Improving Risk Adjustment for Mortality After Pediatric Cardiac Surgery: The UK PRAiS2 Model

Libby Rogers, MMath, Katherine L. Brown, MPH, MD, Rodney C. Franklin, FRCP, MD, Gareth Ambler, PhD, David Anderson, FRCS, MD, David J. Barron, FRCS, MD, Sonya Crowe, PhD, Kate English, MRCP, MBChB, John Stickley, BS, Shane Tibby, MBChB, MSc, Victor Tsang, FRCS, MS, Martin Utley, PhD, Thomas Witter, RN, and Christina Pagel, PhD

Clinical Operational Research Unit, University College London, London; Cardiac, Critical Care and Respiratory Division, Great Ormond Street Hospital NHS Foundation Trust, London; Pediatric Cardiology, Royal Brompton and Harefield NHS Foundation Trust, London; Department of Statistical Science, University College London, London; Cardiology and Critical Care, Evelina London Children's Hospital, London; Cardiothoracic Surgery, Birmingham Children's Hospital, Birmingham; and Department of Congenital Cardiology, Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom

Background. **Partial Risk Adjustment in Surgery (PRAiS), a risk model for 30-day mortality after children's heart surgery, has been used by the UK National Congenital Heart Disease Audit to report expected risk-adjusted survival since 2013. This study aimed to improve the model by incorporating additional comorbidity and diagnostic information.**

Methods. **The model development dataset was all procedures performed between 2009 and 2014 in all UK and Ireland congenital cardiac centers. The outcome measure was death within each 30-day surgical episode. Model development followed an iterative process of clinical discussion and development and assessment of models using logistic regression under 25 × 5 cross-validation. Performance was measured using Akaike information criterion, the area under the receiver-operating characteristic curve (AUC), and calibration. The final model was assessed in an external 2014 to 2015 validation dataset.**

Results. **The development dataset comprised 21,838 30-day surgical episodes, with 539 deaths (mortality, 2.5%). The validation dataset comprised 4,207 episodes, with 97 deaths (mortality, 2.3%). The updated risk model included 15 procedural, 11 diagnostic, and 4 comorbidity groupings, and nonlinear functions of age and weight. Performance under cross-validation was: median AUC of 0.83 (range, 0.82 to 0.83), median calibration slope and intercept of 0.92 (range, 0.64 to 1.25) and –0.23 (range, –1.08 to 0.85) respectively. In the validation dataset, the AUC was 0.86 (95% confidence interval [CI], 0.82 to 0.89), and the calibration slope and intercept were 1.01 (95% CI, 0.83 to 1.18) and 0.11 (95% CI, –0.45 to 0.67), respectively, showing excellent performance.**

Conclusions. **A more sophisticated PRAiS2 risk model for UK use was developed with additional comorbidity and diagnostic information, alongside age and weight as nonlinear variables.**

(Ann Thorac Surg 2017;104:211–9)

Approximately 3,500 children under 16 years of age have heart surgery each year in the United Kingdom [1] and since 2000 all cardiac centers have contributed procedure data to the National Congenital Heart Disease Audit (NCHDA). Center-specific mortality outcomes for individual procedures have been published online since 2007 by the NCHDA [2].

With the accumulation of registry based data, consensus-based risk stratification methods for pediatric cardiac surgery, such as the Risk Adjusted classification for Congenital Heart Surgery (RACHS-1) [3] and Aristotle [4] methods, have given way to empirical approaches including the Society of Thoracic Surgeons–European Association of Cardiothoracic Surgery (STS-EACTS) score [5]. The Partial Risk Adjustment in Surgery (PRAiS) risk model for 30-day mortality after pediatric cardiac surgery was developed in 2011, using 10 years of UK audit data [6, 7] alongside accompanying software to

The Supplemental Material can be viewed in the online version of this article [http://dx.doi.org/10.1016/j.athoracsur.2016.12.014] at http://www.annalsthoracicsurgery.org.

Accepted for publication Dec 12, 2016.

Address correspondence to Dr Rogers, Clinical Operational Research Unit, University College London, 4 Taviton St, London, WC1H 0BT, United Kingdom; email: libby.rogers@ucl.ac.uk.

0003-4975/$36.00
http://dx.doi.org/10.1016/j.athoracsur.2016.12.014

implement it for routine monitoring [8], and has been used by the NCHDA since 2013.

The PRAiS model incorporated information about procedure (29 categories), cardiac diagnosis (3 risk categories), number of functioning ventricles, age category (neonate, infant, child), age, and weight as continuous variables, as well as presence of a non–Down syndrome comorbidity and whether surgery was performed on cardiopulmonary bypass. As the PRAiS model began to be used to monitor outcomes, data quality for information previously collected but not actively used (eg, comorbidity and diagnosis codes) improved rapidly.

