free text, one author (Dr. Archer) manually assigned each record a primary diagnosis and procedure. Each database lists diagnoses and procedures differently; for this study, we did not map these lists to one another in such a way as to allow a qualitative comparison. We also determined the distribution of diagnoses for patients with records in only one of the two databases.

To evaluate the concordance of pre-operative and patient-level factors, we compared following fields across the patient groups in each dataset: gestational age, birth weight, and the percentage of patients with multiple gestation, extracardiac malformation, and aneuploidy or genetic syndrome. We then determined and compared between databases the overall mean mortality at discharge and length of stay. Statistical analysis was with an unpaired t-test for numerical data and Fisher's exact test for percentages. In the case of length of stay, the statistical test was performed on the logarithm of the length of stay due to a non-normal distribution.

Finally, we performed a qualitative analysis of the patient population captured in each database, the variables in each database, as well as the discrete lists of cardiac diagnoses and surgical procedures in each database, to identify information unique to each database.

Institutional approval

The University of Florida Institutional Review Board approved this research protocol.

Results

There were 255 total records meeting inclusion criteria from either database, 239 (93.7%) from the Society of Thoracic Surgeons database and 212 (83.1%) from the Vermont Oxford database. A total of 196 (76.8%) had records in both databases, with 43 (16.8%) only in the Society for Thoracic Surgeons database and 16 (6.3%) only in the Vermont Oxford database.

Table 1. Primary diagnosis of patients in the STS database

Diagnosis	n	%
Hypoplastic left heart syndrome	63	26.4%
Coarctation of aorta	31	13.0%
TGA, IVS	16	6.7%
VSD + Coarctation of aorta	10	4.2%
TGA, VSD	10	4.2%
Patent ductus arteriosus	9	3.8%
Truncus arteriosus	6	2.5%
Pulmonary atresia, IVS + Coronary artery anomaly, Fistula	6	2.5%
Total anomalous pulmonary venous connection (TAPVC), type 1 (supracardiac)	5	2.1%
Pulmonary atresia, VSD (Including TOF, PA)	5	2.1%
AVC (AVSD), Complete	4	1.7%
DORV, TOF type	4	1.7%
DORV, TGA type	4	1.7%
Aortic arch hypoplasia	4	1.7%
Total anomalous pulmonary venous connection, Type 3 (infracardiac)	4	1.7%
Single ventricle, heterotaxia syndrome	4	1.7%
Single ventricle, mitral atresia	4	1.7%
TOF	3	1.3%
Pulmonary atresia	3	1.3%
Total anomalous pulmonary venous connection, type 2 (cardiac)	3	1.3%
Vascular ring	3	1.3%
VSD, type 2 (perimembranous) (paramembranous) (conoventricular)	3	1.3%
Ebstein's anomaly	2	0.8%
Single ventricle, DILV	2	0.8%
Arrhythmia, heart block	2	0.8%
Miscellaneous, other	2	0.8%
Cardiac tumour	2	0.8%
TOF, pulmonary stenosis	2	0.8%
Interrupted aortic arch + VSD	2	0.8%
Single ventricle, tricuspid atresia	2	0.8%
Single ventricle, unbalanced AV canal	2	0.8%
Pulmonary atresia, IVS	2	0.8%
Pulmonary artery stenosis, branch, central (within the hilar bifurcation)	1	0.4%
Truncus arteriosus + interrupted aortic arch	1	0.4%
Tricuspid regurgitation and tricuspid stenosis	1	0.4%
Arrhythmia, atrial, other	1	0.4%
Arrhythmia	1	0.4%
Aortic stenosis, valvar	1	0.4%
Coronary artery anomaly, other	1	0.4%

(Continued)

Table 1. (*Continued*)

Diagnosis	n	%
TOF, absent pulmonary valve	1	0.4%
Arrhythmia, heart block, congenital	1	0.4%
TOF, AVC (AVSD)	1	0.4%
Open sternum with open skin (includes membrane placed to close skin)	1	0.4%
Hypoplastic RV	1	0.4%
Pericardial effusion	1	0.4%
Interrupted aortic arch	1	0.4%
Pulmonary stenosis, Valvar	1	0.4%
Total	**239**	

