Development of a diagnosis- and procedure-based risk model for 30-day outcome after pediatric cardiac surgery

Sonya Crowe, PhD,[a] Kate L. Brown, MPH, MRCP,[b] Christina Pagel, PhD,[a] Nagarajan Muthialu, FRCS,[b] David Cunningham, PhD,[c] John Gibbs, FRCP,[c] Catherine Bull, MRCP,[b] Rodney Franklin, MD, FRCP,[d] Martin Utley, PhD,[a] and Victor T. Tsang, MD, FRCS[b]

Objective: The study objective was to develop a risk model incorporating diagnostic information to adjust for case-mix severity during routine monitoring of outcomes for pediatric cardiac surgery.

Methods: Data from the Central Cardiac Audit Database for all pediatric cardiac surgery procedures performed in the United Kingdom between 2000 and 2010 were included: 70% for model development and 30% for validation. Units of analysis were 30-day episodes after the first surgical procedure. We used logistic regression for 30-day mortality. Risk factors considered included procedural information based on Central Cardiac Audit Database ''specific procedures,'' diagnostic information defined by 24 ''primary'' cardiac diagnoses and ''univentricular'' status, and other patient characteristics.

Results: Of the 27,140 30-day episodes in the development set, 25,613 were survivals, 834 were deaths, and 693 were of unknown status (mortality, 3.2%). The risk model includes procedure, cardiac diagnosis, univentricular status, age band (neonate, infant, child), continuous age, continuous weight, presence of non–Down syndrome comorbidity, bypass, and year of operation 2007 or later (because of decreasing mortality). A risk score was calculated for 95% of cases in the validation set (weight missing in 5%). The model discriminated well; the C-index for validation set was 0.77 (0.81 for post-2007 data). Removal of all but procedural information gave a reduced C-index of 0.72. The model performed well across the spectrum of predicted risk, but there was evidence of underestimation of mortality risk in neonates undergoing operation from 2007.

Conclusions: The risk model performs well. Diagnostic information added useful discriminatory power. A future application is risk adjustment during routine monitoring of outcomes in the United Kingdom to assist quality assurance. (J Thorac Cardiovasc Surg 2013;145:1270-8)

CHD

Since one UK center experienced a number of ''excess deaths'' in children after cardiac surgery,[1] a culture of audit and quality improvement has emerged in the United Kingdom, with particular interest in monitoring outcomes and center performance within pediatric cardiac surgery.[2,3] A major review of pediatric cardiac surgery services in the United Kingdom[4] recently stressed the need for national processes for reporting outcomes to be timely and meaningful. Yet to do such routine monitoring fairly and effectively, one needs to account for the case mix of each center.[5] Adjusting for risk in pediatric cardiac surgery is challenging because of the diversity of the patient population in terms of the diagnoses, operations performed, age at operation, and other factors.[6]

A worldwide effort to collect data for quality assurance and benchmarking[7-9] has seen the evolution of a number of multi-institutional databases. This activity has been underpinned by ongoing work on congenital cardiac diagnostic and procedural coding toward the development of universally applicable codes to describe the pediatric cardiac case mix.[10-12] Accrual of standardized data on case mix and outcomes has led to a shift from the use of consensus-based risk stratification tools (eg, RACHS-1 [Risk Adjustment for Congenital Heart Surgery-1] categories[13] and Aristotle Basic Complexity Levels [ABC Levels][14]) to risk estimates based on empirical data.[15] Of note, this previous work has focused on outcomes according to the procedure performed, without account taken of the range of cardiac diagnoses for which some procedures are performed.

The current article reports the development of the Partial Risk Adjustment in Surgery (PRAiS) model for pediatric cardiac surgery, which is based on empirical data, with

From the Clinical Operational Research Unit,[a] University College London, London, United Kingdom; Cardiac Unit,[b] Great Ormond Street Hospital, London, United Kingdom; Central Cardiac Audit Database,[c] National Institute for Cardiovascular Outcomes Research, University College London, London, United Kingdom; and Royal Brompton and Harefield Hospitals,[d] London, United Kingdom.

Funding: the National Institute for Health Research Health Services Research program (Project Number 09/2001/03). See the Health Services Research Web site (http://www.netscc.ac.uk/hsdr/projdetails.php?ref=09-2001-13) for more information. The views and opinions expressed are those of the authors and do not necessarily reflect those of the Health Services Research program, National Institute for Health Research, National Health Service, or the Department of Health.

Disclosures: Authors have nothing to disclose with regard to commercial support.

Received for publication Feb 14, 2012; revisions received April 30, 2012; accepted for publication June 12, 2012; available ahead of print July 23, 2012.

Address for reprints: Sonya Crowe, PhD, Clinical Operational Research Unit, University College London, 4 Taviton Street, London WC1H 0BT, United Kingdom (E-mail: sonya.crowe@ucl.ac.uk).

0022-5223/$36.00

http://dx.doi.org/10.1016/j.jtcvs.2012.06.023

Abbreviations and Acronyms
CCAD = Central Cardiac Audit Database
EACTS = European Association for Cardio-Thoracic Surgery
IPCCC = International Paediatric and Congenital Cardiac Code
STS = Society of Thoracic Surgeons

procedural information augmented by information on cardiac diagnosis in addition to age, weight, and comorbidities. The motivation was to develop a model fit for the purpose of adjusting for case-mix severity during routine monitoring of short-term outcomes after pediatric cardiac surgery in the United Kingdom.

MATERIALS AND METHODS

Data Source and Study Population

The pseudonymized dataset used in this study was provided by the Central Cardiac Audit Database (CCAD).[16] Since 2000, mandatory data submissions to CCAD have been requested every 3 months from all hospitals performing cardiac surgery in the United Kingdom, including details about patient diagnoses and comorbidities, and the operation performed. The data are validated and subject to a quality assurance process, with all units undergoing annual inspection in which local records are examined to ensure every case performed in the center has been submitted and a random sample of case notes is examined in detail to assess data quality.[9] Patients' survival status is independently verified through periodic requests to the National Health Service Central Register, as approved by the National Information and Governance Board for Health and Social Care, with consent requested from patients/parents for participation in national audit of outcomes.