This study uses recent NCHDA data to improve the PRAiS model by incorporating more detailed information about comorbid conditions, acuity, and diagnosis. In this paper, the original model will be referred to as the PRAiS1 model and the new model as the PRAiS2 model.

Material and Methods

Data

NCHDA data are subject to external validation and survival status is independently verified for patients resident in England or Wales by the Health and Social Care Information Centre.

The development dataset consisted of all pediatric cardiac procedures performed in the United Kingdom and Ireland from April 2009 to March 2014. Records were excluded if the patient was over 16 years of age; the procedure was noncardiac, nonsurgical, or minor; or it was performed at 1 small-volume center that ceased pediatric cardiac surgery in 2010.

In February 2016, the NCHDA provided a further year of procedure data for April 2014 to March 2015: this was used as an external validation dataset. The final model was recalibrated on all data from 2009 to 2015.

The unit of analysis was a 30-day surgical episode [6, 7]. For each patient an episode started with their first surgical procedure. Any further surgical procedures that the same patient underwent within 30 days of this first procedure were not included. The next surgical procedure recorded for the same patient more than 30 days after the first surgical procedure was treated as the start of a new, independent 30-day episode. The outcome measure was death within each 30-day surgical episode. Where the life status at 30 days could not be determined, discharge status prior to 30 days was used.

Data Cleaning

Duplicate records and inconsistencies in age or life status within or across records relating to the same patient were identified, removed, or resolved.

Episodes in the development dataset with an absolute weight-for-age Z-score of 5 or over or with a weight judged to be infeasible by a clinician or a missing weight were assigned the mean weight-for-age. To mimic prospective use, no adjustments of weight were made when an episode was being used for validation.

Risk Factors Considered

PROCEDURES. A Specific Procedure hierarchical algorithm was developed by the NCHDA to sort submitted individual European Pediatric Cardiac Code (EPCC) [9] codes into a Specific Procedure category. We used the Specific Procedure Algorithm Coding v5.05 from May 2016 [10], which includes 49 Specific Procedures.

In the PRAiS1 model, only bypass and nonbypass surgical procedures were included. For the PRAiS2 model the hybrid procedure for hypoplastic left heart syndrome (HLHS) [11] was also incorporated.

DIAGNOSIS. A scheme that maps the individual EPCC diagnosis codes recorded for each episode to 1 of 29 hierarchical diagnostic categories was used [12, 13] (see Supplemental Material 1 for details). An increase in the number of diagnostic groupings from the 3 broad groups used in the PRAiS1 model [6] was explored.

A surgical episode was determined to relate to a patient with a functionally univentricular heart if they had a diagnosis or procedure code suggesting single-ventricle physiology.

COMORBIDITY. Each episode can have several different EPCC comorbidity codes recorded. Possible ways to include more comorbidity information were explored as discussed in detail in Brown and colleagues [14].

AGE AND WEIGHT. The association of both age and weight with the log odds of mortality is nonlinear. The PRAiS1 model used a continuous linear weight and age and neonate, infant and child age bands as categorical risk factors [6]. However, this resulted in unrealistic jumps in predicted risk at age band boundaries. To better capture this nonlinearity and explore any interaction between age and weight, we considered age and weight categories in combination with continuous age, weight, and weight-for-age Z-scores similarly to the STS model [15], cubic splines [16], and fractional polynomials [17].

Model Development Strategy

As raw mortality rate is low (<3%), there is a practical limit to how many parameters can be reasonably included in an empirical risk model. A common rule of thumb is that overfitting is alleviated if the number of deaths in the dataset is at least 10 times larger than the number of parameters [18]. This gave a practical limit of about 40 parameters that could be included in the PRAiS2 model, necessitating a trade-off between adding comorbidity and diagnostic information and reducing Specific Procedure information, particularly given the increase in the number of individual procedure categories included in the Specific Procedure algorithm, including several lower volume categories.

An expert advisory panel from 5 hospitals comprising 3 surgeons (V.T., D.A., D.J.B.), 2 cardiologists (K.E., R.C.F.), 2 intensivists (K.L.B., S.T.), and 2 data management experts (T.W., J.S.) was assembled to consider the relative importance of comorbidities and potential trade-offs in included risk factors.