Abbreviations: Atrioventricular (AV), Atrioventricular canal (AVC), Atrioventricular septal defect (AVSD), Double-inlet left ventricle (DILV), Double outlet right ventricle (DORV), Hypoplastic left heart syndrome (HLHS), Intact ventricular septum (IVS), Right ventricle (RV), Society of Thoracic Surgeons (STS), Totally anomalous pulmonary venous connection (TAPVC), Transposition of the great arteries (TGA), Tetralogy of Fallot (TOF), Ventricular septal defect (VSD).

Of the records meeting inclusion criteria in the Society for Thoracic Surgeons database, there were 325 records discarded as they contained non-"index" operations (205 (63.1%) delayed sternal closure, 50 (15.3%) mediastinal exploration, 31 (9.5%) mechanical circulatory support-related procedures, and the remaining 39 (12%) other secondary procedures). This left 239 unique patient records.

During manual assignment of free text fields, as described in the methods, some diagnoses or procedures were assigned to a category of multiple diagnoses (n = 8, 3.8%) or procedures (n = 1, <1%) when no clear primary single entry was appropriate. This manual assignment resulted in 26 unique primary diagnoses and 16 unique primary cardiac surgeries reported in our Vermont Oxford-derived dataset.

The distribution of primary cardiac diagnoses for patients in the Society of Thoracic Surgeons and the Vermont Oxford-derived datasets is shown in Tables 1 and 2, respectively. The most common diagnoses in each dataset were hypoplastic left heart syndrome (26.4 and 22.6%), variations of coarctation of the aorta, interrupted aortic arch, and aortic arch hypoplasia (20.0 and 17.5%), and transposition of the great arteries with concordant atrioventricular connections and discordant ventriculoarterial connections (16 and 12.7%).

Of the 44 unique primary cardiac surgical operations listed in the Society for Thoracic Surgeons-derived dataset, the top three were Norwood operation (21.3%), aortic arch repair (16.7%), and arterial switch operations (11.7%). As expected, this corresponded to the top three diagnoses. In the Vermont Oxford data, however, 95 (44.8%) were listed as "Repair of palliation of CHD," "other open heart or vascular surgery," or did not have any cardiac surgical procedure listed. Because of this high percentage of non-specific and missing data in the Vermont Oxford data, we did not attempt further analysis of the procedure distribution from this dataset.

Of the 43 patients with a record in the Society for Thoracic Surgeons database but not in the Vermont Oxford Network database, 37 (86.0%) were transferred from outside institutions to the paediatric cardiac ICU and thus bypassed the neonatal ICU and

Table 2. Primary diagnosis of patients in the VON database

Diagnosis	n	%
Hypoplastic left heart syndrome	48	22.6%
Transposition of the great vessels	27	12.7%
Interrupted aortic arch	19	9.0%
Coarctation of aorta	18	8.5%
Tetralogy of Fallot	17	8.0%
Diagnosis not listed	14	6.6%
Double outlet right ventricle	13	6.1%
Total anomalous pulmonary venous return	11	5.2%
Multiple complex diagnoses	8	3.8%
Complete atrioventricular septal defect	5	2.4%
Truncus arteriosus	5	2.4%
Pulmonary atresia with intact ventricular septum	5	2.4%
Tricuspid atresia	3	1.4%
Ventricular septal defect	3	1.4%
Atrial septal defect	2	0.9%
Hypoplastic right ventricle	2	0.9%
Vascular ring	2	0.9%
Pulmonary stenosis	2	0.9%
Partially anomalous pulmonary venous return	1	0.5%
Atrial tumour	1	0.5%
Flail leaflets of the Tricuspid valve	1	0.5%
Aortic stenosis	1	0.5%
Arrhythmia	1	0.5%
Ebstein's anomaly	1	0.5%
Rhabdomyoma	1	0.5%
Other diagnosis	1	0.5%
Total	**212**	

Abbreviation: Vermont Oxford Network (VON).