The data used concerned surgical operations conducted before October 31, 2010, in patients aged less than 16 years. Official transition to adult services in the United Kingdom occurs at 16 years of age, and guidelines recommend the treatment of individuals aged 16 years or more to be in an adult center. The dataset was then split into development (70% of patients) and validation (30% of patients) samples using random allocation stratified by year and institution of first procedure. The development sample contained 34,385 records, corresponding to 22,449 unique patients. The validation sample containing 14,316 records (9354 unique patients) was set aside and not used in risk model development.

Defining Episodes of Surgical Management

To obviate ambiguities in assigning short-term outcomes to operations performed close together in time, we defined 30-day episodes of surgical management. The first such episode for a patient started with his/her first surgical operation and was assigned an outcome of alive or dead according to the vital status of the patient at 30 days. Any reintervention within this 30-day episode was not included in model development but was noted as a secondary outcome of the episode for the purposes of monitoring (not reported in this article). The patient's next surgical operation *more than 30 days after* the start of this first episode was treated as the start of a new episode and so forth. Each episode was treated as independent within the analysis.

Grouping Operations Using the Central Cardiac Audit Database "Specific Procedure" Algorithm

A combination of up to 8 individual procedural International Paediatric and Congenital Cardiac Codes (IPCCCs)[17] may be submitted to CCAD to describe each operation. The Steering Committee of CCAD, which includes experienced pediatric cardiac surgeons and cardiologists, have developed a specific procedure algorithm that links the combinations of individual IPCCCs in a record to at most 1 of 36 recognizable operations. The list of 36 operations (hereafter referred to as "specific procedures") includes generally accepted benchmark operations[18] along with others that were determined by the CCAD Steering Committee between 2000 and 2010. The algorithm imposes a hierarchy with the record assigned the most complex specific procedure consistent with the combination of codes submitted. The 36 specific procedures capture 83% of operations in the data and center-specific outcomes for these specific procedures have been published by CCAD on the Internet[16] and are well known as a core output of CCAD.

Classification of Primary Diagnosis

Each CCAD record contains up to 6 IPCCC diagnostic codes. To explore the potential for this information to add discriminatory power to risk adjustment, we developed a new hierarchical scheme that links the combination of IPCCC diagnostic codes available for a record to at most 1 of 24 primary cardiac diagnoses. We also identified those combinations of IPCCCs that indicated that the patient had a functionally univentricular heart. The process for developing these diagnostic categories is described in detail by Brown and colleagues.[19]

Other Factors Considered

Given the planned use of the model in quality assurance, only preoperative factors were considered for inclusion in the risk model. In addition to specific procedure and diagnostic information, the factors considered on the basis of potential clinical relevance and availability within the dataset were year of surgery; whether the procedure was performed on bypass; patient sex, age, weight; whether there was an antenatal diagnosis; ethnicity; the Townsend score of socioeconomic deprivationc[20]; and comorbidity.

IPCCCs defining comorbid conditions were grouped into 4 categories: premature (gestational age <37 weeks); Down syndrome; congenital non–Down syndrome comorbidity (all genetic syndromes, clinical constellations of features that constitute a recognized syndrome, and congenital structural defects of organs other than the heart[21]); and acquired comorbidity (including preoperative comorbidities acquired as a result of heart disease or its treatments, eg, renal failure or necrotizing enterocolitis).[22] For a given patient record, comorbid conditions appearing as IPCCCs in any of the comorbidity or diagnosis fields were classed as comorbidities. We treated records where no comorbidities were entered as though that patient did not have any comorbidity.

Missing and Unknown Data

Episodes with missing 30-day outcome were removed. Weight-for-age z scores were calculated for each episode on the basis of a subdivision of the development dataset into 23 age bands (narrower at younger ages). Episodes in the development set with an absolute z score of 3 or more were considered infeasible and, along with episodes with missing weights, assigned the mean weight of their corresponding age band. To mimic prospective use, no adjustment of weights of this nature was made in the validation set. Where inconsistencies in any of the data were suspected, for example, between episodes relating to the same patient, the data were confirmed with CCAD.

Model Development

After descriptive analyses that were performed to characterize the development dataset, univariate 30-day, episode-level mortality rates were calculated for the candidate preoperative risk factors, with some removed from consideration on the basis of this univariate analysis. Some risk factors were removed because of considerations of data completeness.

Multiple logistic regression analysis was conducted within PASW Statistics 18, Release Version 18.0.0 (SPSS, Inc, 2009, Chicago, Ill), using

backward stepwise and ''enter'' regression methods to identify potential models and to parameterize prespecified models, respectively.[23] The area under the receiver operating characteristic curve (C-index), the Hosmer–Lemeshow chi-square statistic, and MADCAP charts[24] were used to assess the discrimination and accuracy of the candidate models developed. The MADCAP charts show cumulative predicted and observed deaths versus episode number, with episodes ordered by increasing predicted risk, enabling visual identification of patterns of systematic over- or underestimation of risk. Comparison of MADCAP charts was used to gauge value added or lost by adopting different approaches to analyzing variables and in assessing stability of model parameterization when using different random subsets of the development dataset.

Instability of model parameterization across random subsets of the development data was taken to indicate a risk of overfitting the data. In these instances, we simplified variables by reducing the number of categories and assessed tradeoffs between model performance and stability of parameterization.

Model development followed an iterative process of multiple logistic regression: assessment of model performance and stability, discussion between clinicians and analysts, and variable simplification. Ultimately, model choice was influenced by considerations of uptake by CCAD and UK centers, as well as statistical performance.

The discrimination and accuracy of the final model were assessed in the independent validation dataset. For interest, 2 additional models were evaluated—one based solely on specific procedure and the other comprising all factors in the final model except specific procedure.

