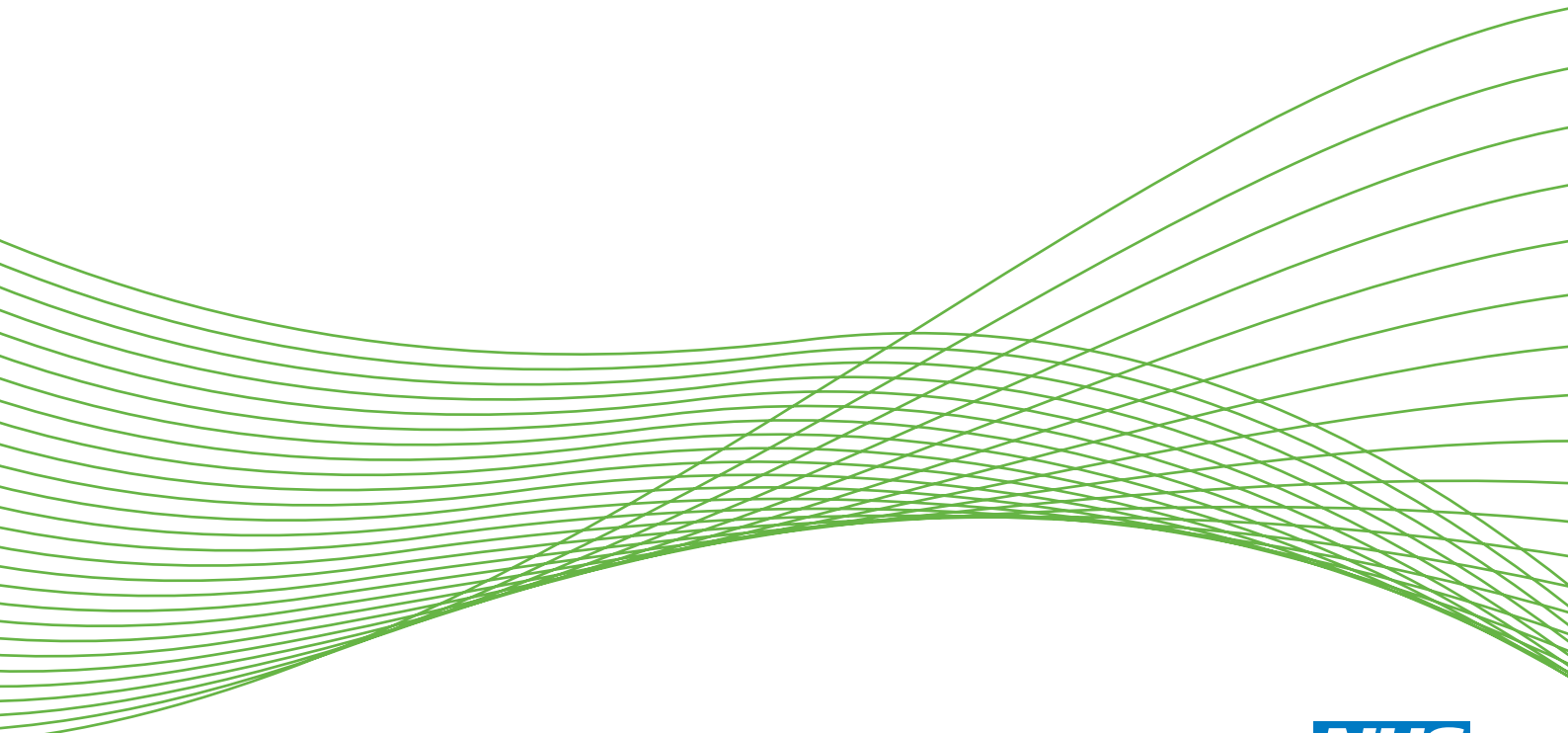


## A mortality risk model to adjust for case mix in UK paediatric cardiac surgery

*C Pagel, KL Brown, S Crowe, M Utley, D Cunningham and VT Tsang*



***National Institute for  
Health Research***



# A mortality risk model to adjust for case mix in UK paediatric cardiac surgery

C Pagel,<sup>1</sup> KL Brown,<sup>2\*</sup> S Crowe,<sup>1</sup> M Utley,<sup>1</sup>  
D Cunningham<sup>3</sup> and VT Tsang<sup>2</sup>

<sup>1</sup>Clinical Operational Research Unit, University College London, London, UK

<sup>2</sup>Cardiac Unit, Great Ormond Street Hospital, London, UK

<sup>3</sup>National Institute for Cardiovascular Outcomes Research (NICOR), Centre for Cardiovascular Prevention and Outcomes, University College London, London, UK

\*Corresponding author

**Declared competing interests of authors:** The authors declare (1) no financial support for the submitted work from anyone other than their employer and NIHR as listed above; (2) no financial relationships with commercial entities that might have an interest in the submitted work; (3) no spouses, partners or children with relationships with commercial entities that might have an interest in the submitted work; and (4) no non-financial interests that may be relevant to the submitted work.

Published May 2013

DOI: 10.3310/hsdr01010

This report should be referenced as follows:

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).



# Health Services and Delivery Research

ISSN 2050-4349 (Print)

ISSN 2050-4357 (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) ([www.publicationethics.org/](http://www.publicationethics.org/)).

Editorial contact: [nihredit@southampton.ac.uk](mailto:nihredit@southampton.ac.uk)

The full HS&DR archive is freely available to view online at [www.journalslibrary.nihr.ac.uk/hsdr](http://www.journalslibrary.nihr.ac.uk/hsdr). Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: [www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

## Criteria for inclusion in the *Health Services and Delivery Research* journal

Reports are published in *Health Services and Delivery Research* (HS&DR) if (1) they have resulted from work for the HS&DR programme or programmes which preceded the HS&DR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

## HS&DR programme

The Health Services and Delivery Research (HS&DR) programme, part of the National Institute for Health Research (NIHR), was established to fund a broad range of research. It combines the strengths and contributions of two previous NIHR research programmes: the Health Services Research (HSR) programme and the Service Delivery and Organisation (SDO) programme, which were merged in January 2012.

The HS&DR programme aims to produce rigorous and relevant evidence on the quality, access and organisation of health services including costs and outcomes, as well as research on implementation. The programme will enhance the strategic focus on research that matters to the NHS and is keen to support ambitious evaluative research to improve health services.

For more information about the HS&DR programme please visit the website: [www.netscc.ac.uk/hsdr/](http://www.netscc.ac.uk/hsdr/)

## This report

The research reported in this issue of the journal was funded by the HS&DR programme or one of its proceeding programmes as project number 09/2001/13. The contractual start date was in July 2010. The final report began editorial review in October 2011 and was accepted for publication in June 2012. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HS&DR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Pagel *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library ([www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)), produced by Prepress Projects Ltd, Perth, Scotland ([www.prepress-projects.co.uk](http://www.prepress-projects.co.uk)).

## **Health Services and Delivery Research Editor-in-Chief**

**Professor Ray Fitzpatrick** Professor of Public Health and Primary Care, University of Oxford, UK

## **NIHR Journals Library Editor-in-Chief**

**Professor Tom Walley** Director, NIHR Evaluation, Trials and Studies and Director of the HTA Programme, UK

## **NIHR Journals Library Editors**

**Professor Ken Stein** Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

**Professor Andree Le May** Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

**Dr Martin Ashton-Key** Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

**Professor Matthias Beck** Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

**Professor Aileen Clarke** Professor of Health Sciences, Warwick Medical School, University of Warwick, UK

**Dr Tessa Crilly** Director, Crystal Blue Consulting Ltd, UK

**Dr Peter Davidson** Director of NETSCC, HTA, UK

**Ms Tara Lamont** Scientific Advisor, NETSCC, UK

**Dr Tom Marshall** Reader in Primary Care, School of Health and Population Sciences, University of Birmingham, UK

**Professor William McGuire** Professor of Child Health, Hull York Medical School, University of York, UK

**Professor Geoffrey Meads** Honorary Professor, Business School, Winchester University and Medical School, University of Warwick, UK

**Professor Jane Norman** Professor of Maternal and Fetal Health, University of Edinburgh, UK

**Professor John Powell** Senior Clinical Researcher, Department of Primary Care, University of Oxford, UK

**Professor James Raftery** Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

**Dr Rob Riemsma** Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

**Professor Helen Roberts** Professorial Research Associate, University College London, UK

**Professor Helen Snooks** Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Please visit the website for a list of members of the NIHR Journals Library Board:  
[www.journalslibrary.nihr.ac.uk/about/editors](http://www.journalslibrary.nihr.ac.uk/about/editors)

**Editorial contact:** [nihredit@southampton.ac.uk](mailto:nihredit@southampton.ac.uk)

# Abstract

## A mortality risk model to adjust for case mix in UK paediatric cardiac surgery

C Pagel,<sup>1</sup> KL Brown,<sup>2\*</sup> S Crowe,<sup>1</sup> M Utley,<sup>1</sup> D Cunningham<sup>3</sup>  
and VT Tsang<sup>2</sup>

<sup>1</sup>Clinical Operational Research Unit, University College London, London, UK

<sup>2</sup>Cardiac Unit, Great Ormond Street Hospital, London, UK

<sup>3</sup>National Institute for Cardiovascular Outcomes Research (NICOR), Centre for Cardiovascular Prevention and Outcomes, University College London, London, UK

\*Corresponding author

**Background:** Congenital heart disease (CHD) is a relatively common disorder in childhood, affecting approximately 8–9 per 1000 live-born infants annually in the UK. CHD often involves serious abnormalities and is an important cause of childhood mortality, morbidity and disability. It is generally recognised that it is important and valuable to monitor outcomes in cardiac surgery and that, to do so fairly and effectively, one needs to risk stratify the case load of each unit. There is evidence that, since outcome monitoring in adult cardiac surgery became mandatory and routine, outcomes have improved. At present, no process for routinely monitoring risk-adjusted outcomes in paediatric cardiac surgery exists.

**Objectives:** To establish whether or not a risk model can be developed that is fit for the purpose of adjusting for case mix severity to facilitate routine monitoring of outcomes for paediatric cardiac surgery in the UK and to assess whether or not and how diagnostic information can augment procedural information in risk adjustment.

**Methods:** Data from the Central Cardiac Audit Database (CCAD) for all cardiac surgery procedures, excluding reoperations within 30 days, performed in the UK for patients < 16 years between 2000 and 2010 (38,597 patient episodes) were included: 70% for model development and 30% quarantined for validation. The outcome was 30-day survival, as supplied to CCAD through the Central Register of NHS patients (now the Medical Research Information Service). The CCAD defines 36 'specific procedures'. Nine of these were merged as a 'low-volume specific procedure' group (< 90 cases each in the entire development set). Unassigned cases were grouped as 'not a specific procedure'. Twenty-four 'primary' cardiac diagnoses and separately a categorisation of 'univentricular' status were defined using a hierarchical algorithm developed by the study team based on International Paediatric and Congenital Cardiac codes. Comorbidities considered included prematurity (< 37 weeks' gestation), Down syndrome, constellations of features that constitute a recognised syndrome, congenital structural defects of organs other than the heart and acquired conditions. Other candidate variables included use of bypass and patient age, weight and sex. Data were analysed using logistic regression.

**Results:** In the development set, there were 25,665 episodes that resulted in survival to 30 days, 693 episodes for which the vital status at 30 days was unknown and 854 episodes that resulted in death within 30 days in the development set (mortality 3.2% overall). The risk model developed includes the following factors: specific procedure, primary cardiac diagnosis grouped into low-, medium- and high-risk categories, univentricular heart status, age band (neonate, infant, child), continuous age, continuous weight, presence of a comorbidity other than Down syndrome and use of bypass. To account for decreasing mortality over time in the development set, a binary indicator for operations performed after 1 January 2007 is

also included in the model. We were able to calculate a risk score for 95% of cases in the test set: weight was missing in 5% of cases. Data completeness improved over time. The proposed model discriminated well: the area under the receiver operating characteristic curve (AUC) for the test set was 0.77 (0.81 for post-2007 data). Removal of all but procedural information gave a reduced AUC of 0.72. The model performed well across the spectrum of predicted risk in the entire data set, but there was underestimation of mortality risk in the test set among neonates operated from 2007.

**Limitations:** An important limitation is that the model pertains to short-term 30-day outcomes (not long-term outcomes) and is designed for the purpose of routine monitoring for quality assurance rather than bedside-type predictions for individual patients. Over the recent period in the validation set (since 2007), the model was found to underestimate risk at the very high-risk end (> 10% risk), in particular among neonates. This indicates that risk adjustment based on the current parameterisation of the model will potentially give an unduly negative impression of outcomes at those centres with a high proportion of high-risk cases. Finally, any risk model used for ongoing quality improvement initiatives needs to be regularly updated as data quality improves and clinical practice evolves.

**Conclusions:** For the first time diagnostic information has been successfully incorporated into risk adjustment for short-term outcomes in this patient group, which added discriminatory power. The risk model is fit for purpose, although the underestimation of risk in recent neonates is an important caveat. Several centres have expressed an interest in piloting the risk model and the accompanying monitoring tool. Future work includes developing software to generate variable life-adjusted display charts within units using the risk model; using the risk model to explore trends in case mix over time; and informing future work in evaluating long-term outcomes for children with CHD.

**Funding:** The National Institute for Health Research Health Services and Delivery Research programme.



# Contents

<b>Glossary</b>	<b>ix</b>
<b>List of abbreviations</b>	<b>xi</b>
<b>Scientific summary</b>	<b>xiii</b>
<b>Chapter 1 Background and our research objectives</b>	<b>1</b>
Congenital heart disease	1
The <i>Safe and Sustainable</i> review	1
Monitoring risk-adjusted outcomes in cardiac surgery	2
Our research objectives	2
<b>Chapter 2 Brief summary of the relevant literature</b>	<b>5</b>
Subjective risk stratification schemes	5
Empirical risk stratification schemes	5
Criteria for the assessment of risk models	6
<b>Chapter 3 The Central Cardiac Audit Database congenital heart disease data set</b>	<b>9</b>
Origins and process	9
Data fields supplied to the research team	9
Specific procedure algorithm	9
Compatibility of the Central Cardiac Audit Database data set and the Risk Adjustment for Congenital Heart Surgery-1 scoring system	11
<b>Chapter 4 Data preparation and descriptive analyses</b>	<b>13</b>
Separation of development and test sets	13
Removal of duplicate records	13
Data cleaning	14
Defining comorbidity	15
Defining episodes of surgical management for analysis	15
Field completeness and descriptive analyses	17
<b>Chapter 5 Defining primary cardiac diagnosis and identifying univentricular heart status</b>	<b>23</b>
Preamble on diagnosis classification	23
Diagnosis classification scheme 1 (mapping International Paediatric Congenital and Cardiac codes to primary diagnosis)	23
Diagnosis classification scheme 2 (mapping International Paediatric Congenital and Cardiac codes to diagnostic groups defined by ventricle number and aortic obstruction)	24
The resultant diagnostic classification schemes	26
Summary of work on diagnosis classification	26

<b>Chapter 6 Univariate associations with mortality</b>	<b>29</b>
Initial exploratory analyses	29
<b>Chapter 7 Model development and selection</b>	<b>35</b>
Model building process	35
Diagnostic information	36
Collapsing comorbidity information	36
Age and weight	37
Low-volume specific procedures	38
Year of surgery	38
Final parameter selection and model performance in the development set	39
<b>Chapter 8 Model performance</b>	<b>41</b>
Global performance in the test set	41
Features of the test set compared with the development set	43
Global performance of model from 2007 onwards	46
Institutional case mix and model performance	47
Proposed monitoring tool	52
<b>Chapter 9 Discussion</b>	<b>55</b>
Model overview and attributes	55
Model limitations and cautions	55
Implications for clinical practice	56
Implications and recommendations for research	58
<b>Chapter 10 Conclusion</b>	<b>61</b>
<b>Acknowledgements</b>	<b>63</b>
<b>References</b>	<b>65</b>
<b>Appendix 1 Dealing with duplicate data</b>	<b>69</b>
<b>Appendix 2 Moving from International Paediatric Congenital and Cardiac codes to diagnostic groups</b>	<b>71</b>
<b>Appendix 3 Comparison of diagnostic groups and specific procedure</b>	<b>119</b>
<b>Appendix 4 Final logistic regression risk model</b>	<b>121</b>
<b>Appendix 5 Protocol</b>	<b>123</b>

# Glossary

**Aristotle** A consensus-based cardiac surgery risk-adjustment scheme.

**Atrial septal defect** Type of heart disease.

**Aortic valve replacement** Type of heart operation.

**Atrioventricular septal defect** Type of heart disease.

**Biventricular heart** Two pumping chambers (normal number) rather than one.

**Classification and regression tree analysis** Used in data analysis.

**Double-outlet right ventricle** Type of heart disease.

**Hypoplastic left heart syndrome** Type of heart disease.

**Intact ventricular septum** A component of the heart.

**Pulmonary atresia** Type of heart disease.

**Paediatric cerebral performance category** Basic evaluation and grading of the child's neurological condition.

**Patent ductus arteriosus** Type of heart disease.

**RACHS-1 (Risk Adjustment for Congenital Heart Surgery-1)** A consensus-based cardiac surgery risk-adjustment scheme.

**Transposition of the great arteries** Type of heart disease.

**Univentricular heart** One pumping chamber (abnormal number) rather than two.

**Ventricular septal defect** Type of heart disease.

**z-score** Standardised or normal scores, such as weight values, in a population.

+ Denotes the presence of left ventricular or aortic obstruction in the heart (abnormal), for example UVH(+).