Analysis

Groupings of comorbidity codes were determined using the clinical and coding expertise of the panel, informed by

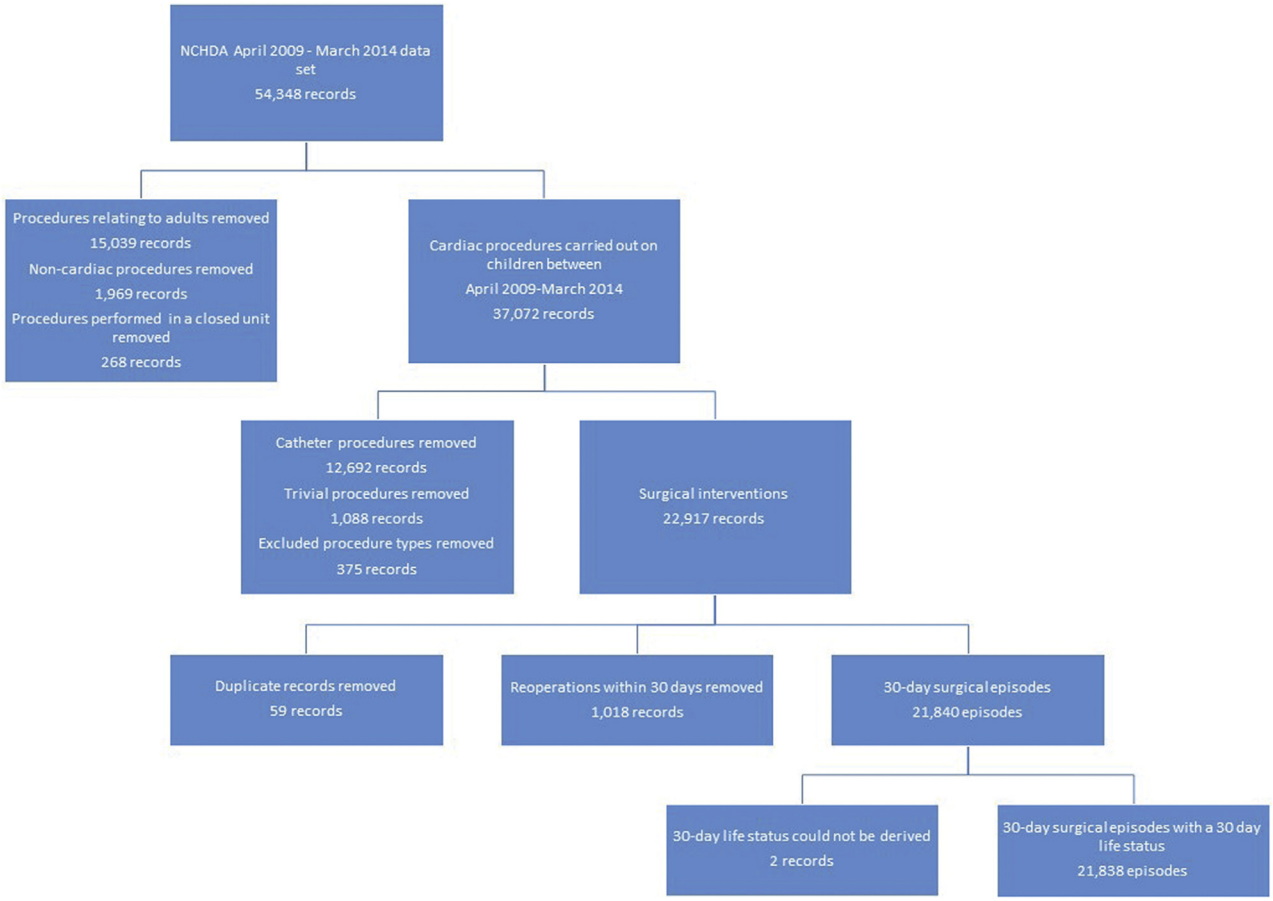

Fig 1. Inclusions and exclusions from the development dataset. (NCHDA = National Congenital Heart Disease Audit.)

univariate and multivariate associations of candidate comorbidity groupings with mortality [14].

Classification and regression tree analysis [19] was used to define initial broader groupings of procedure and diagnostic categories, taking into account patient age and observed mortality. These groups were adjusted iteratively after consultation with the clinical expert panel and testing of the performance of the resultant risk models under cross-validation. The expert panel provided advice on procedures and diagnoses where the observed mortality was not considered representative of the risk, or where procedures or diagnoses that were qualitatively different should not be combined in a single group.

Multiple logistic regression was used within the Stata statistical software package (StataCorp 2013, Release 13, College Station, TX).

Comparison of candidate models was performed within the development dataset using the area under the receiver-operating characteristic curve (AUC) and Akaike information criterion.