The distribution of predicted risk in the development and validation sets was also compared to assess stability of case mix.

RESULTS

Development and Validation Sets

A total of 693 episodes with missing 30-day status (90% of which occurred before 2002) and a further 72 episodes with missing patient age were removed from the development set. A total of 1485 episodes with missing or anomalous patient weight were assigned the mean weight for the appropriate age. The final development set comprised 26,447 episodes corresponding to 21,610 unique patients. Of these, 834 episodes (3.2%; confidence interval, 3.0-3.4) had a 30-day outcome of death. A total of 1181 episodes contained at least 1 surgical reintervention within 30 days, and 466 episodes contained at least 1 catheter reintervention.

Preliminary Analysis

During preliminary model development, the following risk factors were eliminated from further consideration because of levels of missing data: antenatal diagnosis (38% missing), ethnicity (26%), and the Townsend deprivation score[20] (26%). Patient sex showed no univariate association with 30-day mortality and was not considered further in model development.

The association between age and mortality is nonlinear, and, having explored several options, we chose to include in the model both continuous age and 3 age bands: neonate (<30 days), infant (30 days to 1 year), and child (>1 year).

Because of instability across random subsets of the data (see ''Model Development'' section), we grouped the 9 specific procedures with the lowest volumes (all with <70 episodes in the development set) into a ''low-volume'' specific procedure group comprising aortic root replacement (not Ross), aortopulmonary window repair, atrioventricular septal defect and tetralogy repair, cor triatriatum repair, multiple ventricular septal defect closure, Senning or Mustard procedure, Tetralogy with absent pulmonary valve repair, tricuspid valve replacement, and truncus and interruption repair. We also grouped those diagnostic categories with similar mortality rates into low-, medium-, and high-risk groups (Appendix 1 shows details on the mappings) and grouped the non–Down syndrome comorbidities to give a variable indicating the presence of comorbidity other than Down syndrome (Appendix 2). Down syndrome was not associated with increased risk.

Applying an initial model that did not include year of surgery showed a clear trend of improvement in risk-adjusted outcomes over time. We added a binary variable to indicate whether an episode occurred pre-2007 or from 2007 onward. Although there is no clinical mechanism for such a threshold effect, it enabled the entire development set to be used in the model parameterization while also increasing the likelihood that the model is fit for prospective use.

Descriptive Analysis

The observed 30-day mortality rates in the development and validation sets for those parameters included in the model are shown in Table 1. Thirty-day mortality for specific procedures is shown in Figure 1 (development and validation sets). We note that episodes with missing or anomalous age or weight, or a 30-day status of ''unknown'' were not included in the calculation of univariate mortality rates or within model development or evaluation.

To illustrate the value that diagnostic information can bring to risk adjustment, consider the arterial shunt, one procedure performed in patients with differing anatomic substrates. In the development set, the 30-day mortality for episodes with arterial shunt was 7.1% overall; 12.3% for those with a univentricular heart versus 5.8% for those with a biventricular heart; and 4.1%, 7.2%, and 11.4% in the low-, medium-, and high-risk diagnostic categories, respectively.

Final Risk Model

The final risk factors included in the logistic regression model were age (both as a continuous measure and as neonate/infant/child bands), weight, specific procedure (including a ''low-volume'' group), procedure type (bypass or nonbypass), diagnosis group (low, medium, or high risk), univentricular heart attribute, presence/absence of a recorded non–Down syndrome comorbidity, and episode pre- or post-2007. Details of the regression model are shown in Appendix 3.

TABLE 1. Proportional breakdown of episodes within categories for risk factors included in the final model, along with their associated 30-day mortality

	Development set (used for building model)		Validation set (used for model evaluation)	
Model parameter	Proportion of episodes (95% CI)	30-d mortality (95% CI)	Proportion of episodes (95% CI)	30-d mortality (95% CI)
Age continuous	–	Not shown	–	Not shown
Age band				
Neonates	21.1% (20.0-22.2)	6.8% (6.2-7.5)	21.0% (19.4-22.8)	7.8% (6.8-9.0)
Infants	37.2% (36.3-38.2)	3.1% (2.8-3.5)	36.7% (35.2-38.2)	2.5% (2.1-3.0)
Children	41.7% (40.8-42.7)	1.3% (1.1-1.6)	42.3% (40.9-43.8)	1.4% (1.1-1.8)
Weight (continuous)	–	Not shown	–	Not shown
Specific procedure	–	See Figure 1	–	See Figure 1
Procedural type				
Bypass	74.2% (73.6-74.9)	3.0% (2.8-3.3)	74.8% (73.9-75.8)	3.0% (2.6-3.4)
Nonbypass	25.8% (24.7-26.8)	3.5% (3.1-3.9)	25.2% (23.6-26.9)	3.7% (3.0-4.5)
Diagnostic grouping				
Low-risk diagnosis	39.6% (38.7-40.5)	1.2% (1.0-1.5)	39.9% (38.5-41.4)	1.2% (0.9-1.6)
Medium-risk diagnosis	51.4% (50.6-52.3)	3.7% (3.4-4.1)	50.6% (49.3-52.0)	3.8% (3.3-4.4)
High-risk diagnosis	9.0% (7.9-10.2)	9.0% (7.9-10.2)	9.4% (7.8-11.5)	7.9% (6.4-9.8)
Ventricular status				
Not univentricular heart	85.2% (84.7-85.7)	2.6% (2.4-2.8)	85.0% (84.3-85.8)	2.5% (2.2-2.9)
Univentricular heart	14.8% (13.8-16.0)	6.8% (6.1-7.7)	15.0% (13.4-16.9)	6.8% (5.7-8.2)
Comorbidities				
No comorbidities*	89.0% (88.6-89.4)	2.9% (2.7-3.2)	88.6% (88.0-89.3)	2.9% (2.6-3.3)
At least 1 comorbidity*	11.0% (10.0-12.2)	5.5% (4.7-6.4)	11.4% (9.8-13.4)	4.9% (3.8-6.3)
2007 indicator				
Pre-2007	64.4% (63.7-65.2)	3.4% (3.2-3.7)	63.1% (62.0-64.3)	3.1% (2.7-3.5)
2007 onward	35.6% (34.6-36.6)	2.9% (2.5-3.2)	36.9% (35.4-38.4)	3.3% (2.8-3.9)
Overall	100%	3.2% (3.0-3.4)	100%	3.2% (2.8-3.5)

Figures are shown for the development and validation sets after excluding episodes with missing 30-day status. *CI*, Confidence interval. *Does not include Down syndrome.