Table 3. Primary diagnoses of patients in the STS but not in the VON database

Diagnosis	Number	Per cent
Hypoplastic left heart syndrome	14	32.6%
Coarctation of aorta	12	27.9%
TGA, VSD	4	9.3%
Truncus arteriosus	2	4.7%
Arrhythmia, heart block	3	7.0%
Miscellaneous, other	1	2.3%
Hypoplastic RV	1	2.3%
Total anomalous pulmonary venous connection (TAPVC), type 1 (supracardiac)	1	2.3%
Single ventricle, DILV	1	2.3%
Vascular ring	1	2.3%
Coronary artery anomaly, other	1	2.3%
VSD, type 2 (perimembranous) (paramembranous) (conoventricular)	1	2.3%
Pulmonary atresia, IVS + coronary artery anomaly, Fistula	1	2.3%

Abbreviations: Atrioventricular (AV), Double-inlet left ventricle (DILV), Hypoplastic left heart syndrome (HLHS), Intact ventricular septum (IVS), Right ventricle (RV), Society for Thoracic Surgeons (STS), Totally anomalous pulmonary venous connection (TAPVC), Vermont Oxford Network (VON); Ventricular septal defect (VSD).

Table 4. Primary diagnosis of patients in the VON database but not in the STS database

Diagnosis	Number	Per cent
Tetralogy of Fallot	3	20.0%
Coarctation of aorta	2	13.3%
Pulmonary atresia with intact ventricular septum	2	13.3%
Double outlet right ventricle	2	13.3%
Double outlet right ventricle + atrioventricular septal defect	1	6.7%
Hypoplastic left heart syndrome	1	6.7%
Partially anomalous pulmonary venous connection	1	6.7%
Pulmonary stenosis	1	6.7%
Ventricular septal defect	1	6.7%
Atrioventricular septal defect	1	6.7%

Abbreviations: Society for Thoracic Surgeons (STS), Vermont Oxford Network (VON).

were not entered in the Vermont Oxford database; of the remainder, two were admitted to the neonatal ICU but erroneously not entered in the Vermont Oxford database, one was discharged from the newborn nursery and readmitted through the Emergency Department, and two had errors in the Medical Record Number and could not be matched. The distribution of diagnoses is shown in Table 3; of note, the majority had either a single ventricle lesion (n = 17, 39.5%) or coarctation of the aorta (n = 12, 27.9%).

Of the 16 patients with a record in the Vermont Oxford Network database but not in the Society for Thoracic Surgeons database, one was found to have a duplicate medical record number and actually did have a matching record in Society for Thoracic Surgeons. Of the remaining 15, most (n = 12, 80%) were ultimately found to have records in the Society of Thoracic Surgeons database that did not meet our original inclusion criteria because the surgery occurred outside of the neonatal period and the infant was admitted to our hospital after the day of birth. All but one of these were born in other hospitals and transferred into our neonatal ICU and thus entered into the Vermont Oxford database, explaining the

seeming discrepancy. The distribution of diagnoses in this group is relatively even [Table 4].

Three (20%) did not have a record in the Society for Thoracic Surgeons database. One of these underwent an interventional cardiac catheterisation with pulmonary balloon valvuloplasty, but not cardiac surgery. The discrepancy in this case was due to coding the catheter-based procedure as a surgery rather than a catheter-based procedure and resulting in the record erroneously meeting

Table 5. Data errors and lack of Specificity.

Error or non-specific item	STS	VON
Incorrect DOB	2	0
Incorrect MRN	4	2
Incorrect procedure code	0	1
Non-specific diagnosis	2	1
Non-specific procedure	0	95

Abbreviations: Society for Thoracic Surgeons (STS), Vermont Oxford Network (VON), date of birth (DOB), medical record number (MRN).

Table 6. Comparison of pre-operative risk factors.