Fivefold cross-validation repeated 25 times (25×5 cross-validation) was then carried out [20], with the data splits stratified by year and unit to ensure a representative case mix. The median and range of the AUC over the 25 repeats of the cross-validation was calculated [21]. The calibration slope (β) and intercept (α) were estimated in each of the 125 test sets [22], and the median and range were calculated. β gives an indication of whether the model is underfitting or overfitting the data, and α indicates whether the model is underpredicting or overpredicting deaths. If a model is perfectly calibrated, $\beta = 1$ and $\alpha = 0$ [23].

A comparison was also made to the PRAiS1 model recalibrated on the development dataset to ensure that improvements in model performance were achieved.

Finally, any episodes that had particularly large leverage or influence [16] were investigated to see whether they were genuine outliers or data errors.

Performance of the final selected model is reported as the median and range of the AUC, calibration slope, and calibration intercept over the cross-validated test sets and the AUC, calibration slope, and intercept are reported with 95% confidence intervals (CI) in the external validation set.

Results

The cleaned development dataset comprised 21,838 30-day surgical episodes, of which 539 resulted in death within 30 days (mortality rate, 2.5%). Figure 1 gives a summary of the records excluded from the development dataset.

Table 1. Specific Procedures Included in Each Specific Procedure Group, With Their Respective Frequencies and 30-Day Mortalities

Specific Procedure Groups	Frequency	30-Day Mortality
Group 1	633 (2.9)	11.1
Norwood Procedure (Stage 1)	589 (2.7)	10.7
HLHS hybrid approach	44 (0.2)	15.9
Group 2	414 (1.9)	7.2
TAPVC repair + arterial shunt	10 (0.0)	60.0
Truncus and interruption repair	15 (0.1)	6.7
Truncus arteriosus repair	190 (0.9)	5.3
Interrupted aortic arch repair	118 (0.5)	5.1
Arterial switch + aortic arch obstruction repair (with-without VSD closure)	81 (0.4)	8.6
Group 3	760 (3.5)	7.8
Arterial shunt	760 (3.5)	7.8
Group 4	1,171 (5.4)	3.9
Repair of total anomalous pulmonary venous connection	329 (1.5)	5.2
Arterial switch + VSD closure	311 (1.4)	2.6
Isolated pulmonary artery band	531 (2.4)	4.0
Group 5	1,885 (8.6)	4.1
PDA ligation (surgical)	1,885 (8.6)	4.1
Group 6	1,996 (9.1)	1.2
Arterial switch (for isolated transposition)	724 (3.3)	1.5
Isolated coarctation/hypoplastic aortic arch repair	1,236 (5.7)	1.0
Aortopulmonary window repair	36 (0.2)	0.0
Group 7	667 (3.1)	4.6
Senning or Mustard procedure	16 (0.1)	12.5
Ross-Konno procedure	69 (0.3)	2.9
Mitral valve replacement	164 (0.8)	3.7
Pulmonary vein stenosis procedure	96 (0.4)	6.3
Pulmonary atresia VSD repair	201 (0.9)	4.5
Tetralogy with absent pulmonary valve repair	48 (0.2)	4.2
Unifocalization procedure (with/without shunt)	73 (0.3)	5.5
Group 8	1,508 (6.9)	2.7
Heart transplant	152 (0.7)	3.3
Tricuspid valve replacement	17 (0.1)	5.9
Aortic valve repair	292 (1.3)	2.1
Pulmonary valve replacement	328 (1.5)	1.8
Aortic root replacement (not Ross)	59 (0.3)	3.4
Cardiac conduit replacement	167 (0.8)	1.8
Isolated RV to PA conduit construction	400 (1.8)	3.3
Tricuspid valve repair	93 (0.4)	4.3
Group 9	369 (1.7)	3.5
Multiple VSD closure	59 (0.3)	1.7
Atrioventricular septal defect and tetralogy repair	50 (0.2)	2.0
Cor triatriatum repair	54 (0.2)	5.6
Supravalvar aortic stenosis repair	102 (0.5)	3.9
Rastelli-REV procedure	104 (0.5)	3.8
Group 10	1,156 (5.3)	1.7
Bidirectional cavopulmonary shunt	1,156 (5.3)	1.7
Group 11	889 (4.1)	1.0
Atrioventricular septal defect (complete) repair	889 (4.1)	1.0
Group 12	989 (4.5)	1.0
Fontan procedure	989 (4.5)	1.0

(*Continued*)

Table 1. Continued

Specific Procedure Groups	Frequency	30-Day Mortality
Group 13	1,264 (5.8)	0.4
Aortic valve replacement: Ross	151 (0.7)	0.7
Subvalvar aortic stenosis repair	609 (2.8)	0.3
Mitral valve repair	235 (1.1)	0.4
Sinus venosus ASD and/or PAPVC repair	269 (1.2)	0.4
Group 14	2,152 (9.9)	0.6
Atrioventricular septal defect (partial) repair	396 (1.8)	0.5
Tetralogy of Fallot–type DORV repair	1,521 (7.0)	0.7
Vascular ring procedure	235 (1.1)	0.4
Group 15	2,868 (13.1)	0.1
Anomalous coronary artery repair	94 (0.4)	0.0
Aortic valve replacement: non-Ross	93 (0.4)	0.0
ASD repair	941 (4.3)	0.0
VSD repair	1,740 (8.0)	0.2
No specific procedure group	3,117 (14.3)	2.9
No specific procedure	3,117 (14.3)	2.9

Values are n (%) or %.