This model was parameterized across the entire development set, giving a C-index of 0.78 (indicating reasonable discrimination) and Hosmer–Lemeshow chi-square of 9.2 ($P = .325$) (indicating no statistically significant differences between observed and expected number of deaths when calculated in deciles of predicted risk).

Evaluation of the Risk Model

A risk score could be calculated for 95% of episodes in the validation set: Age was missing in 0.2% of episodes, and weight was missing in 4.7% of episodes. Thirty-day outcome was missing in 226 episodes. Among the remaining 10,597 episodes (in 7849 patients), there were 335 deaths within 30 days (3.2%; confidence interval, 2.8-3.5). A total of 468 episodes included at least 1 surgical reintervention, and 181 episodes included at least 1 catheter reintervention.

Figure 2 is a MADCAP chart showing the performance of the model across the spectrum of predicted risk in the validation set. The C-index is 0.77 compared with 0.78 in the development set. This good discrimination can be seen in the MADCAP chart with a shallow climb of cumulative observed deaths (stepped line) at low predicted risk and a steeper climb at high predicted risk. The Hosmer–Lemeshow chi-square statistic is 22.7, indicating that the discrepancies between observed and predicted mortality in deciles of predicted risk are statistically significant ($P = .004$), which was predominantly due to the higher than predicted number of deaths in the first and fourth deciles of predicted risk.

These discrepancies are evident in the portions of the MADCAP chart where the stepped line climbs at a higher or lower rate than the smooth line (predicted deaths). The overall number of predicted deaths was 329.3 compared with the 335 observed.

Figure 3 shows a receiver operating characteristic curve that illustrates the additional discriminatory power of diagnostic and other patient information: A model based solely on specific procedure gave a C-index of 0.72, and a model based on all factors in the final model *except* specific procedure has a C-index of 0.74 in the validation set.

The risk model is intended for future use in routine monitoring. Given this and the observed improvement in outcomes over time and our adjustment for this in the model, it is performance of the model in episodes that occurred during or after 2007 that is most informative concerning its fitness for purpose. Figure 4 shows a MADCAP chart of model performance in episodes occurring after January 1,

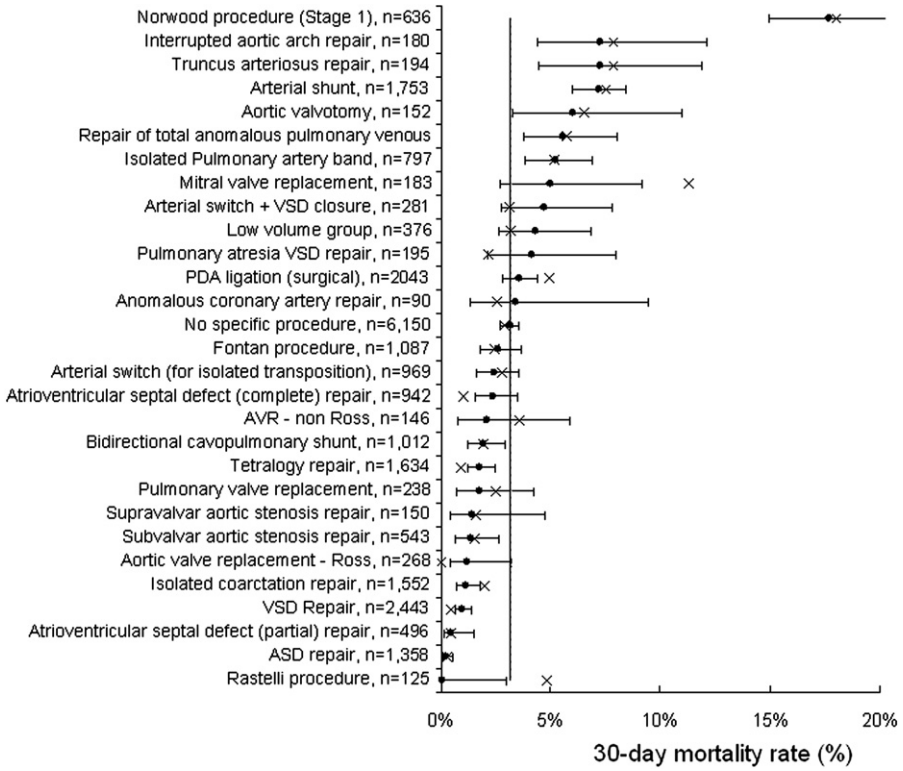

FIGURE 1. Observed 30-day mortality for specific procedures in the development set (*circles*) and validation set (*crosses*). The numbers (*n*) of episodes for each specific procedure in the development set are listed. The low-volume specific procedure group includes aortic root replacement (not Ross), aortopulmonary window repair, atrioventricular septal defect and tetralogy repair, cor triatriatum repair, multiple ventricular septal defect closure, Senning or Mustard procedure, tetralogy with absent pulmonary valve repair, tricuspid valve replacement, and truncus and interruption repair. The *vertical lines* denote the mean 30-day mortality in the development set (*black*) and validation set (*grey dashed*): Note that these are almost identical. Data are ordered in decreasing 30-day mortality for the development set. Note that the Rastelli procedure is defined as an intraventricular left ventricle to aorta tunnel and right ventricle to pulmonary artery conduit. *VSD*, Ventricular septal defect; *PDA*, patent ductus arteriosus; *AVR*, aortic valve replacement; *ASD*, atrial septal defect.