Pre-operative factor	Units	STS	VON	p
Gestational age at birth[1]	weeks (σ)	37.3 (2.8)	37.4 (2.9)	0.739
Birth weight[2]	kg (σ)	2.9 (0.7)	2.9 (0.8)	0.506
Multiple birth	n (%)	19 (7.9%)	20 (9.4%)	0.617
Extracardiac lesion	n (%)	34 (14.2%)	15 (7.1%)	0.016*[3]
Genetic abnormality/ syndrome	n (%)	44 (18.4%)	18 (8.5%)	0.002*[3]

Key: *: statistically significant; σ: standard deviation.
[1]For the 196 matched pairs, the average gestational age discrepancy between VON and STS was 0.44 weeks (approximately 3 days), with 10 records having a difference >= 1 week and one record in which the GA was not entered in STS. For the remainder, the difference when present was due to the fact that STS records GA in weeks while VON records GA in weeks+days.
[2]For the 196 matched pairs, the average birth weight discrepancy between VON and STS was 0.015 kg. There was one record having a difference > 0.5 kg and one record where the birth weight was not entered in STS. For all but 8 of 196 matched pairs, the birth weight discrepancy was less than 10 grams and mostly due to rounding in the STS records.
[3]The significant differences in the per cent of patients with extracardiac lesions and genetic abnormality/syndrome are due primarily to data discrepancies between the two databases: There were 16 matched records that had an extracardiac lesion recorded in STS but not VON. There were 26 matched records with a genetic abnormality/syndrome recorded in STS but not VON and two with one recorded in VON but not STS. Note that these totals are slightly less than the actual differences observed because the full analysis also includes unmatched records, those present in one database without a corresponding record in the other.
Abbreviations: GA (gestational age), Society for Thoracic Surgeons (STS), Vermont Oxford Network (VON).

Table 7. Unique features of each database.

Information in VON but not in STS
Records of neonates who have heart disease but not undergo surgery
Details of delivery room resuscitation
Transcatheter interventions
Increased granularity of issues of pre-maturity
-NEC
-PDA, including medical and transcatheter therapy
-ROP
-BPD/CLD
Information in STS but not VON
Records of neonates not admitted through the NICU
Increased granularity of cardiac surgical procedures
Increased granularity of CHD diagnoses

Abbreviations: bronchopulmonary dysplasia (BPD), chronic lung disease (CLD), necrotising enterocolitis (NEC), neonatal ICU (NICU), patent ductus arteriosus (PDA), retinopathy of pre-maturity (ROP), Society for Thoracic Surgeons (STS), Vermont Oxford Network (VON).

inclusion criteria for this study. The other two underwent undergoing aortic arch repair and double outlet right ventricle repair respectively and did have records entered in the Society of Thoracic Surgeons database; however, in each case a misentered birth date resulted in the appearance that the surgery did not occur during the birth hospitalisation, and thus, the records erroneously did not meet inclusion criteria.

Table 5 summarises the data entry errors found, as well as the instances of a lack of specificity in coding diagnoses and procedures in both databases.

The comparison of pre-operative patient-specific factors is shown in Table 6. The mean gestational age, birth weight, and presence of multiple births did not differ significantly between datasets. There were however significantly more patients with extracardiac lesions and genetic abnormalities or syndromes recorded in the Society of Thoracic Surgeons dataset than in the Vermont Oxford dataset (14.2 versus 7.1%, p = 0.016; and 18.4 Versus 8.5%, p = 0.002, respectively). This difference was due primarily to patients for whom the presence of an extracardiac lesion or genetic abnormality was recorded in the Society of Thoracic Surgeons dataset and not the Vermont Oxford dataset.

The overall in-hospital mortality rate was 4.7% in the Vermont Oxford Network-derived dataset and 4.6% in the Society for Thoracic Surgeons-derived dataset (p = 1.0). The mean length of stay was 65.3 (σ = 64.6) days in the Vermont Oxford Network-derived dataset and 62.7 (σ = 65.3) days in the Society for Thoracic Surgeons-derived dataset (p = 0.227).