ASD = atrial septal defect; DORV = double outlet right ventricle; HLHS = hypoplastic left heart syndrome; PA = pulmonary artery; PDA = patent ductus arteriosus; PAPVC = partial anomalous pulmonary venous connection; REV = réparation à létage ventriculaire; RV = right ventricle; TAPVC = total anomalous pulmonary venous connection; VSD = ventricular septal defect.

Missing data was not a significant issue [6]—fewer than 0.6% of records were missing a relevant field.

Of the final 21,838 episodes, 5 episodes had their weights adjusted from grams to kilograms and 33 episodes with anomalous or missing weight were assigned the mean weight for their age.

For 3,165 episodes (14.5%), the discharge status prior to 30 days was used to generate outcome as certified life status was unavailable (overwhelmingly for overseas patients lacking a UK national health identifier).

The validation dataset comprised 4,207 episodes with 97 deaths (mortality rate, 2.3%). To mimic prospective use, no adjustments to missing or anomalous weights were made. For 724 episodes (17.2%), the discharge status prior to 30 days was used as a proxy for 30-day life status.

New Comorbidity, Procedure, and Diagnosis Groups

The comorbidity factors included in the model were: congenital comorbidity; acquired comorbidity; severity of illness; and additional cardiac risk factor.

The 49 specific procedures were combined into 15 procedure groups as well as a no Specific Procedure category; these are shown in Table 1. The 29 diagnosis categories were combined into 11 diagnosis groups; these are shown in Table 2.

The groups were determined on a basis of complexity, risk, age, any association between procedures and diagnosis, and the clinical validity of the groups. For instance, arterial shunts, which are performed on a wide variety of patients, were retained as a separate group.

The PRAiS2 model includes similar additional risk factors to the STS model variables [24] of noncardiac congenital anatomic abnormality, chromosomal abnormality or syndrome, and preoperative factors. The treatment of procedural and diagnostic information in the 2 models differs, with the STS model including a separate intercept parameter for each combination of age group and primary procedure, and no diagnostic information. The PRAiS2 procedural and diagnostic groups were determined accounting for interactions between procedures, diagnoses, risk and age, and clinical validity.

Other Risk Factors

The HLHS hybrid approach procedures were included in the nonbypass procedure category.

Fractional polynomials of the form $ax + b\sqrt{x}$ were best able to capture the relationship of both age and weight with mortality, and were much preferred by the clinical expert panel to avoid artifactual changes in estimates of risk at age or weight category boundaries. Similarly to the STS model, age-for-weight Z-scores did not improve the model.

Mortality has been falling in the United Kingdom and Ireland [25], with particular reduction post-2012. Similar to the PRAiS1 model, a binary post-2012 epoch variable was included.

Once the final model risk factors had been chosen, shrinkage [22] was considered. There was little indication of overfitting and shrinkage did not significantly improve calibration under cross-validation.

Final Risk Model

The final risk model was a logistic regression model with the following variables: age; weight; procedure group; diagnosis group; procedure type; univentricular heart indicator; congenital comorbidity indicator, acquired comorbidity indicator; severity of illness indicator; additional

Table 2. Diagnoses Included in Each Diagnosis Group, With Their Respective Frequencies and 30-Day Mortality Rates