2007, in the validation set. The corresponding C-index is 0.81 compared with 0.77 across all years.

Although the model shows better discrimination among data from 2007 onward, it underestimates risk at the very high risk end of the spectrum of predicted risk (Figure 4, right). It is, as a result, less accurate overall in these more recent data than in the full development set.

The distributions of predicted risk in the development and validation sets were found to be similar: Approximately 30% of episodes have 1% or less predicted risk of 30-day mortality, 80% of episodes have 4% or less predicted risk, and 5% of episodes have a predicted risk of more than 10%.

DISCUSSION

We have developed a risk model for use in monitoring 30-day mortality in pediatric cardiac surgery that incorporated diagnostic information in addition to procedure, age, weight, and comorbidity. The model shows reasonable accuracy and good discrimination between groups of patients with high and low mortality, with a C-index of 0.77 when evaluated across the entire validation data and a C-index of 0.81 for post-2007 data. This discrimination is similar to that of other published risk-adjustment tools from the same field of practice.[15] As we have shown, supplementing procedural information with diagnostic, age, weight, and comorbidity characteristics increased the discriminatory performance of the risk model: The C-index across the entire validation data was 0.72 when only specific procedure was included.

It has been observed that procedure categories developed for use in risk adjustment may have incomplete coverage, leaving some operations excluded from outcome analyses.[6,13,15] The empirically based tool for analyzing mortality, the Society of Thoracic Surgeons European Association for Cardio-Thoracic Surgery (STS-EACTS) Congenital Heart Surgery Mortality Score and the STS-EACTS Congenital Heart Surgery Mortality Categories, published by the STS and EACTS in 2009, increased the coverage of records by including 148 types of operation and using a Bayesian model to adjust for small denominators.[15] As discussed in the ''Materials and Methods'' section, the specific procedure categories reported by CCAD online[16] are well established and accepted within the United

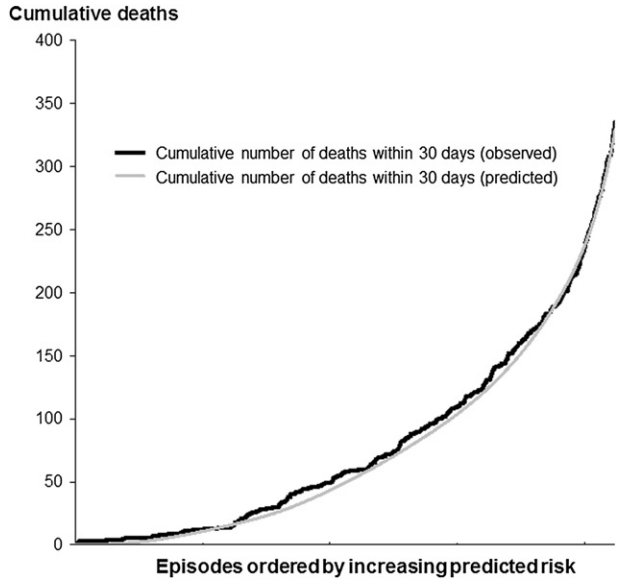

FIGURE 2. Cumulative deaths among the entire validation set plotted against episode number with episodes ordered by increasing risk as predicted by the risk model.

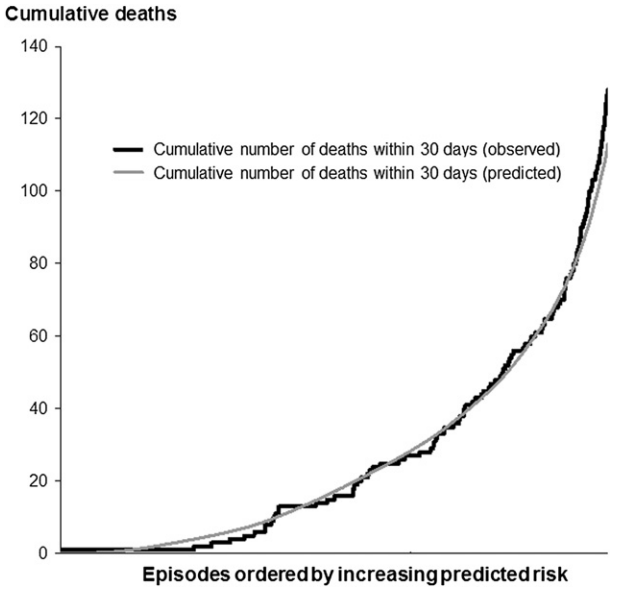

FIGURE 4. Performance of the model in the validation set for all episodes occurring after January 1, 2007.

Kingdom for benchmarking: These cover 83% of records in the dataset. No additional subjective procedural complexity ratings were used. We gathered additional information for use in mortality predictions by ascertaining cardiac diagnosis, which could be allocated to 97.1% of records classed as ''not a specific procedure,'' the most common diagnosis (11.6%) being ''acquired.''

The risk model is intended for future use in routine monitoring of risk-adjusted outcomes within UK pediatric cardiac centers. It is the performance of the model during or after 2007 that is most informative concerning its fitness for this purpose. In this period, the model was found to underestimate risk at the very high-risk end. This indicates that risk adjustment based on the current parameterization of the model will potentially give an unfair assessment of outcomes at those centers with a high proportion of high-risk cases. This is an important caveat to interpretation of risk-adjusted outcomes within and between centers that will need to be considered as the work is taken forward. It is important to understand how differences in case mix and differential performance of the risk model in different subgroups could combine to give an artefactual impression of better or worse risk-adjusted outcomes at one center compared with another. This issue is of particular importance, given the level of scrutiny to which these types of outcome data are exposed. Although the United Kingdom is currently the only country that displays unit-specific pediatric cardiac surgical outcomes of procedures online,[16] there has been considerable debate of this issue in the professional journals, with the suggestion that program-level reporting of unit-specific outcomes across a range of domains may evolve internationally over the coming years.[25,26]