– Denotes the absence of left ventricular or aortic obstruction in the heart (normal), for example BVH(–).



## List of abbreviations

AUC	area under the receiver operating characteristic curve	MADCAP	mean adjusted deaths compared against predicted
AVSD	atrioventricular septal defect	NICOR	National Institute for Cardiovascular Outcomes Research
BVH	biventricular heart	PDA	patent ductus arteriosus
CCAD	Central Cardiac Audit Database	RACHS-1	Risk Adjustment for Congenital Heart Surgery-1
CHD	congenital heart disease	STS	Society of Thoracic Surgeons
CI	confidence interval	TGA	transposition of the great arteries
EACTS	European Association for Cardio-Thoracic Surgery	UVH	univentricular heart
HLHS	hypoplastic left heart syndrome	VSD	ventricular septal defect
IPCC	International Paediatric Congenital and Cardiac	VLAD	variable life-adjusted display

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.



# Scientific summary

## Background

Congenital heart disease (CHD) is a relatively common disorder in childhood, affecting approximately 8–9 per 1000 live-born infants annually in the UK. The crude prevalence of CHD was estimated in 2008 to be 3.05 per 1000 patients registered with UK general practitioners. CHD often involves serious abnormalities and is an important cause of childhood mortality, morbidity and disability. Around one-third of deaths due to CHD occur before 14 years of age. Infancy is the highest risk period, with 21% of CHD deaths occurring in the first year of life, but the proportion of deaths in adults > 45 years is increasing. In 2006, CHD accounted for 3% of child deaths.

It is generally recognised that it is important and valuable to monitor outcomes in cardiac surgery and that, to do so fairly and effectively, one needs to risk stratify the case load of each unit. Risk stratification of adult cardiac surgery patients is an essential part of the audit process, which reduces the prospect of unfair assessment of outcomes attributed to a surgeon or team whose mortality rate is relatively high simply because it reflects patients who were inherently higher risk. There is evidence that, since outcome monitoring in adult cardiac surgery became mandatory and routine, outcomes have improved, and there has been no consequent negative effect in terms of centres turning away high-risk cases, as was originally feared. There were consequent benefits to patients and their families in terms of the quality of care and the improved information they received. Analytical methods for outcome monitoring are well advanced for adult cardiac surgery and the use of graphical techniques such as the variable life-adjusted display (known by the acronym VLAD) to display risk-adjusted outcome charts is now common in the quality assurance process.

At present, no process for routinely monitoring risk-adjusted outcomes in paediatric cardiac surgery exists, and achieving this is clearly desirable. The challenge is due to the great diversity of the patient population in terms of the diagnoses, indications for surgery, operations performed, age at operation and other factors as well as the logistics of co-ordinating such an endeavour across many disparate cardiac centres in a geographical area.

## Objectives

Our objectives for this project were

1. to test an existing risk model based on the Risk Adjustment for Congenital Heart Surgery-1 (RACHS-1) score and patient age, derived from outcomes at one centre (Great Ormond Street Hospital), in the Central Cardiac Audit Database (CCAD) data from all centres across the UK
2. to understand the contribution that diagnostic information can make to risk estimation and monitoring of outcomes, establish whether or not information concerning comorbidities can contribute to improved methods of risk estimation and, if indicated and possible, revise the existing risk model such that it is suitable for use at other centres and by CCAD
3. to examine the implications of reporting mortality outcomes by diagnosis as well as by procedure category
4. to disseminate our findings and any risk models and monitoring tools developed to UK centres and to CCAD, so that it can consider how best to share the information with stakeholders, including through its 'public portal' web pages.

## Methods

### Data

Since 2000, quality assurance in paediatric cardiac surgery in the UK has been underpinned by CCAD, which provides aggregated national and institution-specific data. Mandatory data submissions to CCAD are requested every 3 months from hospitals performing cardiac surgery in the UK, including details about patients and the operations performed. Patient identifiers, including NHS number, are provided to CCAD, which periodically requests information on patients' survival status from the Central Register of NHS patients. This process is approved by the National Information Governance Board for Health and Social Care with consent requested from parents for participation in national audit of outcomes. There is an extensive quality assurance process incorporating a rolling programme of professionally led site visits.

We received a data set of pseudonymised records from CCAD in December 2010. After removing all records in which the patient was > 16 years at the time of the procedure and all records for patients who underwent only catheter procedures, we split the data set into development (70% of patients) and test (30% of patients) data sets using random allocation stratified by year and institution of first procedure for a patient. All further analysis was performed on the development set, which contained 34,385 records, corresponding to 22,449 unique patients. The quarantined test set contained 14,316 records, corresponding to 9354 unique patients, and played no part in risk model development.

### Data cleaning

Considerable effort was required to prepare the data set for analysis. This data cleaning process involved identifying and removing duplicate records, identifying and resolving instances in which the pseudonymisation process had resulted in different patients being given the same ID number, and identifying and resolving instances of inconsistent patient details within a sequence of records. The inclusion of the CCAD senior strategist, Dr David Cunningham, on the project team was essential to resolving issues of data quality and interpretation in a timely manner.

### Outcome measure

The unit of analysis was a '30-day episode'. 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 in the model development. The next surgical procedure recorded for the same patient > 30 days after the first surgical procedure was treated as the start of a new 30-day episode.

The outcome measure for each episode was death within 30 days of the start of that episode.

### Candidate risk factors

For each episode there were data fields available providing the specific procedure performed (one of 36 groups defined by CCAD or 'unassigned'), whether or not the operation was performed on bypass, up to six separate diagnostic codes, data on comorbidity, whether or not the patient had an antenatal diagnosis, patient sex, age and weight, ethnicity and a measure of deprivation. Comorbidities considered included prematurity (< 37 weeks' gestation), Down syndrome, genetic syndromes and clinical constellations of features that constitute a recognised syndrome, congenital structural defects of organs other than the heart and acquired conditions.

### Preparation of diagnostic information

There were up to six raw diagnostic codes for each patient record and each raw diagnostic code can take on one of several hundred values. The research team enlisted the support of other clinical experts in developing a new hierarchy of 24 primary cardiac diagnoses based on the raw diagnostic codes available for each episode.



Additionally, we identified within the raw codes those combinations which indicated that the patient had a single functioning ventricle.

### Analysis

Candidate risk factors were first assessed in terms of data quality, clinical face validity and univariate association with death within 30 days. Classification and regression tree analysis was used to identify groupings of primary cardiac diagnoses that were associated with death within 30 days.

Model development followed an iterative process involving multiple logistic regression, assessment of model performance, assessment of model stability, clinical discussion and variable simplification. Ultimately, model choice was not a purely statistical consideration. Consideration of uptake by CCAD and centres supported the use of as many specific procedure groups as possible and we also took account of the opportunity presented by the risk model to drive improvements in completeness and data quality concerning diagnosis and comorbidity.

## Results

### Mortality rates

In the development set there were 25,665 episodes that resulted in survival to 30 days, 693 episodes in which the vital status at 30 days was unknown and 854 episodes that resulted in death within 30 days (mortality 3.2% overall).

### The risk model

The final risk model, decided on jointly by the clinical and analytical teams, was a logistic regression model with the following variables:

- age (both as a continuous measure and as neonate/infant/child bands)
- weight (as a continuous measure)
- specific procedure (one of 27 CCAD groups, 'no specific procedure' or 'low-volume specific procedure')
- procedure type (bypass or non-bypass)
- broad diagnosis group (low-, medium- or high-risk group)
- univentricular heart attribute (indicator variable)
- presence/absence of a recorded non-Down syndrome comorbidity
- indicator variable for whether an episode occurred pre 2007 or from 2007 onwards.

This last variable was introduced to account for decreasing mortality over time in the development set and to ensure that prospective use of the risk model would provide a means of benchmarking against recent national outcomes.

### Evaluation of the risk model

In the test set it was possible to calculate a risk score in 95% of cases. Weight was missing in 4.9% of cases and age in 0.2% of cases. There were also 392 (3.4%) episodes without diagnostic information but these were included in the analysis within the diagnostic category 'empty/missing diagnosis'. Additionally, the 30-day life status was 'unknown' for 226 (2.0%) episodes in the test set. There were differences in the mortality rates associated with individual risk factors between the test set and the development set. In particular, there was a higher mortality rate among neonates in the test set than among those in the development set.

The area under the receiver operating curve (AUC) for the model in the test set across all years is 0.77, similar to the value calculated for the development set (0.78). This indicates good discrimination between groups of patients with high and low mortality. The total number of observed deaths was 335 compared with 329 predicted. There were statistically significant discrepancies between observed and predicted mortality in deciles of predicted risk.

Given the intention for this model to be used prospectively, we also evaluated the model in test episodes since 2007. Here, there was greater discrimination (0.81 vs 0.78), but the model underestimates risk among very high-risk episodes and is, as a result, less accurate overall in these more recent data (128 observed deaths vs 113 predicted post 2007).

### **Case mix**

The distribution of predicted risk in the two data sets is very similar and shows that just over 30% of episodes have a predicted risk of 30-day mortality of  $\leq 1\%$  and 80% have a predicted risk of  $\leq 4\%$ , but that around 5% of episodes have a predicted risk of  $> 10\%$ .

Many of the institutions have quite similar profiles of predicted risk but there are some marked differences. For instance, at one institution 12% of episodes have a predicted risk of 30-day mortality of  $> 10\%$ , whereas at another institution this proportion is  $< 1\%$ .

## **Conclusions**

A risk model has been developed for monitoring short-term surgical outcomes following paediatric cardiac surgery that, for the first time, makes use of diagnostic information as well as procedural data. The model shows good discrimination between groups of patients with high and low mortality and reasonable accuracy.

Given that the model underestimated risk at the very high-risk end of the recent (since 2007) test data, risk adjustment using the model as currently parameterised will potentially give an unfair assessment of outcomes at those centres with a high proportion of high-risk cases. This is an important caveat to the interpretation of risk-adjusted outcomes within and across institutions that will need to be discussed with CCAD and the institutions as this work is taken forward.

## **Recommendations for future research**

The completion of this project has highlighted a number of research priorities and opened possibilities for future research:

1. It would be desirable for any implementation of routine monitoring of risk-adjusted outcomes in paediatric cardiac surgery to be piloted and, if appropriate, rolled out and evaluated, with evaluation including qualitative research on drivers and barriers to adoption.
2. The availability of a risk model makes possible research into differences between case mix across centres nationally and into trends in case mix over time.
3. The scheme developed as part of this project for the classification of primary cardiac diagnosis based on the codes available within the CCAD data set enables important research to be conducted on the long-term outcomes achieved for groups of surgical CHD patients defined by diagnosis. This would provide valuable information to patients, carers, commissioners and clinicians alike.

## Our next steps

One of the co-investigators has, over the course of the project, joined the CCAD steering committee. This has further improved the link between CCAD and the project team, which was already robust given the presence of the senior strategist for the UK cardiac audits as a co-investigator.

This project and future use of the risk model was discussed at the CCAD steering committee meeting in October 2011. The CCAD steering committee has a strong interest in the work given their commitment to audit and quality assurance. CCAD and several centres have expressed an interest in piloting the risk model and the accompanying monitoring tool.

## Funding

The National Institute for Health Research Health Services and Delivery Research programme.



# Chapter 1 Background and our research objectives

## Congenital heart disease

Congenital heart disease (CHD) is a relatively common disorder in childhood, affecting approximately 8–9 per 1000 live-born infants annually in the UK.<sup>1</sup> The crude prevalence of CHD was estimated in 2008 to be 3.05 per 1000 patients registered with UK general practitioners.<sup>2</sup> CHD often involves serious abnormalities and is an important cause of childhood mortality, morbidity and disability.<sup>3,4</sup> Around one-third of deaths due to CHD occur before 14 years of age. Infancy is the highest risk period, with 21% of CHD deaths occurring in the first year of life,<sup>5,6</sup> but the proportion of deaths in adults > 45 years is increasing.<sup>7</sup> In 2006, CHD accounted for 3% of child deaths.<sup>5</sup>

Currently, around 9000 procedures (including both cardiac surgery and interventional cardiac catheterisation) are performed for CHD in the UK each year, the majority of them in children.<sup>8</sup> Recent studies document greater numbers of children surviving to lead adult lives with repaired or palliated CHD, who experienced an increasing level of hospital admissions between 1994 and 2007<sup>7</sup> and used more primary care facilities than matched control subjects.<sup>2</sup> Reoperations for CHD, long-term survival for individual cardiac diagnoses and quality of life are viewed as important issues by parents and professionals.<sup>9</sup>

## The *Safe and Sustainable* review

This work has been conducted against the backdrop of a major review of paediatric cardiac surgery services in the UK. At the time of writing there are 14 specialist paediatric cardiac surgery centres in the UK, which perform a variable total number of procedures and different individual procedures.<sup>8</sup> The *Safe and Sustainable* review of paediatric cardiac surgery<sup>10</sup> set out to define criteria that characterise a high-quality paediatric cardiac surgery service, review existing units against those criteria and make recommendations concerning the future delivery of the service in England and Wales relating to the current 11 NHS centres in that region. Following a consultation process, its recommendations include a reduction in the number of units providing this care. For some, the process adopted by the *Safe and Sustainable* team presents a blueprint for how other specialist services should be assessed and, if necessary, reconfigured. For others, the review has meant further discord, uncertainty and disruption for a specialty that has, arguably, been the subject of more external scrutiny than any other.<sup>11–16</sup>

The *Safe and Sustainable* review documents published in 2010<sup>10</sup> ranked centres according to health-care delivery, organisation, staff and research metrics including volume of cases and quality of post discharge care. Outcome measures were not included in the review data because the advisory panel concluded that validated methods for adjustment of case complexity were not available, and therefore the raw outcome data available from the Central Cardiac Audit Database (CCAD), albeit complete and validated, were too challenging to interpret for inclusion.

On the basis of the data obtained in the review process, the Joint Commission of Primary Care Trusts proposed four options for service reconfiguration, with closure of four or five surgical centres, leaving six or seven larger ones, on the grounds that this change would make the service safer and more sustainable.

Among its recommendations, the *Safe and Sustainable* review board concluded that CCAD should 'review its process for reporting outcomes such that these were timely and meaningful'.<sup>10</sup> Our project is very timely with respect to this recommendation and, based on the interest shown by surgical units in our findings (see *Chapter 9*), the review would seem to have contributed to a receptive context for this piece of research.

## Monitoring risk-adjusted outcomes in cardiac surgery

In the wake of the Bristol Royal Infirmary Inquiry, which followed one centre experiencing an abnormally high death rate in children after cardiac surgery,<sup>17</sup> there has been an emerging culture of audit and quality improvement in the NHS and particular interest in monitoring outcomes and centre performance within paediatric cardiac surgery.<sup>12,13</sup>

It is generally recognised that it is important and valuable to monitor outcomes in cardiac surgery and that to do so fairly and effectively one needs to risk stratify the case load of each unit.<sup>18</sup> Risk stratification of adult cardiac surgery patients is an essential part of the audit process, which reduces the prospect of unfair assessment of outcomes attributed to a surgeon or team whose mortality rate is relatively high simply because it reflects patients who were inherently higher risk. There is evidence that, since outcome monitoring in adult cardiac surgery became mandatory and routine, outcomes have improved, and there has been no consequent negative effect in terms of centres turning away high-risk cases, as was originally feared.<sup>19</sup> There were consequent benefits to patients and their families in terms of the quality of care and the improved information they received. Analytical methods for monitoring outcomes are well advanced for adult cardiac surgery and the use of graphical techniques such as the variable life-adjusted display (commonly known by the acronym VLAD) to display risk-adjusted outcome charts is now a common part of the quality assurance process.<sup>20–22</sup>

### Our research objectives

At present, no process for routinely monitoring risk-adjusted outcomes in paediatric cardiac surgery exists. Achieving this for paediatric congenital heart surgery is clearly desirable. The challenge is due to the great diversity of the patient population in terms of the diagnoses, indications for surgery, operations performed, age at operation and other factors<sup>23</sup> as well as the logistics of co-ordinating such an endeavour across many disparate cardiac centres in a geographical area.