Diagnosis Groups	Frequency	30-Day Mortality
Group 1	2,034 (9.3)	6.4
HLHS	1,401 (6.4)	6.5
Truncus arteriousus	398 (1.8)	4.8
Pulmonary atresia and IVS	235 (1.1)	8.9
Group 2	2,569 (11.8)	4.6
Functionally UVH	1,436 (6.6)	4.5
Pulmonary atresia and VSD	1,133 (5.2)	4.7
Group 3	1,380 (6.3)	3.3
TGA + VSD/DORV-TGA	1,171 (5.4)	3.3
Interrupted aortic arch	209 (1.0)	2.9
Group 4	1,724 (7.9)	4.1
PDA	1,724 (7.9)	4.1
Group 5	1,869 (8.6)	2.4
Miscellaneous primary congenital diagnosis	1,248 (5.7)	2.2
Tricuspid valve abnormality (including Ebstein's)	219 (1.0)	2.7
TAPVC	302 (1.4)	3.0
No diagnosis given	100 (0.5)	2.0
Group 6	447 (2.0)	2.5
Acquired	447 (2.0)	2.5
Group 7	3,711 (17.0)	1.5
AVSD	1,737 (8.0)	2.0
Fallot/DORV Fallot	1,974 (9.0)	1.1
Group 8	1,273 (5.8)	1.6
Aortic valve stenosis (isolated)	517 (2.4)	1.9
Mitral valve abnormality	496 (2.3)	1.0
Miscellaneous congenital terms	260 (1.2)	1.9
Group 9	544 (2.5)	2.0
TGA + IVS	544 (2.5)	2.0
Group 10	2,022 (9.3)	0.7
Aortic arch obstruction + VSD/ASD	1,582 (7.2)	0.8
Pulmonary stenosis	440 (2.0)	0.5
Group 11	4,265 (19.5)	0.4
Subaortic stenosis (isolated)	265 (1.2)	0.0
Aortic regurgitation	217 (1.0)	0.0
VSD	2,389 (10.9)	0.7
ASD	1,269 (5.8)	0.1
Arrhythmia	125 (0.6)	0.8

Values are n (%) or %.

ASD = atrial septal defect; AVSD = atrioventricular septal defect; DORV = double outlet right ventricle; HLHS = hypoplastic left heart syndrome; IVS = intact ventricular septum; PDA = patent ductus arteriosus; TAPVC = total anomalous pulmonary venous connection; TGA = transposition of the great arteries; UVH = univentricular heart; VSD = ventricular septal defect.

cardiac risk factor indicator; and 2013 onward indicator. The frequency and 30-day mortality rates in the development datasets for the categorical risk factors included in the model are shown in Table 3. As the Specific Procedure and diagnosis groups are grouped by age and similarity as well as risk, the odds ratios compared to the reference group are not informative and so are not included.

Under cross-validation, the model had a median AUC of 0.83 (range, 0.82 to 0.83), showing excellent discrimination. There was only slight evidence of overfitting, with a median calibration slope of 0.92 (range, 0.64 to 1.25) and median calibration intercept of –0.23 (range, 1.08 to 0.85; perfect calibration slope = 1 and intercept = 0), indicating slight underprediction.

The recalibrated PRAiS1 model had a median AUC of 0.80 (range, 0.80 to 0.81), and a median calibration slope and intercept of 0.90 (range, 0.63 to 1.31) and –0.29 (range, –1.15 to 1.06), respectively, and so was outperformed by the PRAiS2 model across all measures.

In the external validation set, there were 97 observed deaths in the test set compared with 89 predicted. The AUC was 0.86 (95% CI, 0.82 to 0.89) (Fig 2) and the calibration slope and intercept were 1.01 (95% CI, 0.83, 1.18) and 0.11 (95% CI, –0.45 to 0.67), respectively. The model

Table 3. The Frequency, Mortality, OR, and p Value for the Final Categorical Risk Factors

Risk Factor	Records (N = 21,838)	Deaths (%)	OR (95% CI)	*p* Value
Specific procedure group				
Group 1	633	70 (11.1)	…	…
Group 2	414	30 (7.2)	…	…
Group 3	760	59 (7.8)	…	…
Group 4	1,171	46 (3.9)	…	…
Group 5	1,885	78 (4.1)	…	…
Group 6	1,996	23 (1.2)	…	…
Group 7	667	31 (4.6)	…	…
Group 8	1,508	40 (2.7)	…	…
Group 9	369	13 (3.5)	…	…
Group 10	1,156	20 (1.7)	…	…
Group 11	889	9 (1.0)	…	…
Group 12	989	10 (1.0)	…	…
Group 13	1,264	5 (0.4)	…	…
Group 14	2,152	13 (0.6)	…	…
Group 15	2,868	3 (0.1)	…	…
No specific procedure	3,117	89 (2.9)	…	…
Diagnosis group			…	…
Group 1	2,034	131 (6.4)	…	…
Group 2	2,569	117 (4.6)	…	…
Group 3	1,380	45 (3.3)	…	…
Group 4	1,724	70 (4.1)	…	…
Group 5	1,869	45 (2.4)	…	…
Group 6	447	11 (2.5)	…	…
Group 7	3,711	56 (1.5)	…	…
Group 8	1,273	20 (1.6)	…	…
Group 9	544	11 (2.0)	…	…
Group 10	2,022	15 (0.7)	…	…
Group 11	4,265	18 (0.4)	…	…
Procedure type				
Nonbypass	5,032	176 (3.5)	Reference	…
Bypass	16,806	363 (2.2)	1.5 (1.1–2.1)	0.013
UVH status				
Not UVH	18,101	352 (1.9)	Reference	…
UVH	3,737	187 (5.0)	1.9 (1.4–2.7)	<0.001
Severity of illness (acuity)				
No indication of severe illness	19,578	369 (1.9)	Reference	…
Indication of severe illness	2,260	170 (7.5)	1.7 (1.4–2.2)	<0.001
Acquired cardiac risk factor				
No indication of acquired cardiac risk factor	20,785	483 (2.3)	Reference	…
Indication of acquired cardiac risk factor	1,053	56 (5.3)	2.2 (1.6–3.0)	<0.001
Acquired comorbidity				
No indication of acquired comorbidity	20,584	458 (2.2)	Reference	…
Indication of acquired comorbidity	1,254	81 (6.5)	2.0 (1.5–2.6)	<0.001
Congenital comorbidity				
No indication of congenital comorbidity	19,393	449 (2.3)	Reference	…
Indication of a congenital comorbidity	2,445	90 (3.7)	1.5 (1.2–1.9)	0.001
Procedure year				
Pre-2013	16,259	429 (2.6)	Reference	…
2013 onward	5,579	110 (2.0)	0.7 (0.6–0.9)	0.003