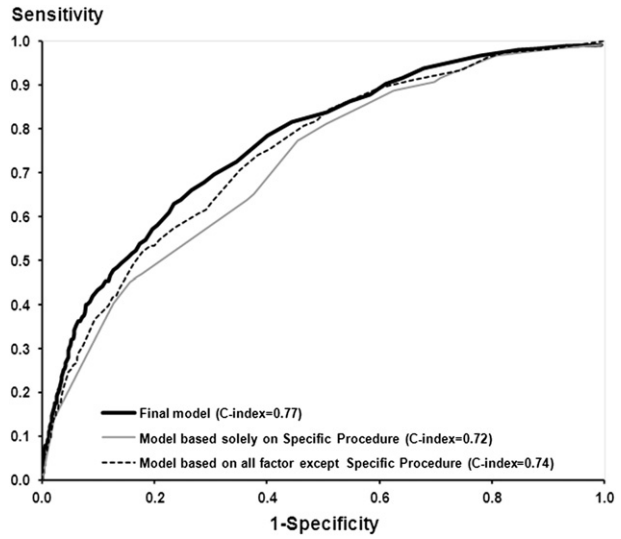

FIGURE 3. Receiver operating characteristic curve for 3 models evaluated in the validation set: the final risk model (C-index = 0.77), a model based solely on specific procedure (C-index = 0.72), and a model based on all factors in the final model *except* specific procedure (C-index of 0.74).

There is a need for a rolling program of recalibration for a model in routine use to account for anticipated improvements in outcomes over time[27] and any other evolving trends. Potential limitations of the current model arising from incomplete data, which tend to reflect the early years of CCAD methodology and user commitment to accurate and full data submission, could also be addressed by future reparameterization. Alongside a recalibration, a growing volume of records over time may support a model with a greater number of variables, for

example, including more diagnostic categories, and it is hoped that the completeness of comorbidity data will increase with time as clinicians perceive the relevance of this information to risk adjustment. Data quality improvements over time for antenatal diagnosis and Townsend deprivation score may allow these factors to be reconsidered.

The next step for this research is to facilitate and evaluate the near real-time routine risk-adjusted monitoring of 30-day outcomes in UK centers. Our ambition is to then complement this work on short-term outcomes by analyzing long-term outcomes among groups of patient defined by primary cardiac diagnosis, making use of the rich and unique source of tracked and validated outcomes available through CCAD. This would provide clinicians with valuable data with which to inform patients and caregivers and to assess services.

CONCLUSIONS

A risk model for pediatric cardiac surgery has been developed that can be used to partially adjust for case mix during routine monitoring of outcomes to assist quality assurance. Diagnostic and other patient information were found to add useful discriminatory power, increasing the amount of clinical data used in the model.

The authors thank Brian Reddy for the early work on the CCAD dataset at the start of this project and acknowledge the contribution of the CCAD steering committee, involved stakeholders, and patient groups for the creation and maintenance of the CCAD, which is a world-class audit database.

References

1. Aylin P, Alves B, Best N, Cook A, Elliott P, Evans SJ, et al. Comparison of UK paediatric cardiac surgical performance by analysis of routinely collected data 1984–96: was Bristol an outlier? *Lancet*. 2001;358:181-7.
2. Stark J, Gallivan S, Lovegrove J, Hamilton J, Monro J, Pollock J, et al. Mortality rates after surgery for congenital heart defects in children and surgeons' performance. *Lancet*. 2000;355:1004-7.
3. Spiegelhalter DJ. Mortality and volume of cases in paediatric cardiac surgery: retrospective study based on routinely collected data. *BMJ*. 2002;324:261-3.
4. National Health Service. Safe and sustainable: children's congenital cardiac services. National Health Service Specialist Services. 2011. Available at: http://www.specialisedservices.nhs.uk/safe_sustainable/childrens-congenital-cardiac-services. Accessed April 12, 2012.
5. Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. *Circulation*. 1989;79(6 Pt 2):I3-12.
6. Kang N, Cole T, Tsang V, Elliott M, de Leval M. Risk stratification in paediatric open-heart surgery. *Eur J Cardiothorac Surg*. 2004;26:3-11.
7. Jacobs ML, Jacobs JP, Franklin RCG, Mavroudis C, Lacour-Gayet F, Tchervenkov CI, 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):101-15.
8. Jenkins KJ, Beekman Iii RH, Bergersen LJ, Everett AD, Forbes TJ, Franklin RCG, 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):116-23.
9. Clarke DR, Breen LS, Jacobs ML, Franklin RCG, Tobota Z, Maruszewski B, et al. Verification of data in congenital cardiac surgery. *Cardiol Young*. 2008; 18(Suppl 2):177-87.
10. Franklin RCG, Jacobs JP, Krogmann ON, Béland MJ, Aiello VD, Colan SD, 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.
11. Bergersen L, Giroud JM, Jacobs JP, Franklin RCG, Béland MJ, Krogmann ON, et al. Report from The International Society for Nomenclature of Paediatric and Congenital Heart Disease: Cardiovascular Catheterisation for Congenital and Paediatric Cardiac Disease (Part 2-Nomenclature of Complications Associated with Interventional Cardiology). *Cardiol Young*. 2011;21:260-5.
12. Jacobs JP, Anderson RH, Weinberg PM, Walters HL 3rd, Tchervenkov CI, Del Duca D, et al. The nomenclature, definition and classification of cardiac structures in the setting of heterotaxy. *Cardiol Young*. 2007;17(Suppl 2):1-28.
13. Jenkins KJ, Gauvreau K, Newburger JW, Spray TL, Moller JH, Iezzoni LI. Consensus-based method for risk adjustment for surgery for congenital heart disease. *J Thorac Cardiovasc Surg*. 2002;123:110-8.
14. Lacour-Gayet F, Clarke D, Jacobs J, Gaynor W, Hamilton L, Jacobs M, et al. The Aristotle score for congenital heart surgery. *Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu*. 2004;7:185-91.
15. O'Brien SM, Clarke DR, Jacobs JP, Jacobs ML, Lacour-Gayet FG, Pizarro C, et al. An empirically based tool for analyzing mortality associated with congenital heart surgery. *J Thorac Cardiovasc Surg*. 2009;138:1139-53.
16. Central Cardiac Audit Database: paediatric analysis home page. Congenital heart disease website. Available at: www.ccad.org.uk. Accessed January 29, 2011.
17. International Paediatric and Congenital Cardiac Code (IPCCC) home page. Available at: www.ipccc.net. Accessed April 12, 2012.
18. Jacobs JP, O'Brien SM, Pasquali SK, Jacobs ML, Lacour-Gayet FG, Tchervenkov CI, et al. Variation in outcomes for benchmark operations: an analysis of the Society of Thoracic Surgeons Congenital Heart Surgery Database. *Ann Thorac Surg*. 2011;92:2184-92.
19. Brown KL, Crowe S, Pagel C, Bull C, Nagarajan M, Gibbs J, et al. Use of diagnostic information submitted to the United Kingdom Central Cardiac Audit Database: development of categorisation and allocation algorithms. *Cardiol Young*. October 2, 2012 [Epub ahead of print].
20. Townsend P, Simpson D, Tibbs N. Inequalities in health in the city of Bristol: a preliminary review of statistical evidence. *Int J Health Serv*. 1985;15:637-63.
21. Wellesley D, Boyd P, Dolk H, Pattenden S. An aetiological classification of birth defects for epidemiological research. *J Med Genet*. 2005;42:54-7.
22. Brown KL, Ridout DA, Hoskote A, Verhulst L, Ricci M, Bull C. Delayed diagnosis of congenital heart disease worsens preoperative condition and outcome of surgery in neonates. *Heart*. 2006;92:1298-302.
23. Harrell FE. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer; 2001.
24. Gallivan S, Utley M, Pagano D, Treasure T. MADCAP: a graphical method for assessing risk scoring systems. *Eur J Cardiothorac Surg*. 2006;29:431-3.
25. Shahian DM, Edwards FH, Jacobs JP, Prager RL, Normand S-LT, Shewan CM, et al. Public reporting of cardiac surgery performance: part 1—history, rationale, consequences. *Ann Thorac Surg*. 2011;92(3 Suppl):S2-11.
26. Shahian DM, Edwards FH, Jacobs JP, Prager RL, Normand S-LT, Shewan CM, et al. Public reporting of cardiac surgery performance: part 2—implementation. *Ann Thorac Surg*. 2011;92(3 Suppl):S12-23.
27. Tsang VT, Brown KL, Synnergren MJ, Kang N, de Leval MR, Gallivan S, et al. Monitoring risk-adjusted outcomes in congenital heart surgery: does the appropriateness of a risk model change with time? *Ann Thorac Surg*. 2009;87:584-7.