Importantly, as outlined in *Chapter 2*, the literature and our own previous work focus on outcomes and risks of paediatric surgery according to the procedure(s) that patients have undergone, augmented by patient-specific information such as age and weight. Up to now, no use has been made of information concerning the nature of the heart defect or the diagnosis. There are some recent examples showing that consideration of the comorbidities that the patient has may augment information about procedural risk, and this is also an important factor to consider in respect of case complexity.<sup>24</sup> Although in many surgical specialties there is a one-to-one mapping between diagnosis and surgical intervention, this is not the case in paediatric congenital heart surgery. Some procedures are performed for a number of diagnoses, which undermines the extent to which the procedures performed by a surgeon or within a unit accurately reflect case mix. Similarly, the same heart defect may be managed using different surgical interventions, with the choice of intervention reflecting other aspects of the patient's condition (age, weight, comorbidities and severity of symptoms), but also with scope for there to be differences in surgical strategy between units based on local experience and, potentially, on a different balance being struck between long-term objectives and short-term risks. For these reasons, we were keen to explore, in our project, the scope for using diagnostic information in risk adjustment.

Our objectives for this project were:

1. to test an existing risk model based on the Risk Adjustment for Congenital Heart Surgery-1 (RACHS-1) score and patient age, derived from outcomes at one centre (Great Ormond Street Hospital, London, UK), in the CCAD data from all centres across the UK
2. to understand the contribution that diagnostic information can make to risk estimation and monitoring of outcomes, establish whether or not information concerning comorbidities can contribute to

improved methods of risk estimation and, if indicated and possible, revise the existing risk model such that it is suitable for use at other centres and by CCAD

3. to examine the implications of reporting mortality outcomes by diagnosis as well as by procedure category
4. to disseminate our findings and any risk models and monitoring tools developed to UK centres and to CCAD, so that it can consider how best to share the information with stakeholders, including through its 'public portal' web pages.





## Chapter 2 Brief summary of the relevant literature

There has been a considerable amount of research effort put into the development of risk stratification schemes and risk scores for paediatric cardiac surgery. Some of this work is based on consensus opinion among cardiac surgeons and other professional groups and other work is based on the analysis of empirical data. The major contributions to the literature are discussed in the following sections.

### Subjective risk stratification schemes

With respect to risk-stratifying patients according to the operation performed, a classification scheme called RACHS-1<sup>25</sup> was developed by a committee of experts and published in 2002, by which 79 different types of paediatric cardiac operation are grouped into six categories ranked in order of increasing risk, as perceived by clinicians. The RACHS-1 scheme appears to be useful as a basis for forecasting risk,<sup>23</sup> has been validated in a range of contexts and, as of October 2011, has been cited in 337 scholarly articles. One limitation of this scheme is that not all of the great range of operations described for CHD are incorporated and, therefore, a significant proportion of procedures do not have a risk category attributed. The scheme considers operative information only, with some minimal information on age ranges for a small number of procedures. A further limitation is that this scheme was developed before large amounts of registry data were available and, therefore, was based on clinician expert opinion rather than empirical information.

A further consensus-based scoring system developed by clinicians to describe perceived surgical risk is known as Aristotle.<sup>26</sup> The Aristotle system is based on three components: perceived risk of mortality, morbidity and perception of technical difficulty. Estimates of these three factors were made by a panel of experts for 145 different procedures. There are two versions of the Aristotle score, the more detailed of which requires collection of 248 variables: this high level of data collection is perceived as a barrier to the score being implemented. The Aristotle system has been widely used for grading the complexity of operations for the purposes of audit, but has been less widely taken up than RACHS-1.

Our research team has previously reported the development of a model for the risk of perioperative mortality using data pertaining to 1083 congenital open-heart operations performed between April 2000 and March 2003 at one institution: Great Ormond Street Hospital, London, UK.<sup>23</sup> This model incorporated only information about procedure type, classified according to the RACHS-1 scheme,<sup>25</sup> and patient age. We reported the use of this model for monitoring outcomes for cases performed between April 2003 and May 2004<sup>27</sup> and used these analyses as the basis for a comparison between the RACHS-1 and Aristotle complexity schemes.<sup>28</sup>

### Empirical risk stratification schemes

Since the start of the millennium, there has been a worldwide focus and effort in the field of paediatric cardiac surgery (as for other specialties) to collect audit data for the purposes of quality assurance and benchmarking.<sup>29–31</sup> This era has seen the evolution of CCAD in the UK, the Society of Thoracic Surgeons (STS) Congenital Heart Surgery Database in North America and the European Association for Cardio-Thoracic Surgery (EACTS) Congenital Heart Surgery Database, which accepts data submissions from various countries in Europe and even a few from North Africa. These are multi-institutional databases serving geographical regions and, as such, must conform to the laws and culture of the regions they serve. As mentioned earlier, of all of these databases CCAD is unique because data submission is mandatory and universal within a clearly defined region, data are validated at each centre annually and outcome is independently verified.

One key factor that has allowed for the evolution of these databases is the underpinning work performed on congenital cardiac diagnostic and procedural coding and the development of universally applicable codes

to describe congenital cardiac case mix,<sup>32–35</sup> which has been relatively recent and indeed is an ongoing endeavour. Accrual of standardised data in these large multi-institutional databases has led to a shift from the use of consensus-based risk stratification tools to risk stratification based on the available empirical data.

In 2009 an empirical risk-adjustment tool was published using information from the large STS (43,934 episodes) and EACTS (33,360 episodes) databases.<sup>24</sup> Procedures were assigned a numerical score, calculated using a Bayesian model, and performance of this model was validated in a test set. One of the aims of this study was to group procedures with similar estimated mortality risks into relatively homogeneous groups, which could then in turn be used to adjust for case mix when analysing outcomes and benchmarking. The risk model performed well and compared favourably with previous schemes used for risk stratification based on consensus such as RACHS-1 and Aristotle.

The outcome measure or end point for the model was hospital mortality, because consistent verification of outcome outside the hospital environment was not available to these databases. One problem with this approach is that patients may die shortly after discharge home and such deaths may be related to the operation.<sup>16</sup> A further problem is that patients have extremely variable lengths of hospital stay, with outlying patients staying weeks or months having a greater mortality risk than those with shorter lengths of stay.<sup>36</sup> Use of hospital discharge (or eventual death) as the end point for such analyses makes it very challenging to include such long-stay patients, who are higher risk, in a VLAD-type contemporaneous analysis of risk-adjusted outcomes. A final problem with this approach is that patients may be discharged from one institution to another, sometimes still on life support, some of whom may die at the second institution; these outcomes would be excluded from the analysis.

A further consideration, which was alluded to by the authors, is that participation in the databases is voluntary, with not all institutions in the same regions taking part. The participating centres are subject to a variable level of validation, with some centres undergoing thorough validation and others undergoing none. This is in part due to the legal obstacles associated with patient confidentiality, which are a barrier to independent verification of data from an outside body. This issue has been averted in the CCAD data that have been used for our study because parent consent for participation in national audit is sought by the centres, which allows universal submission of data and validation to occur.

The coverage of procedure types included a higher proportion in the 2009 study than in the previous consensus-based RACHS-1 score, which did not incorporate all procedure types. However, this newer empirical risk-adjustment scheme did not consider diagnostic information and noted that the approach to augment procedure-specific information with information about other patient-specific factors, such as comorbidities, was as yet undetermined.

## Criteria for the assessment of risk models

In the literature described in the previous sections, the main considerations in judging or comparing the attributes of different risk stratification schemes and risk models are the coverage, that is, the proportion of cases that can be incorporated into the scheme or model, and the discrimination, that is, the extent to which the scheme or model is successful in distinguishing groups of patients with higher and lower postoperative mortality. The standard measure of discrimination adopted in this literature and in this project is the area under the receiver operating characteristic curve (AUC). This measure gives the probability that a patient who died, chosen at random, would have had a higher risk score than a randomly selected survivor.

For models that estimate a percentage risk of postoperative death, another useful measure of model performance is the Hosmer–Lemeshow chi-squared statistic. This gives an indication of how statistically significant any deviations are between predicted and observed mortality rates in groups of patient defined by deciles of predicted risk.

In addition to these measures, we have used the graphical tool MADCAP (mean adjusted deaths compared against predicted chart) to assess the discrimination and accuracy of the candidate models developed in this project. MADCAP enables visual identification of patterns of systematic over- or underestimation of risk on the part of a risk model.<sup>37</sup>



# Chapter 3 The Central Cardiac Audit Database congenital heart disease data set

## Origins and process

Since 2000, quality assurance in paediatric cardiac surgery in the UK has been underpinned by CCAD;<sup>8</sup> this provides aggregated national and institution-specific data. Mandatory data submissions to CCAD are requested every 3 months from hospitals performing cardiac surgery in the UK, including details about the patients and the operations performed. Patient identifiers including NHS number are provided to CCAD, which periodically requests information on patients' survival status from the Central Register of NHS patients. This process is approved by the National Information Governance Board for Health and Social Care, with consent requested from parents for participation in the national audit of outcomes. There is an extensive quality assurance process incorporating a rolling programme of professionally led site visits.

As it has evolved to date, CCAD receives data on about 9000 procedures annually. The 30-day surgical survival rates for 38 clinically recognisable 'specific procedures' are available online via a public portal (see *Specific procedure algorithm* for more details). The data are currently displayed as funnel plots of survival rate against case volume<sup>8</sup> for each of the individual specific procedures for a 3-year period, which permit institutions with different case loads to benchmark their outcomes in relation to others and to the national average.

There is currently no case mix or other risk adjustment used in the presentation of the aggregated data for each institution, but the data set has several features that make it amenable for the development of a risk model. With respect to the data sets used to develop other risk models for this surgery (see *Chapter 2*), the advantages of the CCAD data set are that data submission is mandatory for all UK centres performing this work, the data are subject to a quality assurance process, the survival status of patients is consistently defined and independently established and there is detailed diagnostic and comorbidity information available to potentially augment procedural factors.

## Data fields supplied to the research team

The data fields supplied to us for each record are shown in *Figure 1*.

Appropriately, for reasons of confidentiality, we were not supplied with raw data concerning patient dates of birth and exact dates of procedure, but rather the calculated age at operation (in days) and the month and year of procedure. This information was of equivalent value in the development of the risk model, but the agreed absence of date of birth and date of procedure data did present some challenges when attempting to identify duplicate records and confirm the matching of records pertaining to the same patient (see *Chapter 4, Separation of development and test sets and Removal of duplicate records*).

## Specific procedure algorithm

The concept of benchmark operations has previously been used as an approach to deal with adjustment for case complexity in paediatric cardiac surgery, in the absence of other methods being available. This approach is still valued as a complement to other methods that have been developed. Institutional results for certain more prevalent and recognisable benchmark procedures were published shortly after the Bristol Inquiry in the form of reports (e.g. arterial switch operation, complete repair of tetralogy of Fallot). One of the

Pseudonymised NHS number	Pseudonymised in that the link is retained between records pertaining to same patient and real identities are available to CCAD
Pseudonymised hospital number	
Pseudonymised institution ID	
Age	Calculated by CCAD from date of birth and date of procedure
Weight	
Six separate fields containing diagnostic codes	See Chapter 5, <i>Defining primary cardiac diagnosis and identifying univentricular heart status</i> for more details
AntenatalDx	
Preprocedure seizures	
Preprocedure PCPC	
Comorbidity	
The year and month of procedure	
Procedure type	Bypass or non-bypass
Primary operator status	
Operator 2 status	
Sternotomy sequence	
<del>Bypass time</del>	
<del>Cross-clamp time</del>	
<del>Circ-arrest time</del>	
<del>Cath-proc time</del>	
<del>Cath-fluoro time</del>	
<del>Cath-fluoro dose</del>	
Eight separate fields containing procedure codes	See <i>Specific procedure algorithm</i> for more detail
<del>Post-procedure seizures</del>	
<del>Post-procedure PCPC</del>	
<del>Post-op intubation (days)</del>	
Specific procedure	See <i>Specific procedure algorithm</i> for more detail
<del>Day to discharge</del>	
Discharge status	
<del>Discharge destination</del>	
Days-to-life status	
Lifestatus	
Status 30 days	
<del>Status 1 year</del>	
Reoperated within 30 days surgical	
Reoperated within 30 days catheter	
Reoperated within 30 days any	
Time to reoperation surgical	
Time to reoperation catheter	
Sex	
Ethnicity	
<del>Admin status</del>	
PCT	
Deprivation	

**FIGURE 1** The data fields supplied by CCAD. Peri- and postoperative variables were not considered in the model (these are struck through).

reasons is the stated great diversity and complexity of CHD, which may lead to variations of the same procedure being perceived as differing in complexity and therefore risk. It was viewed as crucial to build consensus within the professional community around the classification of individual operations. A significant amount of effort went into defining each one so that comparisons of performance would both be fair and be seen to be fair.

As CCAD was being set up by the professional organisations involved in the care of children with cardiac disease in the late 1990s, the approach used for comparison of institutional performance was through a core group of benchmark operations. Over the first few years of CCAD data collection, the steering committee of CCAD, which consists of a mixture of experienced paediatric cardiac surgeons and paediatric cardiologists, worked on extending the reach of the audit in terms of presentation of results from an increasingly large pool of recognisable operations. This process necessitated the development of the specific procedure algorithm to ensure consistency and transparency with respect to procedure definitions.

The specific procedure algorithm is a means to link the individual International Paediatric Congenital and Cardiac (IPCC) codes submitted to CCAD as part of a patient's procedure record in order to define a recognisable operation. Several individual IPCC codes (up to eight) may be submitted to describe each operation, which when considered in combination give the most accurate account of what the operation involved.

The algorithm defines the series of IPCC codes that may be included to identify an individual operation, and in a proportion of operations in which the definition may be more complicated also defines the list of IPCC codes that must be excluded. The algorithm contains a hierarchy that ranks the recognisable procedures in order, with the most complex at the top (grouped first) and the least complex at the bottom (grouped last).

The algorithm has been refined and improved by the CCAD steering committee year on year, such that the definitions of each operation are tight and consistent.

The CCAD public portal was launched in 2005 with publication of paediatric cardiac surgery results by recognisable procedures in the form of funnel plots. Following public display of these results, the level of engagement from the UK institutions caring for children with heart disease increased, as professionals from the centres took a great interest in the results as they appeared on the public site. This engagement and comment has led to further refinements and improvements in the specific procedure definitions, which are themselves provided on the website as part of the supporting materials. The algorithm has been periodically improved and updated based on feedback and comment from professionals in the community.

### **Compatibility of the Central Cardiac Audit Database data set and the Risk Adjustment for Congenital Heart Surgery-1 scoring system**

One of our stated aims was to explore the utility of a previously developed risk score based on RACHS-1 and age.<sup>23</sup> Because we did not receive the definitive CCAD data set for this project until December 2010, we undertook this task using an older version of the CCAD data set supplied previously for pilot work.

We wrote an algorithm that could map the procedure codes as they appear in the database to RACHS-1 categories. The CCAD surgical codes are European short codes (referred to as IPCC short codes), of which there are over 1000. These IPCC short codes appear in different combinations in each database record, which, when considered together, provide information about the operation that the patient underwent.

The RACHS-1 groupings of procedures represent lists of clinical descriptions of actual operations, each of which may be described by several different individual IPCC codes or combinations of IPCC codes. Some of these IPCC short codes can correspond to more than one RACHS-1 category and, hence, could not be assigned unambiguously to a RACHS-1 category. It is also a known feature of RACHS-1 that not all surgical procedures are encompassed by the scheme; some unusual or rarer complex operations do not appear at all.

In 64% of CCAD records, all recorded surgical procedures could be classified and given an unambiguous score under the RACHS-1 scheme. Another 27% of records had at least one recorded surgical procedure that was not classifiable with RACHS-1, 7% of records had at least one recorded surgical procedure that could not be assigned to an unambiguous RACHS-1 category and 2% of records had at least one unclassifiable procedure as well as one procedure that could not be assigned to an unambiguous RACHS-1 category.

Given that over one-third of CCAD records could not be assigned to an unambiguous RACHS-1 category, we decided not to pursue using RACHS-1 in an updated risk model. Instead, we chose to adopt the CCAD classification of 'specific procedure' discussed in the previous section.