CI = confidence interval; OR = odds ratio; UVH = univentricular heart.

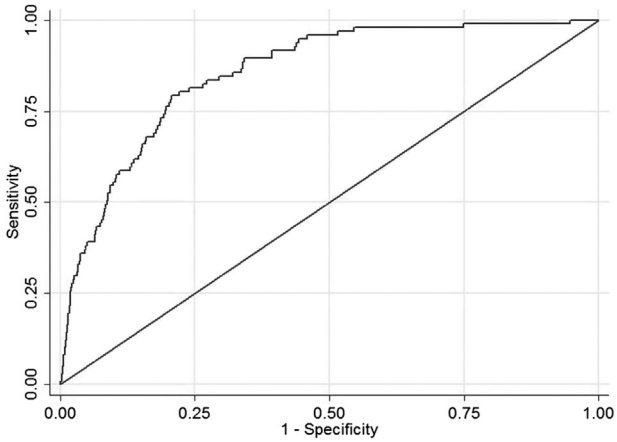

Fig 2. The receiver-operating characteristic curve of the model tested on the validation dataset (area under the receiver-operating characteristic curve = 0.86).

showed excellent performance, comparing well with the STS score (AUC, 0.86) [15].

The final model was recalibrated on all of the data from 2009 to 2015. Full details can be found in Supplemental Material 2.

Comment

This study has produced an improved PRAiS risk model by using more information about comorbidity and diagnosis while reducing the total number of model parameters from 38 to 36. Two independent risk models (STS and PRAiS) using different national datasets have now resulted in similar risk factors and a similar methodology of close iterative interaction between analysts and experts from the clinical community.

The clinical validity of the risk factors to be included in the PRAiS2 model was discussed in considerable detail by the expert panel, as were known case mix differences between centers. In particular, the clinical experience of the panel added value to the empirical evidence to define the final comorbidity and additional risk factors and the final broader Specific Procedure and diagnosis groupings. The close involvement of the expert panel representing many hospitals and different specialties was very successful, allowed careful consideration of how individual codes have been, and will be used by centers in practice and built trust in the final model within the clinical community.

Despite the excellent performance of the PRAiS2 model, there is 1 case mix feature present within the model that needs highlighting: risk for the HLHS hybrid procedure, generally performed on the sickest patients, is underpredicted. Of the 60 HLHS hybrid procedures in the 2009 to 2015 data, there were 10 observed deaths but only 6 predicted deaths. This additional risk cannot yet be well captured in the PRAiS2 model due to low case volume. Additionally, this procedure has predominantly been performed at 1 center, which could therefore be unfairly affected in the prospective national audit. The expert panel recommended including HLHS hybrid procedures within the PRAiS2 model despite these issues as this procedure is expected to become more common. It has been proposed to the NCHDA steering committee that the PRAiS2 model for future national audit is run twice on data including and excluding hybrid HLHS procedures in turn so that this effect can be accounted for. As this procedure becomes more common, the PRAiS2 model will be recalibrated to better reflect this particular risk.

This case mix feature highlights the importance of the continual process of use and collaborative development of this and similar models, as the quality and quantity of the data available increases and as clinical practice evolves.