APPENDIX 1. MAPPING PRIMARY CARDIAC DIAGNOSIS TO DIAGNOSIS RISK GROUP

Each CCAD record contains up to 6 IPCCCs. To explore the potential for this information to add discriminatory power to risk adjustment, we developed a new hierarchical scheme that links the combination of IPCCCs available for a record to at most 1 of 24 primary cardiac diagnoses. The process for developing these diagnostic categories is described in detail by Brown and colleagues.[19] The mapping is shown of these 24 primary cardiac diagnoses (and the categories ''procedure'' and ''comorbidity'') to 1 of 3 diagnosis risk groups used in the risk model: low, medium, or high risk.

Primary cardiac diagnosis	Diagnosis risk group
Hypoplastic left heart syndrome	High risk
Functionally univentricular heart	Medium risk
Common arterial trunk (truncus arteriosus)	Medium risk
TGA + VSD/DORV, TGA type	Medium risk
Interrupted aortic arch	High risk
TGA (concordant AV and discordant VA connections) and intact ventricular septum	Medium risk
Pulmonary atresia with an intact ventricular septum	High risk
Pulmonary atresia + VSD (including Fallot type)	Medium risk
Atrioventricular septal defect	Low risk
Fallot/DORV-Fallot type	Low risk
Aortic valve stenosis (isolated)	Medium risk
Tricuspid valve abnormality (including Ebstein's)	Medium risk
Mitral valve abnormality (including supravalvar, subvalvar)	Medium risk
Totally anomalous pulmonary venous connection	Medium risk
Aortic arch obstruction ± VSD/ASD	Low risk
Pulmonary stenosis	Low risk
Subaortic stenosis (isolated)	Low risk
Aortic regurgitation	Low risk
VSD	Low risk
Interatrial communication (ASD)	Low risk
Patent ductus arteriosus	Medium risk
Miscellaneous congenital	Medium risk
Acquired	Medium risk
Procedure	Low risk
Comorbidity	High risk
Noncardiac or uncoded diagnosis	Medium risk

CCAD, Central Cardiac Audit Database; *IPCCC*, International Paediatric and Congenital Cardiac Code; *TGA*, transposition of the great arteries; *VSD*, ventricular septal defect; *DORV*, double outlet right ventricle; *AV*, atrioventricular; *VA*, ventriculoarterial; *ASD*, atrial septal defect.

APPENDIX 2. INTERNATIONAL PAEDIATRIC AND CONGENITAL CARDIAC CODE COMORBIDITY MAPPING

Listed are the IPCCCs recorded in the CCAD that are defined in the model as a non–Down syndrome comorbidity.