## Chapter 4 Data preparation and descriptive analyses

### Separation of development and test sets

We received the CCAD data set for use in this project in December 2010. After removing all records in which the patient was > 16 years at the time of the procedure and all records for patients who underwent only catheter procedures, we split the data set into development (70% of patients) and test (30% of patients) data sets using random allocation stratified by year and institution of first procedure for a patient. From previous work<sup>11</sup> we were aware of the possibility of improving outcomes over time, which motivated stratifying the randomisation by year of procedure. We wanted to have as much data as possible for developing a model while maintaining a sufficient sample for meaningful evaluation of model performance. Given the size of the CCAD data set, we felt that a 70/30 split was suitable. All further analysis was performed on the development set, which contained 34,385 records, corresponding to 22,449 unique patients. The quarantined test set contained 14,316 records, corresponding to 9354 unique patients, and played no part in risk model development. It is necessary to observe such a quarantine process to avoid the bias that inevitably results if the same data that are used to develop a risk model are used to test it, which almost always leads to overoptimistic test findings.

Given that we had previously worked with some of these data (see *Chapter 3, Compatibility of the Central Cardiac Audit Database data set and the Risk Adjustment for Congenital Heart Surgery-1 scoring system*), we wanted to ensure that those records previously analysed formed part of the development set. This required us to match cases across two versions of the data based on the patients' pseudonymised NHS numbers, the pseudonymised institution IDs and the patients' pseudonymised hospital numbers. We encountered some problems in doing this, which were resolved through analysis and discussions between the analytical team and David Cunningham at CCAD.

We note that the necessary pseudonymisation process had the negative effect of essentially blinding the analytical team to missing or anomalous data in the patient and hospital ID fields. This meant that we had to conduct extensive face validity exercises when trying to identify genuine groups of records pertaining to the same patient. This included analysis to check internal consistency of patient histories in terms of the dates of operations and ages at each operation and also some manual checks of the clinical face validity of sequences of operations. The close collaboration between clinical and analytical teams was essential here, as was having the involvement of CCAD, which could, when necessary, go back to its own records for clarification.

### Removal of duplicate records

Some records appeared to be duplicates based on systematic comparison of pseudonymised patient NHS number or hospital number, year and month of operation, age, specific procedure, procedure type, procedure 1 (the first procedure code specified), bypass time and catheter procedure time variables. There is a plausible mechanism for record duplication: if an institution tries to update an individual record within CCAD by changing either the patient's hospital ID number or the procedure date (both of which are key fields), a second, duplicate, record will be created instead of the original record being overwritten. Two members of the clinical team (KB and NM) went through the list of possible duplicates identified by the comparison algorithm developed by the analysts and marked those records that were considered to be likely duplicates. We then used a formal protocol (see *Appendix 1*) to determine which record of each pair to

retain, based on completeness and plausibility of the data contained in the records. A total (across the test and development sets) of 244 duplicate records were removed from the analysis.

## Data cleaning

As discussed in *Chapter 3*, the CCAD data set is widely recognised as being of an extremely high quality (arguably the best of its type in the world), but given its sheer size and the complexity of paediatric cardiac diagnoses and surgery it is perhaps inevitable that there would be some errors and anomalies in the data. This is particularly the case for information that has previously not been widely used, such as diagnostic and comorbidity information. The analytical team spent considerable amounts of time understanding the data set and, where necessary, amending it. At key steps in this process David Cunningham, at CCAD, was consulted to confirm the analyst's interpretation of the data. The cleaning processes observed for key fields are set out below. Note that cleaning in the test set was performed after model development was completed.

### Sex

For patients whose sex was recorded differently in different records, we assigned the most frequently occurring sex for that patient to all records for that patient ( $n=140$  patients in the development set;  $n=78$  patients in the test set). In cases in which neither sex occurred more frequently than the other for a given patient, we assigned a sex of 'unclear' to all records for that patient ( $n=100$  patients in the development set;  $n=42$  patients in the test set).

### 30-day life status

We identified those patient records for which missing life status could be inferred from later (or earlier) 30-day life status for the same patient: when an unknown status was followed by a status of 'alive' the earlier status was adjusted to 'alive' ( $n=113$  records in the development set;  $n=50$  records in the test set); when an unknown status was preceded by a status of 'dead' the later status was adjusted to 'dead' ( $n=2$  records in the development set; this did not occur in the test set).

### Age and weight fields

Some of the ages, weights and combinations of age/weight recorded for an episode were considered to be implausible. When an age or weight was recorded as identically zero it was considered to be missing data. The data set was then subdivided into 23 age bands (narrower at young age) and the mean and standard deviation of the recorded weights within each band were calculated. These were then used to calculate the weight-for-age z-scores for each episode within the data set. Episodes in the development set with missing weights ( $n=1327$ ) or a z-score of either  $\leq -3$  or  $\geq 3$  (considered infeasible) ( $n=171$ ) were assigned the mean weight corresponding to their age band. A further 35 episodes within the development set were identified as having an anomalous weight/age combination by a clinical member of the project team (KB) and were assigned the mean weight for age. To mimic prospective use, no adjustment of weights of this nature was made in the test set.

### Diagnosis fields

Our use of diagnostic information in the model development relied on us being able to identify each recorded diagnosis for a record by its CCAD code (which correspond to IPCC short codes). Some of the information recorded in the six possible diagnosis fields for each record in the data set was not automatically identifiable as an official CCAD code. In cases in which this information was ambiguous, we replaced that diagnosis with 'empty/unknown' ( $n=1819$  records in the development set;  $n=671$  records in the test set). In other instances we were able to replace the information unambiguously by an official CCAD code, for example when there was an anomalous numerical code format that was nonetheless identifiable (e.g. additional spaces or dots) ( $n=1746$  records in the development set;  $n=788$  records in the test set), or the numerical component of the code was missing but the text part was complete ( $n=23$  in the development set;  $n=8$  in the test set). When multiple codes were recorded in a single diagnosis field, we split these and

moved them to separate diagnosis fields in sequential order for that record ( $n=39$  in the development set;  $n=9$  in the test set).

### Procedure fields

In general, the quality of the procedural information in the data set was of a very high standard, although there were some examples of information recorded in the procedure fields that was not automatically identifiable as an official CCAD code. When the code had an anomalous format (e.g. additional spaces or dots) ( $n=62$  records in the development set;  $n=21$  records in the test set) or the numerical component of the code was missing but the text part was complete or unambiguously identifiable as a unique CCAD code ( $n=34$  records in the development set;  $n=4$  records in the test set), we replaced the information with the official CCAD code. When multiple codes were recorded in one procedure field, we split these into separate procedure fields in sequential order for that record ( $n=8$  records in the development set;  $n=4$  records in the test set). In cases in which this information was ambiguous, we replaced it with 'not an official CCAD code' ( $n=371$  in the development set;  $n=184$  in the test set). After cleaning the procedural information, those records containing only non-cardiac procedures (identified by NM), such as pericardiocentesis, were removed from the data set ( $n=552$  in the development set;  $n=243$  in the test set).

### Comorbidity fields

Up to 10 comorbidity codes are stored in a single data field within the CCAD data set as a concatenated string. These were split into separate fields. As with the diagnosis and procedure fields, some of the information recorded was not automatically identifiable as an official CCAD code. In cases in which this information was ambiguous, we replaced it with 'not an official CCAD code' ( $n=8$  records in the development set;  $n=27$  records in the test set). We replaced information with an official CCAD code when there was an anomalous numerical code format (e.g. additional spaces or dots) or the numerical component of the code was missing but the text part was complete or unambiguously identified with an official code.

## Defining comorbidity

International Paediatric Congenital and Cardiac codes defining comorbid conditions were grouped into four categories. Comorbid conditions appearing in any of the comorbidity or diagnosis fields were classed as such:

- Prematurity. This is defined by CCAD as being gestational age < 37 weeks at birth.
- Down syndrome or trisomy 21.
- Congenital comorbidity of all types other than Down syndrome. This essentially includes all genetic syndromes, clinical constellations of features that constitute a recognised syndrome and congenital structural defects of organs other than the heart.<sup>38</sup>
- Acquired comorbidities. This is a group of codes increasingly used over time in CCAD to define acquired non-cardiac abnormalities of other organ systems that were present preoperatively. These include conditions generally arising secondary to heart disease affecting other organ systems.<sup>39</sup> Examples include renal failure and necrotising enterocolitis.

## Defining episodes of surgical management for analysis

There was extensive discussion among the project team about how to assign outcomes to patients undergoing multiple operations. The problem faced is that, for a patient who has a reoperation within 30 days of their first operation, there is genuine ambiguity as to which operation to attribute a death to that occurs within 30 days of both operations. In the analysis of these cases there is scope for double-counting deaths (and for that matter survivals) or under-reporting surgical activity if reoperation or reoperated patients are ignored. Another alternative would be to use as an outcome measure for the first operation the vital status of the patient 30 days after the last operation in a sequence of reoperations within 30 days. Given the focus of this project on developing a risk model to facilitate timely monitoring of

short-term outcomes, the final decision was to have a '30-day episode' as the unit of analysis. A 30-day episode starts with the first *surgical* procedure on a patient. This episode is then assigned an outcome of alive or dead according to the vital status of the patient 30 days after this first surgical procedure. Any further surgical procedures that the same patient underwent within 30 days of this first procedure were not included in model development. The next procedure recorded for the same patient > 30 days after the first surgical procedure was treated as the start of a new 30-day episode.

Examples of how the 30-day episode was allocated are shown in *Figure 2*.

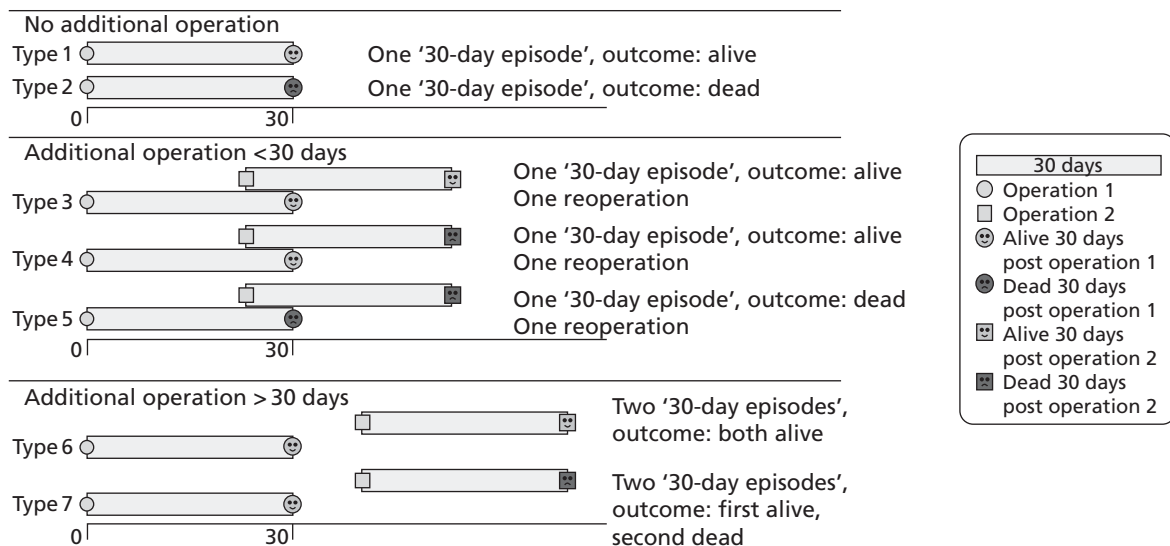
The 30-day episode unit of analysis was allocated using the same protocol in both the development set and the test set.

One drawback of this approach is that not all surgical activity contributes to the risk model development. This does not affect the utility of the risk model as long as, when used for monitoring, it is applied to 30-day episodes as defined above. That said, the level of reintervention within 30 days is rightly seen as important outcome information. For this reason we retained the information on the number of further procedures (both surgical and catheter) recorded for each patient within each 30-day episode and devised ways of presenting this additional information alongside risk-adjusted 30-day mortality.

**Additional issues encountered with matching records pertaining to the same patient**

To construct 30-days episodes it was essential to identify records within the data set pertaining to the same individual patient. As part of our work checking the consistency of the data set, we came across various instances in which the same apparent patient (as identified by pseudonymised NHS number or pseudonymised hospital number) had records in which the data did not seem consistent (e.g. by age or sequence of diagnoses or procedures). We identified all records in the development set in which the pseudonymised NHS number and the pseudonymised institution ID were the same but the patient's pseudonymised hospital number was different (484 records in the development set; 192 records in the test set), as this potentially indicated that the records actually pertained to different patients. Additionally, we identified all records for an apparently single patient (as identified by pseudonymised NHS number) for whom sequential ages were inconsistent with the procedure dates.

Through further investigation, David Cunningham confirmed that these inconsistencies were the result of errors in pseudonymisation or incorrectly entered hospital or NHS numbers. For those records for which the



**FIGURE 2** Examples of how the 30-day episode was allocated.

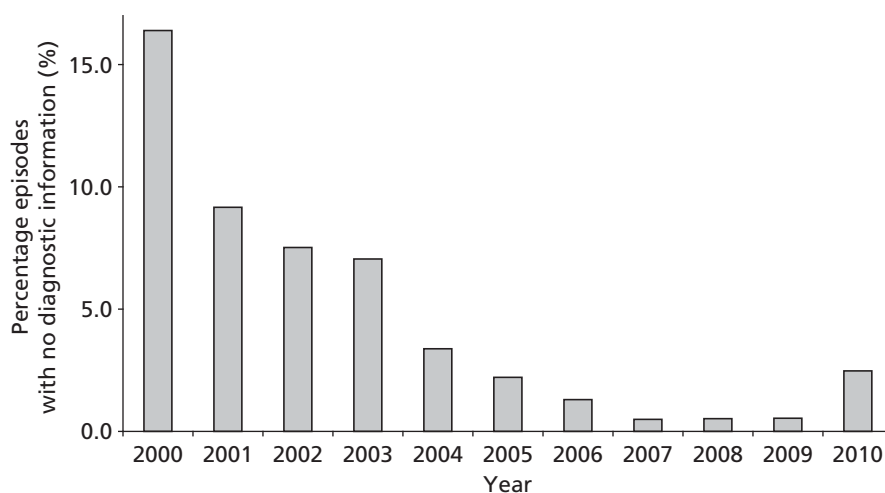
pseudonymised NHS number was invalid, patients were identified using the pseudonymised hospital number, and for all other cases it was determined that the pseudonymised hospital number was incorrect and the records pertained to the same patient (as identified by their pseudonymised NHS number).

Of those records that were originally removed for being associated with patients recorded as having only catheter procedures, 15 were identified as having an incorrect pseudonymised NHS number. However, correcting for this did not result in any of these records being associated with a patient who had undergone a surgical procedure.

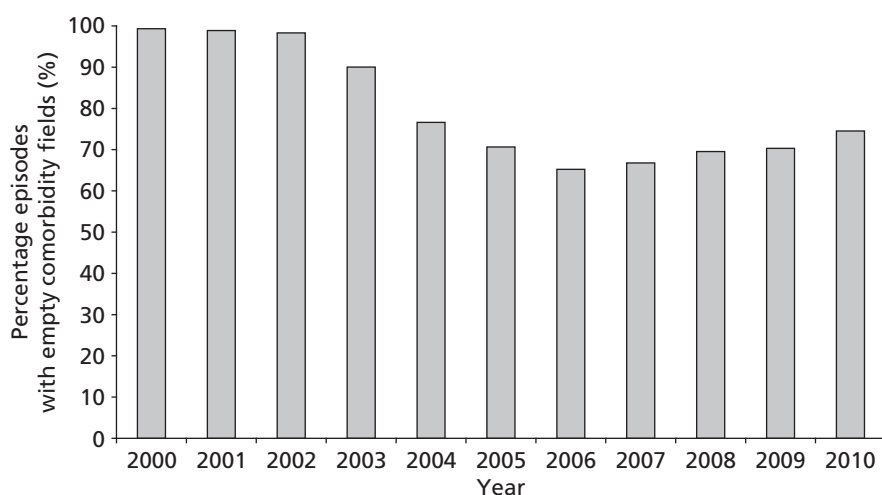
## Field completeness and descriptive analyses

As stated above, the completeness of procedural data was very high. The completeness of diagnosis data fields improved over the period covered by the data set, as can be seen in *Figure 3* (development set).

There is also an apparent improvement in recording of comorbidity data (*Figure 4*), although, whereas diagnostic information should be available for all episodes, an empty comorbidity field can correctly signify 'no comorbidity' rather than missing data.