Table 7 lists the information unique to each database found by reviewing patient populations, variables, and diagnosis/procedure lists in each.

Discussion

We draw several important conclusions from this study. First, concordance of patient records between the two databases is high. For all but three of the patients with records in the Vermont Oxford Network Expanded Database-derived data, there was a corresponding record in the Society of Thoracic Surgeons Congenital Heart Surgery Database-derived data, and the three instances without a corresponding record were due to data entry errors. Although these data entry errors were manually corrected in our single-institution study, such errors might be impossible to identify and correct on a larger scale, multi-institution study. Nevertheless, our results suggest that these errors would not be highly problematic in that they represent 3 of 255 total patients (1.2%). Of 255 total patients, 209 (81.9%) had a record in both databases. The majority of the patients absent from the Vermont Oxford Database were admitted through the units in the hospital other than the neonatal ICU and would thus not be expected to have database records in the Vermont Oxford Database. Actual errors in database entry appeared to be rare, but were found.

Second, the distribution of diagnoses seemed to match well across these two datasets. Because the Society of Thoracic Surgeons Congenital Heart Surgery database contains a more specific list of diagnoses, matched records will have richer but still standardised (discrete) diagnosis versus Vermont Oxford data

alone. On the other hand, some analyses may benefit from the simplicity of fewer categories of diagnosis, and the Vermont Oxford classification may offer an advantage. A direct mapping of diagnoses from one scheme to the other would also allow for a robust, quantitative comparison, permitting further validation of accuracy in both cases. Surgical data in the Vermont Oxford-derived dataset were incomplete or non-specific in a high percentage of records; future larger studies should evaluate whether this reflects institutional practice or is inherent in the Vermont Oxford structure. With the high rate of non-specific procedural data in Vermont Oxford (44.8%), one might question the value of pursuing this linkage. However, the corresponding specific procedural information on these patients can be obtained from the Society for Thoracic Surgeons database, while Vermont Oxford Network will provide additional and complementary information as listed in Table 7 and described in the Results section.

Third, some pre-operative and patient-specific factors such as gestational age, birth weight, and the presence of a multiple gestation were similar between patient records derived from each individual database. Others, such as the presence of extracardiac malformations and genetic syndromes or anomalies, differed. Interestingly, this finding was not because of patients in one database but not the other, but rather because of patients in both databases for which there was a discrepancy in recording these factors. This fact suggests that combined and matched data could be more complete than either database alone. Larger studies will be required to discover if this is a national/international trend or a finding in our centre only.

Limitations of the study

As with all database studies, there are inherent limitations, including a reliance on externally controlled data reporting and integrity, and the potential weaknesses of data encoding and classification schemes. Fortunately, both databases are well-established and have rigorous data collection protocols with an average of 30 years of refinement. Because neither database has restrictive criteria as to which patients are entered (all patients admitted to the neonatal ICU are entered into the Vermont Oxford Network Expanded Database, and all patients undergoing cardiac surgery are entered into the case of the Society of Thoracic Surgeons Congenital Heart Surgery Database), selection bias is minimised. A degree of selection bias does exist based on which centres choose to participate; however, both databases have a large number of participating centres, and the Society for Thoracic Surgeons has a very high penetrance. As we have demonstrated, the inclusion criteria of future studies will need to be carefully refined to ensure that all records that should be included based upon the intent of the study are captured according to the prospectively written study protocol. Finally, the issue of reconciling the two different schemas of recording diagnoses and procedures needs to be carefully addressed in future studies. Varying degrees of specificity may be desirable in certain cases.

Conclusion

In summary, this single-institution study demonstrates the feasibility of both matching and comparing data from the Society of Thoracic Surgeons Congenital Heart Surgery Database and the Vermont Oxford Network Expanded Database. A high degree of concordance of matched records is present, and coded diagnoses as well as pre-operative risk factors are recorded similarly in both databases. Future studies might include analysis of matched data

from all United States of America centres that participate in both databases or comparison of non-matched aggregate data from all participating centres. The data in the Vermont Oxford Network Expanded Database can provide valuable information about neonates with cardiac disease who do not undergo cardiac surgery, serving as a larger population-level denominator and allowing analyses based on a diagnostic rather than procedural cohort.