The NCHDA has approval from the UK Health Research Authority to hold and use data for audit and quality assurance and further requirement for research ethics committee approval was waived. The NCHDA is funded by the National Clinical Audit and Patient Outcomes Programme, administered by the Healthcare Quality Improvement Partnership (HQIP). This project was funded by the National Institute for Health Research Health Services and Delivery Research programme (Project No. 14/19/13). The views and opinions expressed therein are those of the authors and do not necessarily reflect those oaf the NIHR HS&DR programme or the Department of Health. K. Brown and V. Tsang were support by the National Institute for Health Research Biomedical Research Center at Great Ormond Street Hospital for Children NHS Foundation Trust and University College London.

References

1. Cunningham D, Franklin R, Bridgewater B, Deanfield JE. NICOR Investigation of mortality from paediatric cardiac surgery in England 2009-12. 2013. Available at https://nicor4.nicor.org.uk/CHD/an_paeds.nsf/vwContent/Minutes%20and%20Newsletters?Opendocument. Accessed January 11, 2017.
2. NICOR. NICOR, specific procedures national data 2013. Available at https://nicor4.nicor.org.uk/chd/an_paeds.nsf/WBenchmarksYears?openview&RestrictToCategory=2015&start=1&count=500. Accessed January 11, 2017.
3. 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.
4. Lacour-Gayet F, Clarke D, Jacobs J, et al. The Aristotle score for congenital heart surgery. Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu 2004;7:185–91.
5. O'Brien SM, Clarke DR, Jacobs JP, et al. An empirically based tool for analyzing mortality associated with congenital heart surgery. J Thorac Cardiovasc Surg 2009;138:1139–53.
6. Pagel C, Brown KL, Crowe S, Utley M, Cunningham D, Tsang VT. A mortality risk model to adjust for case mix in UK paediatric cardiac surgery. Health Serv Deliv Res 2013; 1.1. Available at https://www.ncbi.nlm.nih.gov/books/NBK259324/. Accessed January 11, 2017.
7. Crowe S, Brown KL, Pagel C, et al. Development of a diagnosis- and procedure-based risk model for 30-day outcome after pediatric cardiac surgery. J Thorac Cardiovasc Surg 2013;145:1270–8.
8. Pagel C, Utley M, Crowe S, et al. Real time monitoring of risk-adjusted paediatric cardiac surgery outcomes using variable

life-adjusted display: implementation in three UK centres. Heart 2013;99:1445–50.
9. International Society for Nomenclature of Paediatric and Congenital Heart Disease AEPC version of the IPCCC - Short List (version April 15, 2015). Available at http://ipccc.net/ipccc-download-form/. Accessed September 20, 2016.
10. NICOR. National Institute for Cardiovascular Outcomes Research. Technical Information. University College London. Available at https://nicor4.nicor.org.uk/chd/an_paeds.nsf/vwContent/Technical%20Information?Opendocument. Accessed May 25, 2016.
11. National Institute for Health Care Excellence. Hybrid procedure for interim management of hypoplastic left heart syndrome in neonates. Guidance and guidelines. 2007. Available at https://www.nice.org.uk/guidance/ipg246. Accessed July 11, 2016.
12. Brown KL, Crowe S, Pagel C, et al. Use of diagnostic information submitted to the United Kingdom Central Cardiac Audit Database: development of categorisation and allocation algorithms. Cardiol Young 2013;23:491–8.
13. Brown KL, Wray J, Knowles RL, et al. Infant deaths in the UK community following successful cardiac surgery: building the evidence base for optimal surveillance, a mixed-methods study. Southampton, UK: NIHR Journals Library, 2016. Available at https://www.ncbi.nlm.nih.gov/pubmed/27252995. Accessed January 11, 2017.
14. Brown KL, Rogers L, Barron DJ, et al. Incorporating comorbidity within risk adjustment for UK pediatric cardiac surgery. Ann Thorac Surg 2017 (in press).
15. 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;100:1054–62.
16. Harrell FE. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer; 2001.
17. Royston P, Altman DG. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. J R Stat Soc Ser C Appl Stat 1994;43:429–67.
18. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol 1995;48:1503–10.
19. Hastie T, Tibshirani R, Friedman JH. Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd edition. New York: Springer-Verlag; 2009.
20. Steyerberg EW, Harrell FE Jr, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54:774–81.
21. Newson R. Confidence intervals for rank statistics: Somers' D and extensions. Stata J 2006;6:309–34; Available at http://www.stata-journal.com/article.html?article=snp15_6. Accessed July 7, 2016.
22. Steyerberg E. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. 2009 edition. New York: Springer; 2008.
23. Cox DR. Two further applications of a model for binary regression. Biometrika 1958;45:562–5.
24. 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;100:1063–8, discussion 1068–70.
25. Brown KL, Crowe S, Franklin R, et al. Trends in 30-day mortality rate and case mix for paediatric cardiac surgery in the UK between 2000 and 2010. Open Heart 2015;2:e000157.