**FIGURE 3** The completeness of diagnosis data fields over time (episode level within the development set).



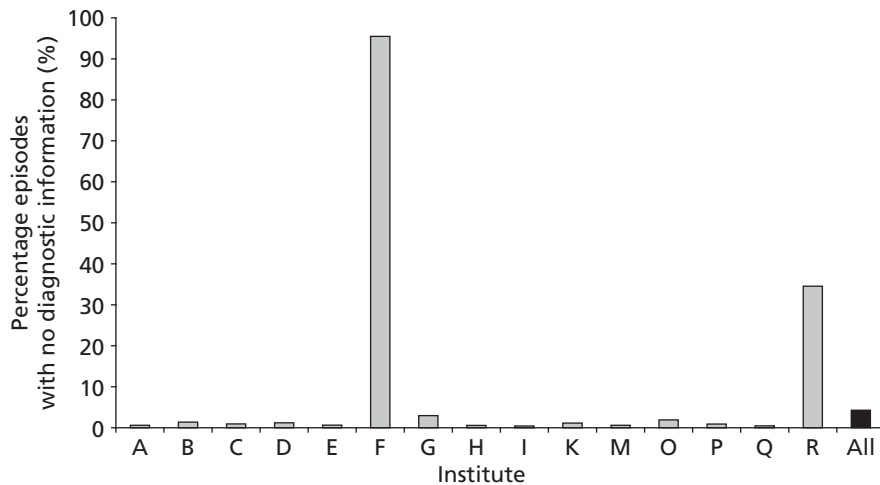
**FIGURE 4** The completeness of comorbidity data fields over time (episode level within the development set).

The completeness of the diagnosis and comorbidity data fields varied across institutions, as depicted in *Figures 5* and *6* respectively. *Table 1* shows how case load varies across these institutions in terms of the number of episodes in the development set.

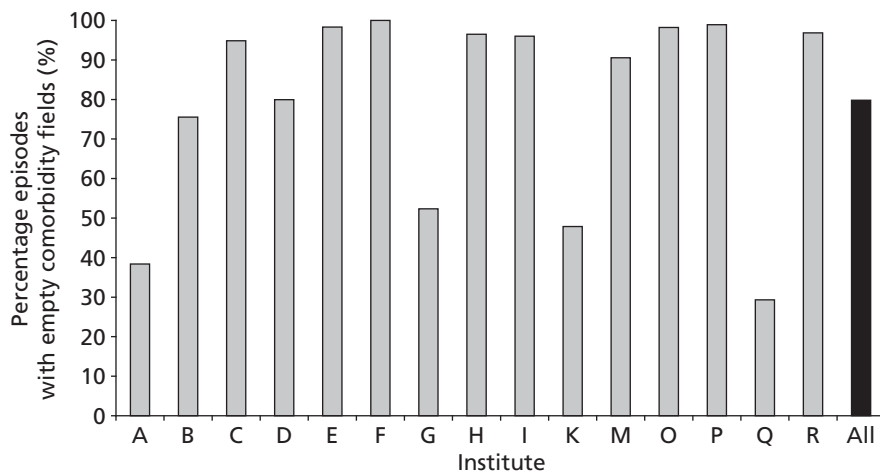
Data completeness for the other fields, as recorded in the entire development set, is given in *Table 2*.

There was a greater proportion of male patient episodes than female in the development set (*Table 3*). The age and weight distributions at episode level are shown in *Figures 7* and *8*, respectively, and a scatterplot of weight against age is given in *Figure 9*.

*Figure 10* shows the number of episodes with missing 30-day status and/or patient age that were removed from the development and test sets, and the number of episodes with missing or anomalous patient weight that were excluded from the test set and replaced in the development set by the mean weight for age. We note that 90% of records with missing status occurred before 2002. We examined the data in the development set to look for differences between records with valid 30-day status and records missing 30-day status and did not find any significant differences.



**FIGURE 5** The completeness of diagnosis data fields by institution (episode level within the development set). Institutions J, L and N have case loads  $\leq 3$  within the entire development set and so are not included.



**FIGURE 6** The completeness of comorbidity data fields by institution (episode level within the development set). Institutions J, L and N have case loads  $\leq 3$  within the entire development set and so are not included.

**TABLE 1** Episode volume by institution  
(entire development set,  $n=27,212$ )

Institute	Number of episodes	% of all episodes
A	672	2.5
B	2558	9.4
C	1784	6.6
D	1233	4.5
E	1381	5.1
F	243	0.9
G	1351	5.0
H	3106	11.4
I	1824	6.7
J	1	0.0
K	1937	7.1
L	3	0.0
M	3710	13.6
N	2	0.0
O	2245	8.3
P	654	2.4
Q	2545	9.4
R	1963	7.2

**TABLE 2** Data completeness in the development set  
( $n=27,212$ )

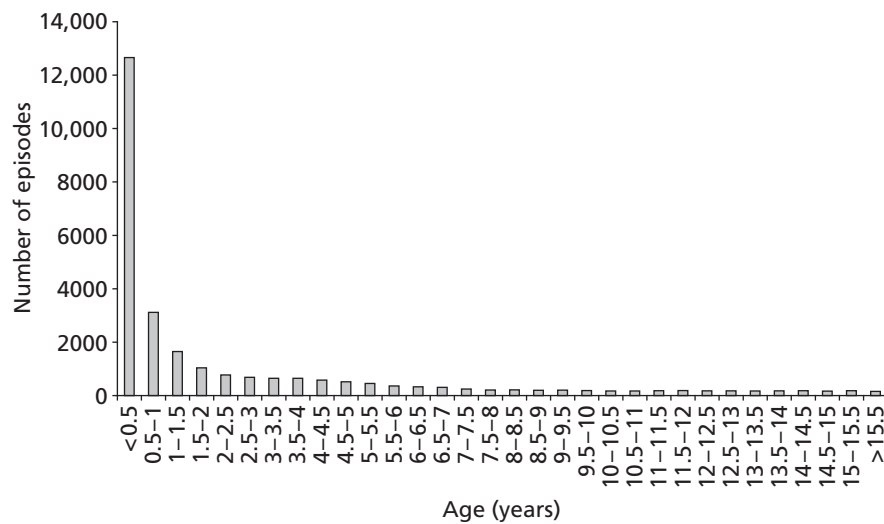
Data field	Number of episodes with missing data	% of episodes with missing data
Age	77	0.3
Weight	1327	4.9
Sex <sup>a</sup>	189	0.7
Antenatal diagnosis	8396	30.9
Sternotomy sequence	8229	30.2
Ethnicity	5848	21.5
Deprivation	7008	25.8

<sup>a</sup> Entries of 'not known', 'unclear' or 'not specified'.

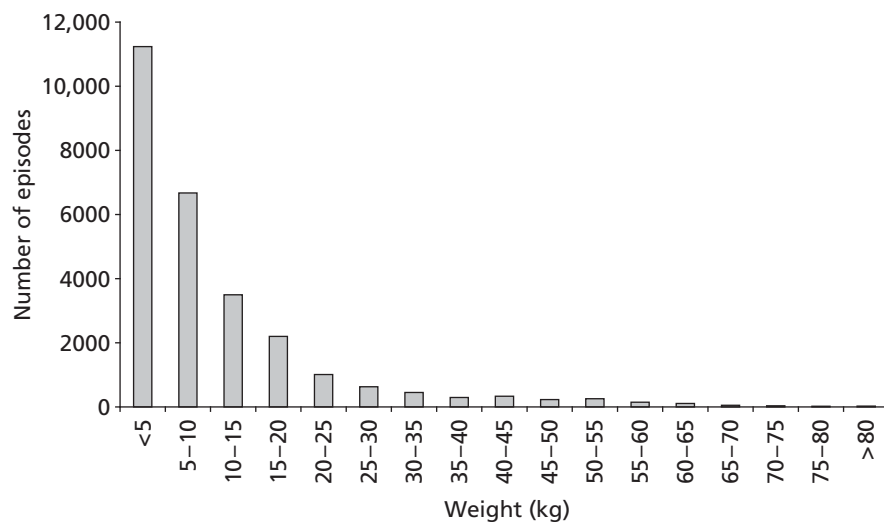
**TABLE 3** Episodes in the development set by sex ( $n=27,212$ )

Sex	Number of episodes	% of episodes
Male	14,978	55.0
Female	12,045	44.3
Unknown <sup>a</sup>	189	0.7

<sup>a</sup> Entries of 'not known', 'unclear' or 'not specified'.

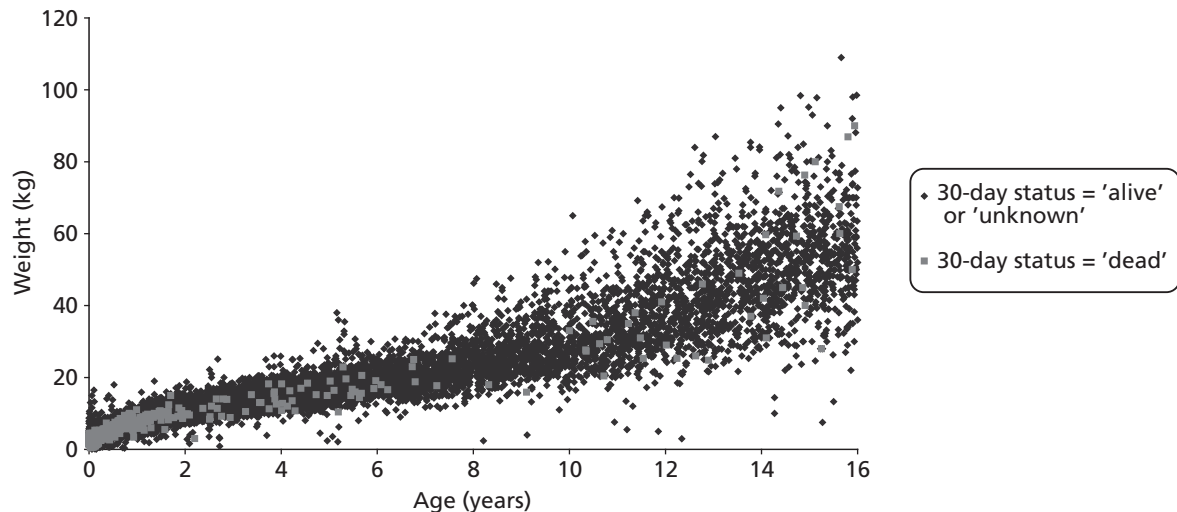


**FIGURE 7** The distribution of patient age at episode level in the development set ( $n=27,212$ ).

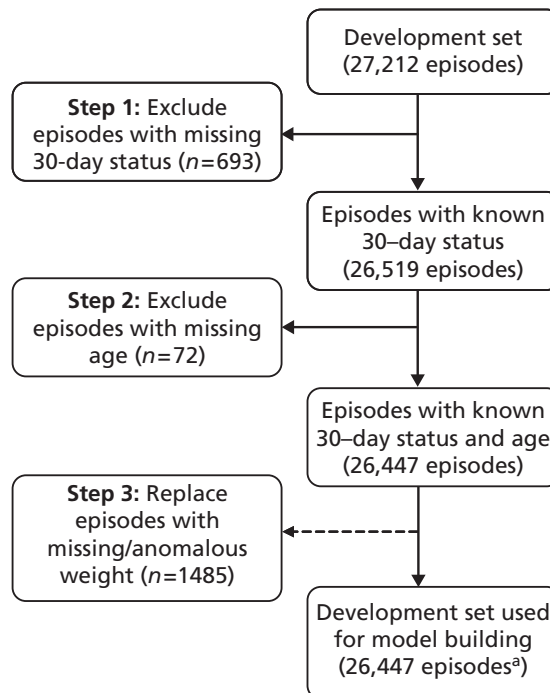


**FIGURE 8** The distribution of patient weight at episode level in the development set ( $n=27,212$ ).





**FIGURE 9** Weight vs age scatterplot for episodes in the development set ( $n=27,212$ ). Black markers are for 30-day status either alive or unknown; grey markers are for 30-day status dead.



**FIGURE 10** Number of episodes with missing 30-day status and/or patient age and number of episodes with missing or anomalous patient weight. Episodes with missing 30-day status and/or patient age were removed from the development set, whereas episodes with missing or anomalous patient weight were replaced by the mean weight for age (see text). a. Includes episodes with missing diagnostic information ( $n=948$ ).



# Chapter 5 Defining primary cardiac diagnosis and identifying univentricular heart status

## Preamble on diagnosis classification

Typically, paediatric cardiac disease has multiple components so that a single patient's CCAD record often has several diagnosis codes. IPCC codes<sup>32,40</sup> are used to list each item contributing to a patient's heart defect, as well as any associated conditions. Allocation of a 'primary' or 'underpinning' diagnosis is challenging and no consensus exists regarding an optimal method. Previous authors have identified the primary diagnosis as the lesion with the greatest anatomic severity,<sup>41</sup> but for some situations such an allocation may be inappropriate. On the other hand, choosing the lesion requiring the earliest intervention<sup>42</sup> also challenges clinical intuition in some scenarios. An alternative approach is to use the underlying physiology of the disease to group diagnoses,<sup>1,43</sup> as in the Extracorporeal Life Support Organisation Registry database;<sup>44</sup> the limitation here is that physiological groupings may encompass a wide range of conditions with correspondingly different prognoses. In principle, diagnostic information could also be categorised in relation to risk, much as short-term surgical data is in systems such as RACHS-1,<sup>25</sup> but actuarial risk rather than short-term risk data would be required to make this pertinent.

Recent epidemiological studies of CHD in the northern region of the UK by Wren *et al.*<sup>45-47</sup> specified possible 'primary diagnoses' and placed them in an explicitly hierarchical list, broadly by clinical severity. An individual patient was then allocated to the highest grouping consistent with their anatomical features, informed if necessary by the lesion precipitating their earliest intervention. Limitations of this categorisation include its failure to consistently distinguish single or biventricular status, and its exclusion of certain diagnoses that still did not fit into the scheme. Our study used the approach of hierarchical categorisation as the starting point, modifying it for the different context of CCAD. Motivated by the possibility that a large number of combinations of diagnostic and procedure categories, some with very small patient numbers, could undermine robust statistical modelling, we also elected to develop a second scheme consisting of larger and less refined groupings based on a categorisation described by Clancy *et al.*,<sup>48</sup> according to the number of functioning ventricles and the presence of aortic obstruction. We anticipated that both schemes would allow patients to be prospectively categorised by clinicians submitting CCAD data, and that there was the potential for either or both schemes to add information to our proposed risk-adjustment model. The diagnosis schemes described in the following sections were developed in close collaboration with Dr Rodney Franklyn and Dr Catherine Bull, both experts in paediatric cardiology.

## Diagnosis classification scheme 1 (mapping International Paediatric Congenital and Cardiac codes to primary diagnosis)

In this section we first describe the methods used to map each of the individual IPCC codes into one of 27 diagnostic groupings. We then describe how the several diagnostic groupings that may appear in any particular record were combined to determine a single 'primary cardiac diagnosis' for that record. We used the entire development set to devise the scheme.

### International Paediatric Congenital and Cardiac diagnosis code to diagnosis groupings

Each IPCC diagnosis code included in CCAD was assigned to a diagnosis category by either RF or CB based on the list described by Wren *et al.*<sup>45,46</sup> A modest expansion of the categories in the Wren system to better meet the specific context of the CCAD data set was agreed by the panel of clinicians (KB, CB, VT, NM and RF). This

included defining an ‘acquired’ cardiac disease category and a ‘congenital miscellaneous’ category for rarer congenital defects. Additionally, some information in the diagnosis fields related to either procedural or comorbidity codes or indicated that the heart was structurally normal. We allocated such diagnoses to ‘procedure’, ‘comorbidity’ and ‘normal’ categories respectively. There was iterative refinement of this mapping of individual codes, and selection of categories, through expert discussion and data inspection to identify anomalies. The full allocation of each IPCC code to a diagnostic category is given in *Appendix 2*.

### ***At the level of a patient record***

Once each diagnostic IPCC code (which could appear in diagnosis fields 1–6 within the CCAD data set) was assigned to a cardiac diagnosis grouping, we decided on a simple hierarchy among the diagnosis groupings to allocate a single ‘primary diagnosis’ to each record. An initial hierarchy was constructed following the example of the Wren *et al.* studies,<sup>46</sup> which orders primary diagnoses based mainly on the clinical assessment of complexity (most complex first, least complex last). We also considered the order of procedures in the CCAD surgical procedure algorithm,<sup>8</sup> which groups procedure codes into ‘specific procedures’ and also relies on a hierarchy that has been subject to various stages of refinement by a group of clinicians. The decision about which primary diagnosis is most important is very clear in certain scenarios but may be subjective in others, relating to the overall management strategy rather than a particular operation, which may be one of a series. Therefore, the panel of clinicians modified the primary diagnosis hierarchy further following an iterative process in which a version of the hierarchy was applied and resulting tabulations of the attributed primary diagnosis categories against related CCAD ‘specific procedures’ were reviewed. Anomalies in the hierarchy were identified based on detection of incongruous matching between diagnoses and procedures. When incongruities were noted, these were explored by revisiting the individual codes within the anomalous records, leading to further revisions of the hierarchy. This process was repeated for three versions of the hierarchy, at which point no further obvious incongruities between diagnosis categories and specific procedures were noted that could not be attributed to coding errors (see *Table 4* for the final hierarchy).