IPCCC (used in the CCAD)
030109. Position or morphology of thoracoabdominal organs abnormal
030305. Tracheobronchial anomaly
030703. Spleen absent (asplenia)
030704. Multiple spleens (polysplenia)
100665. Preprocedural endocarditis
101400. Secondary systemic hypertension
101402. Primary (essential) systemic hypertension
101444. Abdominal aorta aneurysm
101445. Rupture of thoracic aortic aneurysm
101446. Rupture of abdominal aortic aneurysm
101454. Descending aorta dissection and distal propagation (DeBakey type III/Stanford type B)
101460. Systemic arteritis
101505. Necrotizing enterocolitis

(Continued)

Continued

IPCCC (used in the CCAD)
101512. Meconium aspiration
102002. Preprocedural shock
102003. Preprocedural arrhythmia
102005. Preprocedural acidosis
102006. Preprocedural coagulation disorder
102007. Preprocedural renal failure
102008. Preprocedural renal failure requiring dialysis
102009. Preprocedural septicemia
102012. Preprocedural neurologic impairment
102014. Preprocedural mechanical ventilatory support
102015. Preprocedural mechanical circulatory support
102016. Preprocedural pulmonary hypertension
102017. Preprocedural tracheostomy
102018. Preprocedural seizures
102202. Premature birth
102203. Infant of diabetic mother
102205. Premature birth 32-35 wk
102206. Premature birth <32 wk
102300. Hereditary/noncardiac abnormality not apparent
102304. Hereditary disorder associated with heart disease
110635. Preprocedural complete AV block
140101. Chromosomal anomaly
140103. Trisomy 18, Edwards syndrome
140104. Trisomy 13, Patau syndrome
140105. 45XO, Turner syndrome
140121. 22q11 microdeletion
140200. Syndrome-association with cardiac involvement
140206. DiGeorge sequence
140217. Marfan syndrome
140219. Noonan syndrome
140228. Tuberous sclerosis
140230. Williams syndrome (infantile hypercalcemia)
140232. Fetal rubella syndrome
140266. Alagille syndrome: arteriohepatic dysplasia
140300. Noncardiac abnormality associated with heart disease
140304. Noncardiothoracic vascular abnormality
140305. Psychomotor developmental delay
140306. Cystic fibrosis
140307. Congenital diaphragmatic hernia
140308. Tracheoesophageal fistula
140310. Omphalocele
140311. Duodenal stenosis/atresia
140323. Renal abnormality
140329. Thoracic-mediastinal abnormality
140333. Microcephaly
140349. Tracheobronchial malacia
140404. Pectus carinatum
140405. Pectus excavatum
140409. Kyphoscoliosis
140412. Cleft lip or palate
140414. Anterior chest wall (pectus) deformity
140501. Maternal teratogen associated with congenital heart disease
140601. Multiple congenital malformations
160305. Lung disease
161001. Tracheal stenosis

(Continued)

Continued

IPCCC (used in the CCAD)
161009. Tracheal disease
163001. Respiratory failure

IPCCC, International Paediatric and Congenital Cardiac Code; *CCAD*, Central Cardiac Audit Database; *AV*, aortic valve.

APPENDIX 3. FINAL LOGISTIC REGRESSION RISK MODEL

Probability of death within 30 days $= \frac{1}{1+e^{-Z}}$, where

$$Z = -3.905 + 0.089^{*}age - 0.038^{*}weight + \sum_{i=1}^{39} B_i X_i.$$

Parameters $i = 1$–39 are tabulated below along with their corresponding regression coefficients, B_i, and the condition that must be satisfied for $X_i = 1$ ($X_i = 0$ otherwise). Note that patient age must be in units of years and patient weight in units of kilograms.

i	$X_i = 1$ **if condition satisfied** ($X_i = 0$ **otherwise**)	B_i
1	Specific procedure = anomalous coronary artery repair	0.583
2	Specific procedure = aortic valvotomy	1.222
3	Specific procedure = arterial switch (for isolated transposition)	–0.417
4	Specific procedure = arterial shunt	1.528
5	Specific procedure = arterial switch + VSD closure	0.508
6	Specific procedure = ASD repair	–1.234
7	Specific procedure = atrioventricular septal defect (complete) repair	0.135
8	Specific procedure = atrioventricular septal defect (partial) repair	–0.995
9	Specific procedure = aortic valve replacement, non-Ross	1.226
10	Specific procedure = aortic valve replacement, Ross	0.376
11	Specific procedure = bidirectional cavopulmonary shunt	–0.228
12	Specific procedure = Fontan procedure	0.536
13	Specific procedure = interrupted aortic arch repair	0.721
14	Specific procedure = isolated coarctation repair	0.135
15	Specific procedure = isolated pulmonary artery band	1.399
16	Specific procedure = low-volume group	0.879
17	Specific procedure = mitral valve replacement	1.602
18	Specific procedure = no specific procedure	1.114
19	Specific procedure = Norwood procedure (stage 1)	1.171
20	Specific procedure = PDA ligation (surgical)	0.640
21	Specific procedure = pulmonary atresia VSD repair	1.191
22	Specific procedure = pulmonary valve replacement	0.916
23	Specific procedure = Rastelli procedure	–16.501
24	Specific procedure = repair of total anomalous pulmonary venous drainage	0.638
25	Specific procedure = subvalvar aortic stenosis repair	0.789
26	Specific procedure = supravalvar aortic stenosis repair	0.520
27	Specific procedure = truncus arteriosus repair	0.902
28	Specific procedure = tetralogy repair	0.783
29	Specific procedure = VSD repair	–0.139
30	Procedure type = bypass	0.715
31	Diagnosis group = low risk	–0.588
32	Diagnosis group = medium risk	0.222
33	Diagnosis group = high risk	0.366
34	Not identified as univentricular heart	–0.446
35	No recorded non–Down syndrome comorbidities	–0.579
36	Age group = child	–0.797
37	Age group = infant	0.157
38	Age group = neonate	0.640
39	Procedure performed pre-2007	0.257

We note that caution is needed when interpreting individual coefficients because these are not clinically meaningful when taken in isolation of the other risk factors. The predicted risk comes from the combination of procedure, age, weight, severity of diagnosis, and comorbidity information. *VSD*, Ventricular septal defect; *ASD*, atrial septal defect; *PDA*, patent ductus arteriosus.

The most up-to-date version of the PRAiS model specification will always be available from the University College London Clinical Operational Research Unit Web site (http://www.ucl.ac.uk/operational-research/AnalysisTools/PRAiS).