Similarly, the Society of Thoracic Surgeons Congenital Heart Surgery Database provides information about neonates with heart disease who are not admitted via the neonatal ICU. Finally, the combination of more granularity regarding cardiac diagnoses and procedures in the Society of Thoracic Surgeons data and more specific information about neonatal comorbidities and neonatal intensive care in the Vermont Oxford data can answer questions that may not be answerable using either database alone.

References

1. Kumar SR, Mayer JE, Jr., Overman DM, Shashidharan S, Wellnitz C and Jacobs JP. The society of thoracic surgeons congenital heart surgery database: 2021 update on outcomes and research. Ann Thorac Surg 2021; 112: 1753–1762.
2. Pasquali SK, Thibault D, O'Brien SM et al. National variation in congenital heart surgery outcomes. Circulation 2020; 142: 1351–1360.
3. Asante-Korang A and Jacobs JP. Big data and paediatric cardiovascular disease in the era of transparency in healthcare. Cardiol Young 2016; 26: 1597–1602.
4. Jacobs JP, Nelson JS, Fuller S, Scholl FG, Kumar SR and Jacobs ML. Risk adjustment for cardiac surgery in adults with congenital heart disease: what do we know and what do we need to learn? In: Eur J Cardiothorac Surg 2021; 60: 1405–1407.
5. O'Brien SM, Jacobs JP, Pasquali SK et al. The society of thoracic surgeons congenital heart surgery database mortality risk model: part 1-statistical methodology. Ann Thorac Surg 2015; 100: 1054–1062.
6. Jacobs JP, O'Brien SM, Pasquali SK et al. The society of thoracic surgeons congenital heart surgery database mortality risk model: part 2-clinical application. Ann Thorac Surg 2015; 100: 1063–1068; discussion 1068–1070.
7. Edwards EM, Ehret DEY, Soll RF and Horbar JD. Vermont Oxford Network: a worldwide learning community. Transl Pediatr 2019; 8: 182–192.
8. Horbar JD. The Vermont Oxford Network: evidence-based quality improvement for neonatology. Pediatrics 1999; 103: 350–359.
9. https://public.vtoxford.org/member-map/,https://public.vtoxford.org/data-and-reports/database-summaries/,andhttps://public.vtoxford.org/data-and-reports/.Accessed July 6, 2022.
10. Jacobs JP, Overman DM and Tweddell JS. The value of longitudinal follow-up and linked registries. J Am Coll Cardiol 2021; 78: 1714–1716.
11. Jacobs JP, Shahian DM, Grau-Sepulveda M et al. Current penetration, completeness, and representativeness of the society of thoracic surgeons adult cardiac surgery database. Ann Thorac Surg 2022; 113: 1461–1468.

12. Archer JM, Yeager SB, Kenny MJ, Soll RF and Horbar JD. Distribution of and mortality from serious congenital heart disease in very low birth weight infants. Pediatrics 2011; 127: 293–299.

13. Fisher JG, Bairdain S, Sparks EA et al. Serious congenital heart disease and necrotizing enterocolitis in very low birth weight neonates. J Am Coll Surg 2015; 220: 1018–1026.e1014.

14. Vermont Oxford Network Manual of Operations Part 1 v4.1 https://vtoxford.zendesk.com/hc/article_attachments/6539565329299/Manual_of_Operations_Part_1_v4.1.pdf. Accessed July 6, 2022 pp. 6–12, 29–37.

15. Overman DM, Jacobs ML, O'Brien JE, Jr. et al. Ten years of data verification: the society of thoracic surgeons congenital heart surgery database audits. World J Pediatr Congenit Heart Surg 2019; 10: 454–463.