## **Diagnosis classification scheme 2 (mapping International Paediatric Congenital and Cardiac codes to diagnostic groups defined by ventricle number and aortic obstruction)**

We describe first the methods used to categorise individual IPCC codes before describing how the codes relevant to any particular record were combined. We used the entire development set to devise the scheme.

### ***At the level of the International Paediatric Congenital and Cardiac diagnosis code***

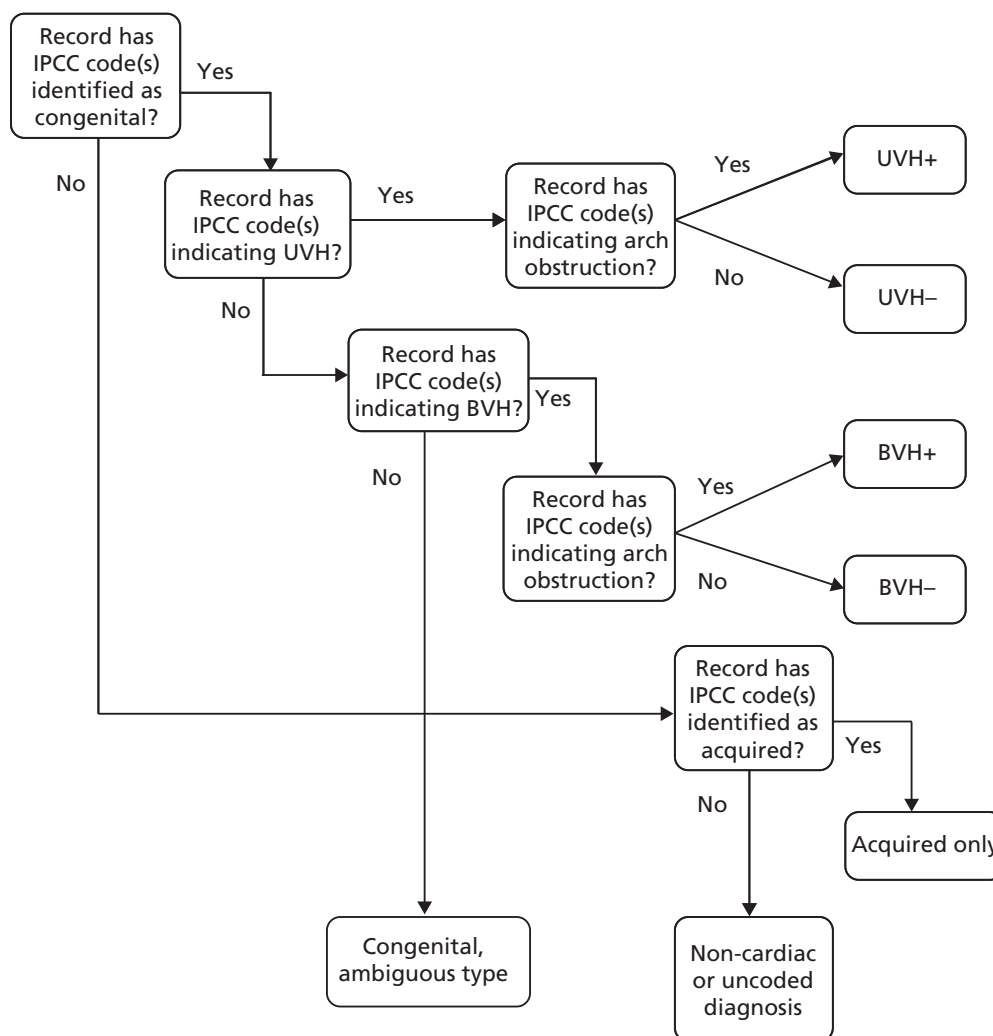
Each individual IPCC code was categorised as a ‘congenital’ diagnosis, ‘acquired’ diagnosis or neither (e.g. when diagnostic fields contained a coding omission or error, or non-cardiac information such as comorbidity). We then identified congenital codes that, in isolation, definitely indicate that a patient has either a univentricular heart (UVH) or a biventricular heart (BVH). Note that, for a minority of congenital diagnosis codes, the ventricular structure is ambiguous (e.g. inferior vena cava abnormality). Finally, we identified those congenital diagnosis codes that additionally indicate the definite presence of an arch or a systemic arterial obstruction.

### ***At the level of a record***

If any of the IPCC codes within a record were identified as a congenital or an acquired diagnosis, or indicated a UVH or the presence of an arch/systemic arterial obstruction, this attribute was carried forward to the entire record to be used as a candidate variable in model development. However, a record was identified with a BVH only if there was at least one IPCC code within the record that indicated a BVH and none that indicated a UVH.

Any given record within the CCAD data set will have one of a large number of possible combinations of these five potential diagnostic attributes, some of which are common and some of which are much rarer. A given combination of diagnostic attributes would not necessarily represent a clinically intuitive patient group and, so, for the purposes of defining subgroups of patient in which to assess the performance of the risk model, we used the methodology set out in *Figure 11* to assign records to one of seven clinically intuitive diagnostic groups. Records with an acquired code but no congenital codes (e.g. 'cardiomyopathy') fell into the 'acquired only' group.

Motivated by the intended prospective use of this mapping scheme, in which the attributes would almost always be known preoperatively, we retrospectively assigned attributes to a minority of records in which the recorded procedure unambiguously defined them. The CCAD specific procedure field was used to assign diagnostic information to a record in which the procedure was deemed to be an unambiguous indication of UVH [Norwood procedure (stage 1); Fontan procedure; and bidirectional cavopulmonary shunt] or the presence of an arch or systemic arterial obstruction [Norwood procedure (stage 1); isolated coarctation repair; interrupted aortic arch repair; truncus and interruption repair; aortic valvotomy; subvalvar aortic stenosis repair; and supra-valvar aortic stenosis repair].



**FIGURE 11** Flow diagram depicting the method for allocating a record to a diagnosis group on the basis of diagnosis attributes assigned to diagnosis codes within that record.

## The resultant diagnostic classification schemes

### Diagnosis classification scheme 1

The final hierarchy is presented in *Table 4*, with hypoplastic left heart syndrome (HLHS) at the top, followed by the more complex CHD diagnoses, many of which may be broken down into several constituent defects. 'Simple' diagnoses that can occur either in isolation or as a component of a more complex lesion, such as ventricular septal defect (VSD) and patent ductus arteriosus (PDA), are towards the bottom of the hierarchy; thus, more complex heart defects are removed at an earlier stage leaving the isolated 'simple defects' in their own separate categories. The 'miscellaneous congenital' category, which consists of rarer defects, lies at the end of the hierarchy followed by acquired diagnoses and then the non-cardiac or uncoded diagnosis groups. The five most common primary diagnoses were VSD (13.0%), PDA (10.4%), Fallot's tetralogy (9.5%), atrioventricular septal defect (AVSD) (8.7%) and aortic arch hypoplasia with or without VSD (7.8%). The comparison between primary diagnoses and procedures at the record level is given in *Appendix 3* and shows good face validity.

### Diagnosis classification scheme 2

The breakdown of patients by diagnosis scheme 2 is given in *Table 5*. The most common cardiac diagnosis category in this scheme was BVH without arch or arterial obstruction [BVH(-)], irrespective of any additional acquired diagnoses, accounting for 68.9% of all patients; the next most common diagnosis category was BVH(+) at 15.4% followed by UVH(+) at 4.8%, UVH(-) at 4.5% and acquired diagnoses only at 2.5%.

There was good concordance between the two primary diagnosis schemes, with, for example, 100% of patients with HLHS falling into the UVH(+) category and 77.9% of patients with a primary diagnosis of functionally UVH falling into the UVH(-) category and the remaining 22.1% falling into the UVH(+) category. Of those patients with an ambiguous number of functional ventricles (2.1% of patients), 79.2% had 'transposition of the great arteries (TGA) + VSD/double-outlet right ventricle – TGA type' and 10.7% 'miscellaneous congenital' disorders.

Importantly, there was evidence that the diagnosis categories add descriptive information when the procedural information is ambiguous with respect to the indication for surgery, such as the 19.0% of all operations classed as 'not a specific procedure', of which 96.5% were allocated a primary cardiac diagnosis (including acquired) in diagnosis scheme 1, the most common being acquired (12.7%), miscellaneous congenital (9.6%) and VSD (9.3%). In diagnosis scheme 2, 94.4% were allocated a primary cardiac diagnosis category (including acquired), the most common being BVH(-) (61.2%). Further specific procedures, for example the Fontan procedure (4.4% of all operations), may benefit from breakdown by diagnosis scheme 1 (100% UVH in diagnosis scheme 2), with 23.3% of procedures performed in patients with HLHS, 32.1% in patients with functionally UVH, 8.9% in patients with pulmonary atresia/intact ventricular septum and 8.8% in miscellaneous congenital patients.

## Summary of work on diagnosis classification

This process has been successful in identifying diagnostic categories likely to add value to procedural information in developing a risk model for 30-day postoperative mortality. This grouping of patient records in CCAD based on diagnostic information enables us to comment for the first time on the frequency or proportions of patients with particular diagnostic features. For example, we can now say that the proportion of patients in CCAD (which captures all UK children who have undergone an intervention) since 2004 who have a UVH is 9.3%.

We selected terms to label the primary diagnosis groupings, such as 'ventricular septal defect' and 'hypoplastic left heart syndrome', that will be recognisable to both laypersons who may have an interest in this topic and professionals. This choice was in keeping with the philosophy of CCAD, which shares information about outcomes with the public; we considered that more technical medically focused schemes for describing and allocating primary diagnoses would be less suitable for this purpose.

In addition to contributing to the development of the risk model, this work on identifying a primary cardiac diagnosis will also inform subsequent analyses of long-term outcome by diagnosis in CCAD and may provide valuable information in terms of national audit of outcomes.

**TABLE 4** The proportions of patients by diagnosis category scheme 1, taken from the development set for 2004 onwards (as diagnostic data quality improves after this date), for those records pertaining to the first instance of a patient in the data set

Diagnosis group	% of patients (95% CI <sup>a</sup> )
HLHS	3.6 (3.3 to 4.0)
Functionally UVH	3.4 (3.1 to 3.7)
Common arterial trunk (truncus arteriosus)	1.3 (1.1 to 1.5)
TGA + VSD/double-outlet right ventricle – TGA type	5.5 (5.1 to 5.9)
Interrupted aortic arch	0.8 (0.7 to 1.0)
TGA (concordant AV and discordant VA connections) and IVS	2.9 (2.6 to 3.2)
Pulmonary atresia (including pulmonary atresia + IVS)	1.4 (1.2 to 1.6)
Pulmonary atresia + VSD (including Fallot type)	2.9 (2.6 to 3.2)
AVSD	8.7 (8.2 to 9.1)
Fallot/double-outlet right ventricle – Fallot type	9.5 (9.0 to 10.0)
Aortic valve stenosis (isolated)	2.3 (2.1 to 2.6)
Tricuspid valve abnormality (including Ebstein's)	1.0 (0.8 to 1.1)
Mitral valve abnormality (including supralvar, subvalvar)	2.2 (1.9 to 2.4)
Totally anomalous pulmonary venous connection	1.8 (1.6 to 2.0)
Aortic arch obstruction with or without VSD/ASD	7.8 (7.4 to 8.3)
Pulmonary stenosis	2.7 (2.4 to 2.9)
Subaortic stenosis (isolated)	1.4 (1.2 to 1.6)
Aortic regurgitation	1.2 (1.0 to 1.4)
VSD	13.0 (12.5 to 13.6)
Interatrial communication (ASD)	7.6 (7.2 to 8.1)
PDA	10.4 (9.9 to 10.9)
Miscellaneous congenital	4.1 (3.8 to 4.5)
Acquired <sup>b</sup>	2.6 (2.3 to 2.9)
Non-cardiac or uncoded diagnosis	2.2 (1.9 to 2.4)
Procedure <sup>b</sup>	
Comorbidity <sup>b</sup>	
Normal <sup>b</sup>	
Empty/unknown <sup>b</sup>	
All patients	100.0

ASD, atrial septal defect; AV, atrioventricular; CI, confidence interval; IVS, intact ventricular septum; TGA, transposition of the great arteries; VA, ventriculoarterial.

a 95% CIs were calculated using Wilson's approximation.

b Non-congenital diagnoses.

**TABLE 5** The proportions of patients by diagnosis category scheme 2, taken from the development set for 2004 onwards (as diagnostic data quality improves after this date), for those records pertaining to the first instance of a patient in the data set

Diagnosis group	% of patients (95% CI <sup>a</sup> )
Congenital UVH(+)	4.8 (4.4 to 5.1)
With acquired	0.5 (0.4 to 0.6)
Without acquired	4.3 (4.0 to 4.7)
Congenital UVH(-)	4.5 (4.1 to 4.8)
With acquired	0.5 (0.4 to 0.6)
Without acquired	4.0 (3.7 to 4.4)
Congenital BVH(+)	15.4 (14.8 to 16.0)
With acquired	1.4 (1.3 to 1.7)
Without acquired	13.9 (13.4 to 14.5)
Congenital BVH(-)	68.9 (68.1 to 69.7)
With acquired	6.2 (5.8 to 6.7)
Without acquired	62.6 (61.8 to 63.5)
Acquired only	2.5 (2.3 to 2.8)
Congenital, ambiguous type	2.1 (1.9 to 2.4)
Non-cardiac or uncoded diagnosis <sup>b</sup>	1.9 (1.7 to 2.1)
All patients	100.0

CI, confidence interval.

a 95% CIs were calculated using Wilson's approximation.

b Non-congenital diagnoses.



## Chapter 6 Univariate associations with mortality

All of the analysis presented in this section was performed at episode level in the entire development set, including all episodes with known 30-day status ( $n=26,519$ ) and excluding those with unknown status ( $n=693$ ). This corresponds to 21,682 patients, of whom 82.5% have only one episode, 13.3% have exactly two episodes and 4.2% have more than two episodes recorded in the development set (Table 6). Note that, following model development, an error was found in the specific procedure data for 1115 records (4% of the development set – see Chapter 8, *Global performance in test set* for details). In the figures and tables presented here we have decided that it is more useful to report the corrected data rather than the slightly different versions inspected by the team as part of the model development process.

### Initial exploratory analyses

First, descriptive analyses were performed to characterise the make-up of the development set population. The development set was then used as the basis for a number of exploratory analyses to learn more about preoperative factors reported to CCAD that are associated with outcome. We also investigated whether or not there were attributes of the data that would affect the development and use of risk scores. Mortality rates based on 30-day status [with 95% confidence intervals (CIs)] were calculated for the following preoperative risk factors, selected as candidate model predictors on the basis of clinical relevance and availability within the data set:

- patient information: sex, age, weight, weight-for-age z-score, ethnicity, sternotomy sequence, antenatal diagnosis
- procedural information: procedure type, specific procedure, episode number for patient
- comorbidity information: premature (from diagnosis and comorbidity fields), Down syndrome (from diagnosis and comorbidity fields), congenital non-Down syndrome comorbidity (from diagnosis and comorbidity fields), acquired comorbidity (from diagnosis and comorbidity fields)
- diagnostic information (see Chapter 5 for details): either scheme 1 diagnostic category or congenital attribute, UVH attribute, BVH attribute, congenital not assigned attribute, arch obstruction attribute, acquired attribute, comorbid diagnosis attribute, not cardiac diagnosis attribute.

For brevity and clarity, mortality rates for only the most relevant groupings of candidate factors are shown in Table 7 and Figures 12–18. In the development set overall ( $n=26,519$ ), the episode-level 30-day mortality rate was 3.2% (95% CI 3.0% to 3.4%).

In discussion with all project team members, potential preprocedural risk factors removed at this stage were:

- Antenatal diagnosis: in addition to concern about the level of missing data, it was felt that, although potentially acting as a surrogate marker for severity of congenital defect, the relationship between antenatal diagnosis and outcome was not straightforward and might not be stable with changing patterns of provision of antenatal diagnostic services.
- Preprocedure seizures: data quality was considered to be poor for this field.
- Preprocedure Paediatric Cerebral Performance Category (PCPC): data quality was considered to be poor for this field.
- Sternotomy sequence number: although this field would seem to offer information about the position in a sequence of operations for the patient, it was felt by the clinicians that data quality for this field was not sufficient to use it within the risk model. Using the number of times that the patient has previously appeared in the data set was explored, but this measure has no meaning for patients with sequences that started before the data set or whose care was not delivered entirely within the UK.
- Ethnicity: it was considered that, with the high level of missing data and a lack of a clear understanding of the mechanisms for genuine clinical differences in mortality risk in different ethnic groups, it would not be appropriate to include ethnicity in the model development process.

- Deprivation: the team considered that to adjust for deprivation could potentially contribute to complacency over possible social gradients in outcomes and that it would be better to exclude measures of deprivation from the risk model. Analysis of risk-adjusted outcomes by quintile of deprivation will be the focus of future work.

We note that episodes with a 30-day status of ‘unknown’ (2.5% of episodes in the development set) were not included in the denominator of univariate mortality rates, nor included within model development.

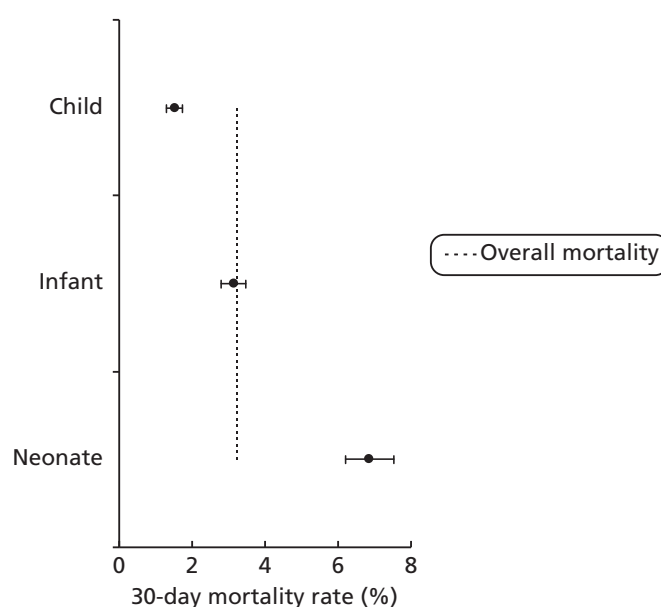
**TABLE 6** The numbers of patients with only one, exactly two and more than two episodes recorded in the development set

Number of episodes	Number of patients in the development set	% of patients in the development set (n=21,682)
Patients with only one episode	17,896	82.5%
Patients with exactly two episodes	2883	13.3%
Patients with more than two episodes	903	4.2%

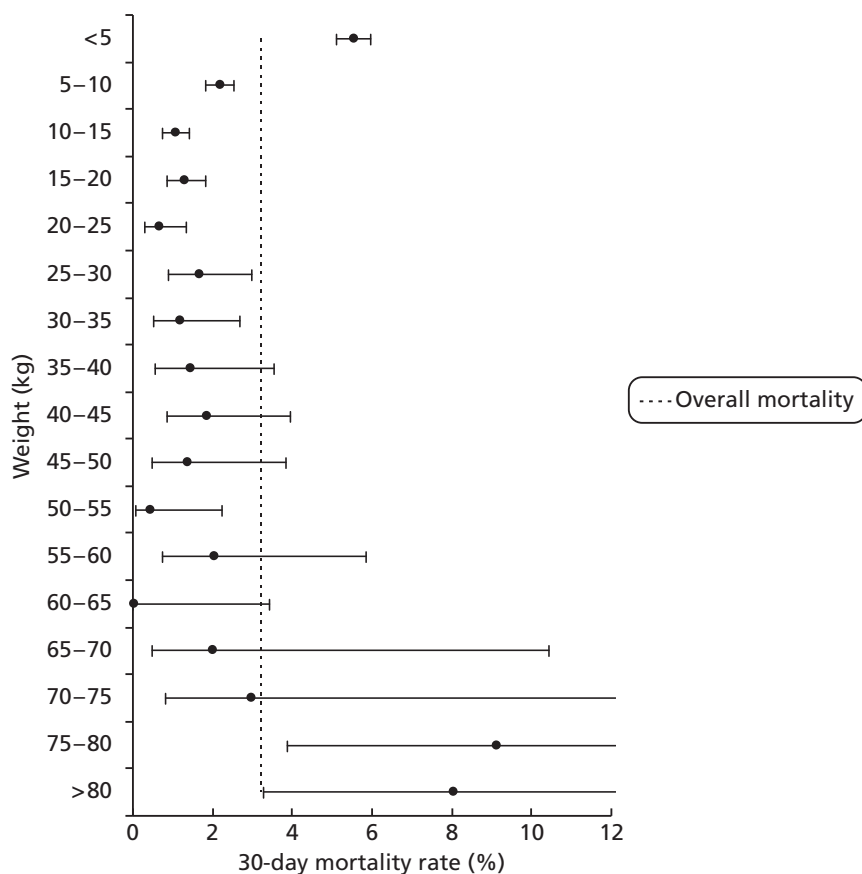
**TABLE 7** Mortality rates based on 30-day status for UVH attribute, procedure type and presence of a non-Down syndrome comorbidity taken from the development set at episode level (n=26,519)

	30-day mortality rate (%)	95% CI <sup>a</sup> (%)
Not UVH	2.6	2.4 to 2.8
UVH	6.8	6.1 to 7.7
Bypass	3.1	2.9 to 3.4
Non-bypass	3.5	3.1 to 4.0
No (non-Down syndrome) comorbidities	2.9	2.7 to 3.2
At least one (non-Down syndrome) comorbidity	5.5	4.7 to 6.4

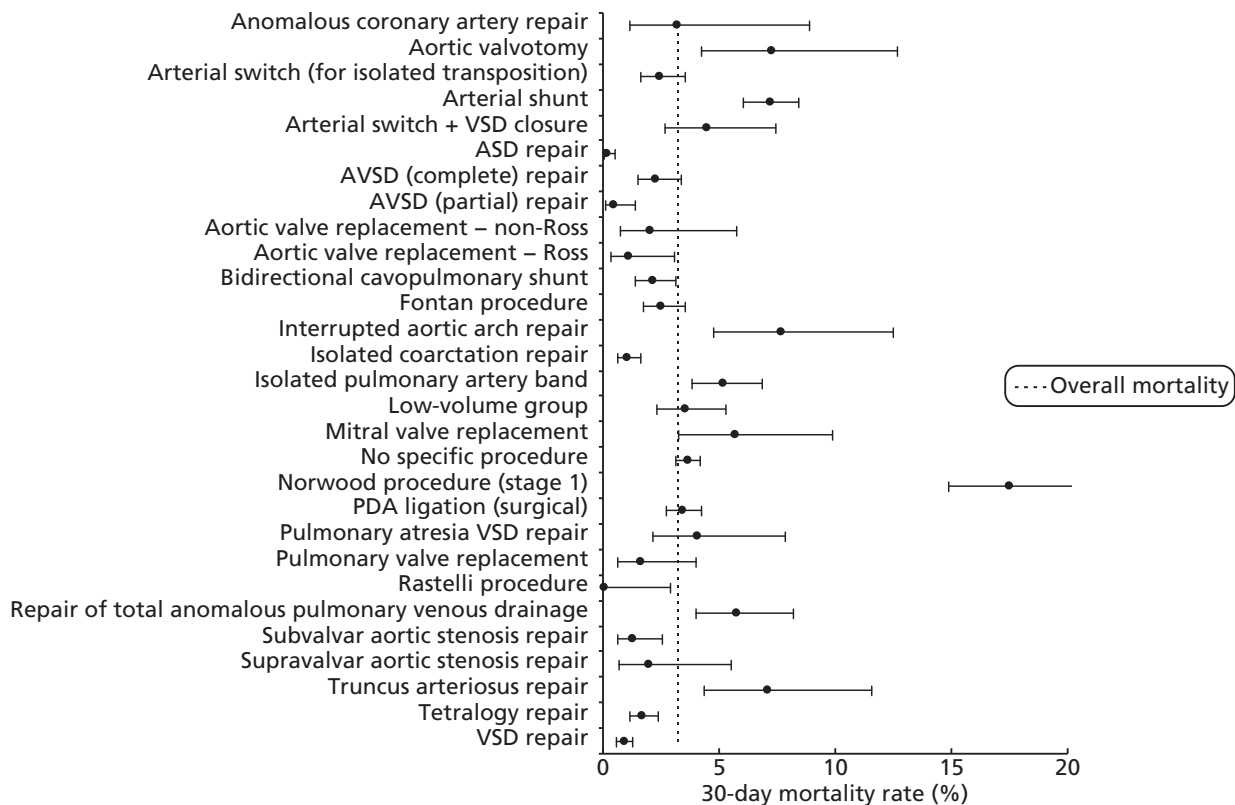
a CIs were calculated using Wilson’s approximation.



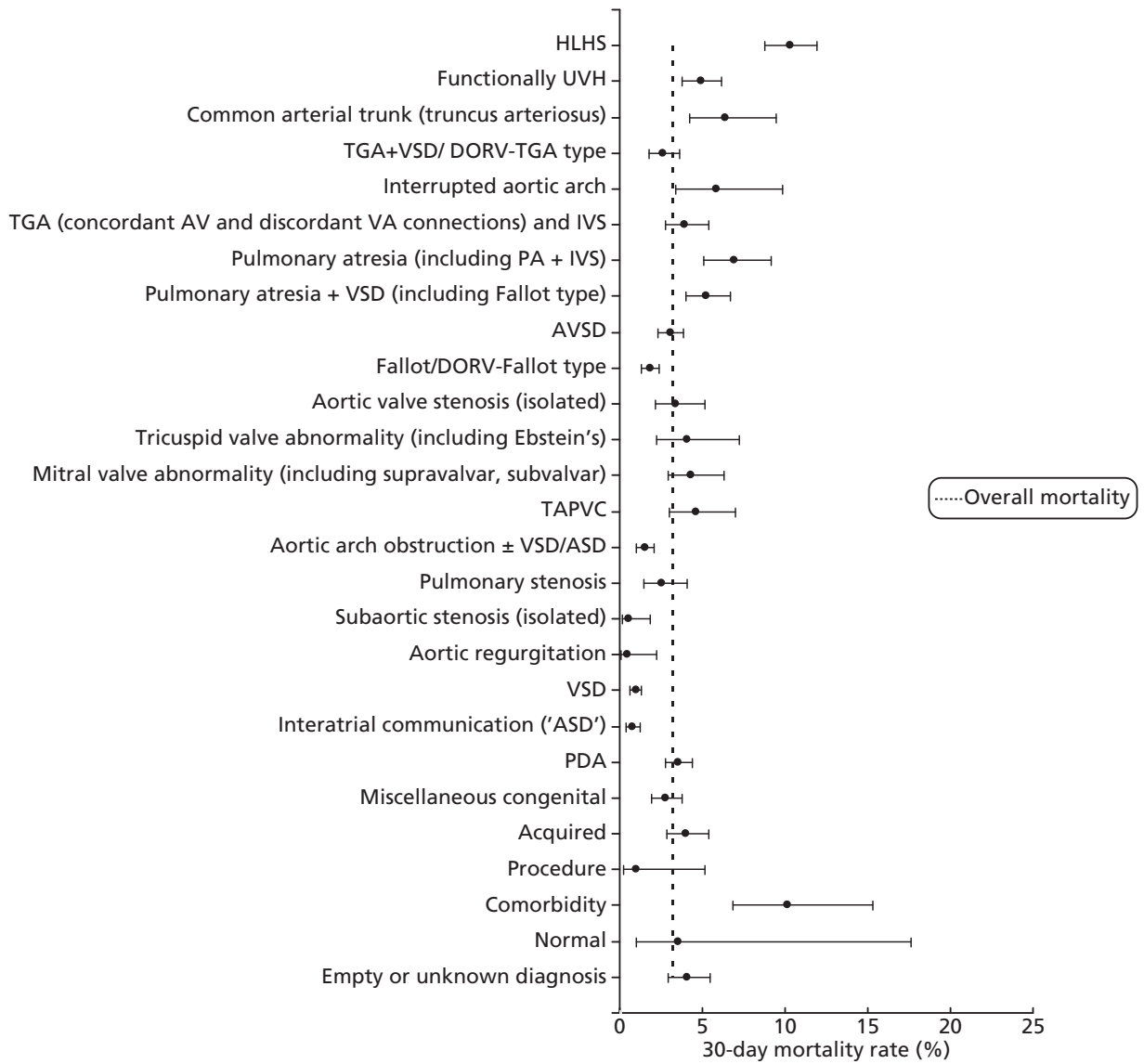
**FIGURE 12** Mortality rates (95% CIs) based on 30-day status for age bands within the development set (n=26,519). Dotted line indicates overall average mortality rate (3.2%).



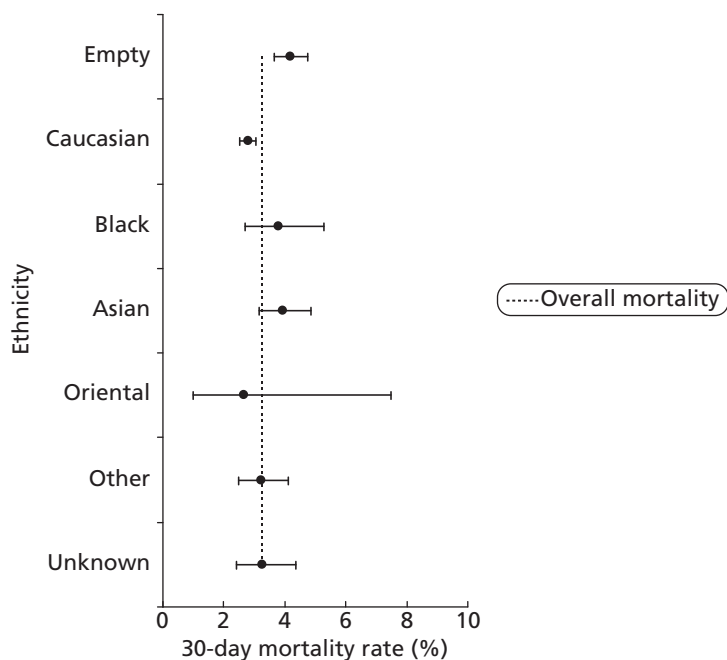
**FIGURE 13** Mortality rates (95% CIs) based on 30-day status for weight bands within the development set ( $n=26,519$ ). Dotted line indicates overall average mortality rate (3.2%).



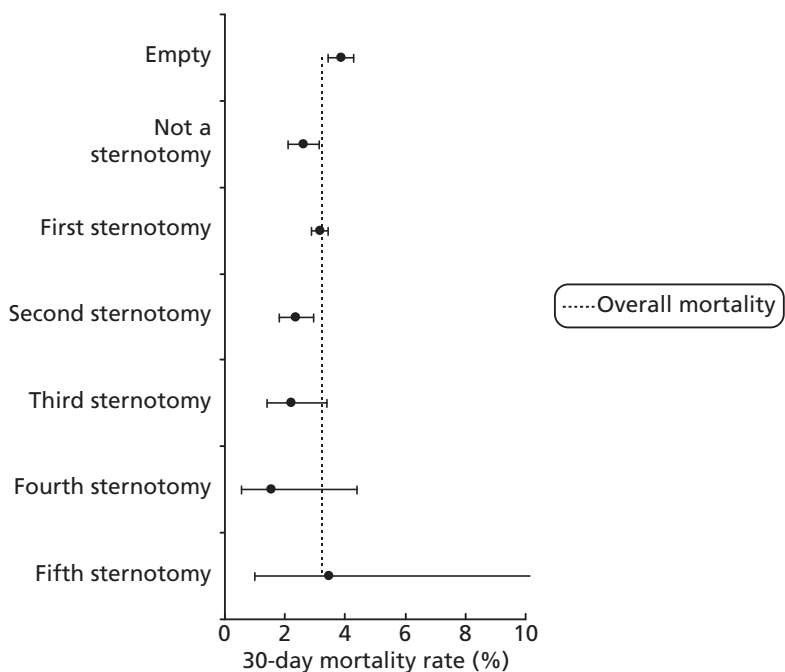
**FIGURE 14** Mortality rates (95% CIs) based on 30-day status for specific procedures within the development set ( $n=26,519$ ). Dotted line indicates overall average mortality rate (3.2%).



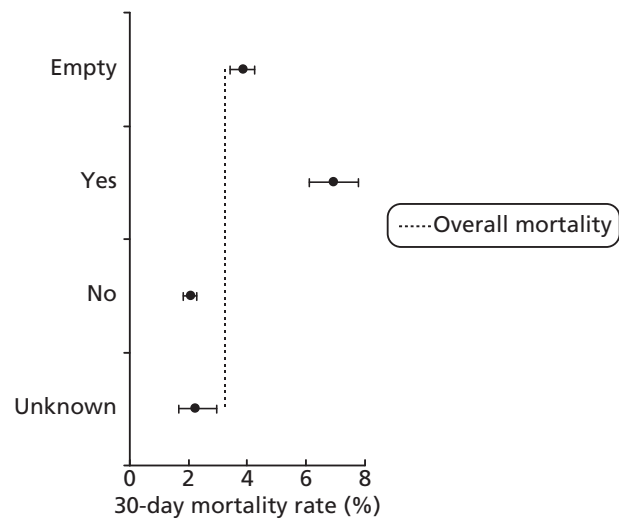
**FIGURE 15** Mortality rates (95% CIs) based on 30-day status for scheme 1 diagnosis groups within the development set ( $n=26,519$ ). Dotted line indicates overall average mortality rate (3.2%). ASD, atrial septal defect; AV, atrioventricular; DORV, double-outlet right ventricle; IVS, intact ventricular septum; PA, pulmonary atresia; TAPVC, totally anomalous pulmonary venous connection; VA, ventriculoarterial.



**FIGURE 16** Mortality rates (95% CIs) based on 30-day status for ethnicity within the development set ( $n=26,519$ ). Dotted line indicates overall average mortality rate (3.2%).



**FIGURE 17** Mortality rates (95% CIs) based on 30-day status for sternotomy sequence within the development set ( $n=26,519$ ). Dotted line indicates overall average mortality rate (3.2%).



**FIGURE 18** Mortality rates (95% CIs) based on 30-day status for antenatal diagnosis within the development set ( $n=26,519$ ). Dotted line indicates overall average mortality rate (3.2%).

# Chapter 7 Model development and selection

## Model building process

The process adopted for constructing the risk model consisted of the iterative use of the following steps.

### *Multivariate logistic regression*

Given that each variable under consideration was judged by the clinical team to have face validity, we employed a backward stepwise method<sup>49</sup> with variables removed from the model at each step if their removal did not significantly ( $p > 0.1$ ) worsen the performance of the model (based on the probability of the likelihood ratio statistic). This reduced list formed the basis of further manual exploration using subsets of the development set to investigate the stability of these risk factors (see below). The list of final risk factors included in the model was compiled based on a combination of their stability, their clinical face validity and their ease of use prospectively. We then used the 'enter' regression method to obtain the parameterisation of the final model. The multiple logistic regression analysis was conducted within PASW Statistics 18, release version 18.0.0 (SPSS Inc., Chicago, IL, USA). The regression output consisted of the factors included in the resulting model, the Hosmer–Lemeshow chi-squared statistic and contingency table, the predicted risk for each episode within the development set, or subset thereof, and the model parameters required to calculate the predicted risk for other episodes.

Being a representative sample of paediatric cardiac surgical procedures performed in the UK, the data set contained more than one record for some patients. Our aim was to develop a risk model for use (at episode rather than patient level) in monitoring risk-adjusted outcomes for an entire paediatric cardiac surgical programme, not just for those undergoing their first operation. For this reason we needed to include records pertaining to patients' second (and in some cases third and fourth) episodes in model development, with equal weight given to each episode rather than each patient.

For this reason it is important to note that the logistic regression output cannot be used to infer odds ratios and CIs associated with particular risk factors at a patient level, as not all observations used in the regression were independent. This inevitable non-independence of some observations in the data set strengthened the motivation for evaluating the goodness of fit and other model characteristics in an entirely distinct set of data rather than relying on the statistical performance of the model among the data set used for its development. Although less efficient in its use of data than other approaches such as repeated bootstrapping, this approach to model evaluation renders the statistical approach adopted in model development much less relevant when judging the goodness of fit or fitness for purpose of a model. Any deleterious effects of non-independent observations, overfitting of the data or other features of the model development process become apparent when the performance of the model is evaluated in the wholly separate test set.

### *Construction and interpretation of mean adjusted deaths compared against predicted plots to assess model performance*

The regression output was used to construct charts, known as MADCAP charts, of cumulative predicted and observed deaths against episode number, with episodes ordered by increasing predicted risk.<sup>37</sup> These give a graphical means of summarising and comparing the performance of the risk model. The end points of the two lines indicate the overall numbers of deaths predicted and observed over a series of cases. The extent to which the slope of the trace of cumulative observed deaths increases with episode number gives information about the discrimination of the model. The greater the 'bowing' towards the bottom right-hand corner of the MADCAP chart, the better the discrimination of the model. The deviations of the 'observed' trace from the 'predicted' trace give an impression of the performance and accuracy of the model across the spectrum of predicted risk.

Comparison of MADCAP charts was used as a way of informing decisions about the value added or lost in adopting different approaches to how particular variables were analysed and also of assessing the stability of model parameterisation when a model was parameterised using different random subsets of the development data set.

Importantly, the differential performance of models in different age bands and diagnostic groups was also studied as part of the development process.

### **Assessment of model stability**

Another step in the model development process was to subdivide the development data set at random into half. One half was used to either construct a model using backward stepwise regression techniques (see above) or parameterise a prespecified model using 'enter' regression techniques (see above), with the performance of the resulting model being evaluated in the other half. To assess the stability of model parameterisation, models were developed or parameterised using several (six to eight) random 50% subsets of the development set with the factors identified for inclusion and the coefficients for each factor tabulated and compared. The stability of model performance was assessed by comparing these groups of six to eight parameterisations using MADCAP charts (see above). We note that we explored the possibility of developing separate models within each age band. Although initially promising, these models were not robust when tested for stability and, so, were not considered further.

### **Simplification of individual candidate variables**

The process of checking the stability of model parameterisation in random subsets of the development data led us to conclude that, for some variables and for candidate models as a whole, there were too many different values (or 'degrees of freedom') that could be included within the model, the risk being that if there is too much freedom in a model then there is a risk of overfitting, giving good model performance in the development set that is not reproduced in test data. In these instances we simplified variables and used the steps outlined above to assess any trade-offs between model performance and stability of parameterisation.

## **Diagnostic information**

We used classification and regression tree analysis to investigate groupings of diagnostic scheme 1 categories with similar mortality rates, resulting in the low-, medium- and high-risk groupings depicted in *Figure 19* (which was developed on 70% of the development set and validated on the remaining 30%).

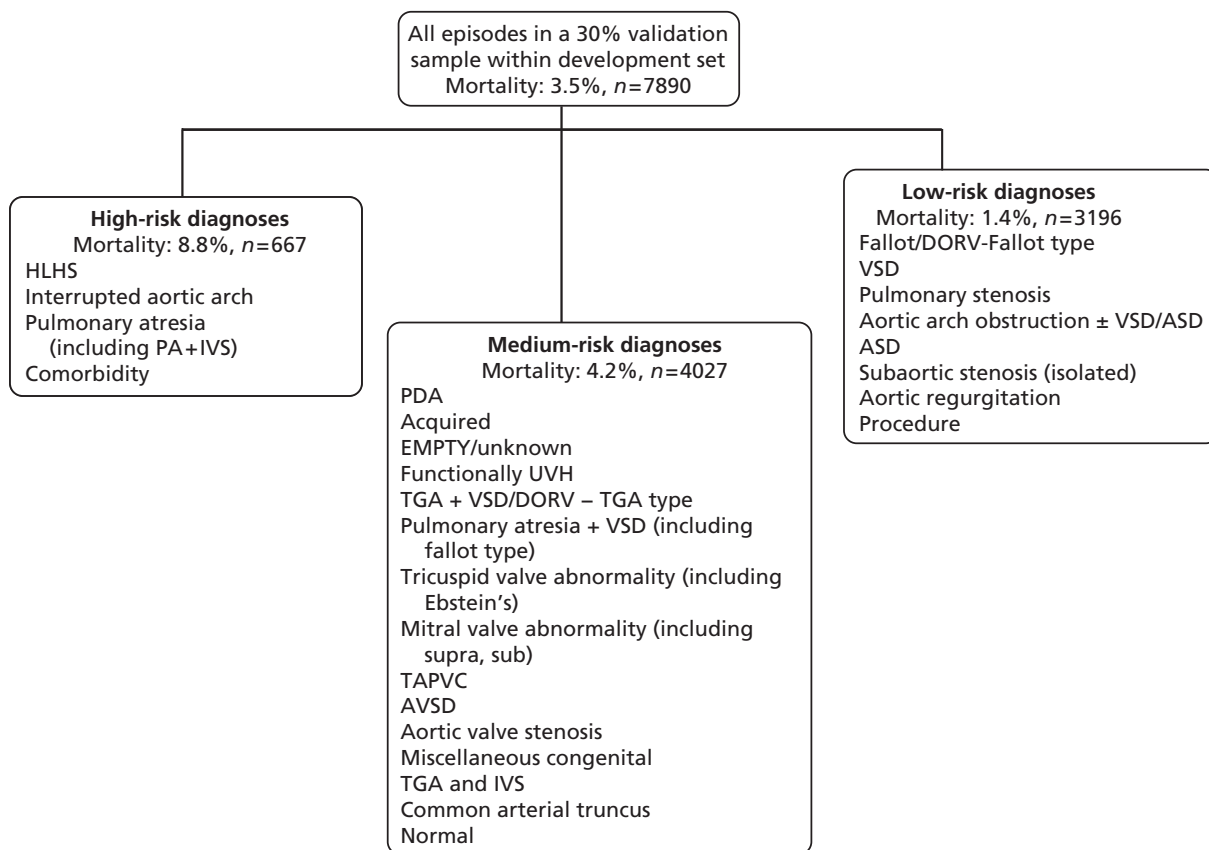
For diagnosis scheme 2, only the UVH attribute consistently remained significant in the multivariate logistic regression analyses.

## **Collapsing comorbidity information**

The sparseness of comorbidity information within the CCAD data set results in low volumes of episodes associated with the four comorbidity categories, even if the comorbidity information derived from both the comorbidity and the diagnosis data fields is pooled. The stability of the model improved notably when, rather than including indicator predictor variables for all four comorbidity groups, we instead grouped the non-Down syndrome comorbidities to produce a single binary indicator for the presence (or absence) of non-Down syndrome comorbidity information for an episode. We note that univariate analysis indicated that the presence of an acquired, premature or congenital non-Down syndrome comorbidity had an adverse effect on 30-day outcome, whereas Down syndrome seemed to have a protective effect (although this was not statistically significant).

An additional consideration in favour of including this single comorbidity indicator in the model is the potential for its presence to drive up data quality prospectively (the comorbidity field has poor completeness, particularly pre 2005).





**FIGURE 19** Grouping of diagnostic categories into low, medium and high risk. ASD, atrial septal defect; DORV, double-outlet right ventricle; IVS, intact ventricular septum; PA, pulmonary atresia; TAPVC, totally anomalous pulmonary venous connection.

## Age and weight

The association between age and mortality is non-linear and defied transformation. We therefore pragmatically chose to include both continuous age and age band [neonate, infant, child (> 1 year)] in the model. The last age, i.e. child (> 1 year), has strong clinical face validity. We explored the possibility of developing a separate model for each age band, but decided that there were insufficient events among children to support the degrees of freedom of specific procedures (even with the low-volume specific procedure group).

We also include continuous age within the model and hence have a stepped age/risk relationship with the same gradient in each of the age bands. We investigated a piecewise linear age/risk relationship with different slopes in each band but this led to more instability in the parameter values across 50% runs within the development set and a higher Hosmer–Lemeshow chi-squared statistic and so was not adopted.

Continuous weight and weight-for-age z-score were also considered as potential risk factors in the model, with the weight-for-age z-score using the development set as a reference population (split into 23 age bands, narrower at young age). In multivariate analyses, weight and weight-for-age z-score were not typically both significant, depending on which other variables were in the model. Of the two variables, continuous weight was thought to be the easier to implement in terms of prospective use, as z-score would require a reference population.

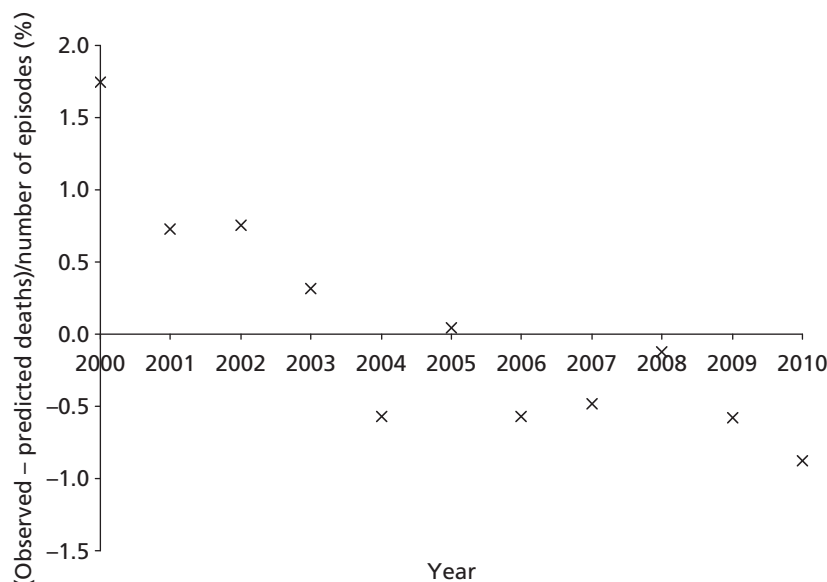
## Low-volume specific procedures

As far as possible we wanted to maintain the specific procedures; however, to develop a robust model we needed to reduce the number of predictor variables associated with procedural information. We therefore decided to group the nine specific procedures with the lowest volumes (all of which have < 70 episodes over the entire development set) into a 'low-volume' specific procedure group. This low-volume group consists of the following specific procedures: aortic root replacement (not Ross); aortopulmonary window repair; AVSD and tetralogy repair; cor triatriatum repair; multiple VSD closure; Senning or Mustard procedure; tetralogy with absent pulmonary valve repair; tricuspid valve replacement; and truncus and interruption repair.

## Year of surgery

We looked at the performance of our initial risk model in each year from 2000 to 2010 and found a clear, but not unexpected,<sup>11</sup> trend of underestimating overall risk in each year to 2005 (by less in later epochs) and increasingly overestimating in the subsequent years. This trend suggests that any model calibrated on data from this entire period would already be out of date, and also highlights the importance for timely recalibration of any risk model prospectively. We considered calibrating the model on a subset of later years in the development set but found that the loss in volume of data led to model instability. To include year of surgery as a variable would be to assume that improvement over time is inevitable and ignores ceiling and threshold effects. We therefore decided to include a variable in the model to indicate whether an episode occurred pre 2007 or from 2007 onwards. Although there is no clinical mechanism for such a risk factor, it enables the entire development set to be used in the model calibration while also ensuring that the calibration is more up to date and fit for purpose prospectively. The inclusion of this indicator variable also highlights the need for a rolling programme of recalibration for a model in routine use.

After finalising the model we reran it in the development set without the variable indicating whether an episode occurred pre 2007 or from 2007 onwards. *Figure 20* shows the difference between observed and predicted deaths from this reparameterised model evaluated in the development set, for each year, weighted by the number of episodes in that year. This indicates a trend of improving risk-adjusted outcomes over time.



**FIGURE 20** Difference between observed and predicted deaths for each year, weighted by the number of episodes in each year, evaluated in the development set using a reparameterised model that excludes the 2007 indicator variable.

## Final parameter selection and model performance in the development set

Based on the analysis and considerations set out above, the final model, decided on jointly by the clinical and analytical teams, was a logistic regression model with the following variables:

- age (both as a continuous measure and as neonate/infant/child bands)
- weight (as a continuous measure)
- specific procedure (one of 27 CCAD groups, 'no specific procedure' or 'low-volume specific procedure')
- procedure type (bypass or non-bypass)
- broad diagnosis group (low-, medium- or high-risk group)
- UVH attribute (indicator variable)
- presence/absence of a recorded non-Down syndrome comorbidity
- indicator variable for whether an episode occurred pre 2007 or from 2007 onwards.

The 'enter' regression model was used to parameterise this model across the entire development data set. The AUC for the model in the development set was 0.78 (95% CI 0.76 to 0.79) and the Hosmer–Lemeshow chi-squared statistic was 9.2 ( $p=0.325$ ), indicating no statistically significant differences between observed and expected number of deaths when calculated in deciles of predicted risk (see Harrell<sup>50</sup> for more information on these measures). We note that data completeness was reasonable for all variables included in the final model.

Although we gave considerable thought to the interaction of age and weight, we did not perform a detailed study of interactions between specific procedures and diagnostic classifications. Consideration of such interactions might have improved model performance in the development set, but we wanted to avoid introducing too many additional degrees of freedom. The characteristics of the model were considered acceptable to the project team, noting that model choice was not a purely statistical consideration and that consideration of uptake by CCAD and centres supported the use of as many specific procedure groups as possible and the opportunity presented by the risk model to drive improvements in completeness and data quality concerning diagnosis and comorbidity entered into our thinking. Details of the model are provided in *Appendix 4*.



## Chapter 8 Model performance

In the test set it was possible to calculate a risk score in 95% of cases: *Figure 21* shows the episodes that were excluded from analysis. Weight was missing in 539 episodes (4.9%) and age in 23 episodes (0.2%). There were also 392 (3.4%) episodes without diagnostic information, but these were included in the analysis within the diagnostic category 'empty/missing diagnosis'. Additionally, the 30-day life status was 'unknown' for 226 (2.0%) episodes in the test set.

### Global performance in the test set

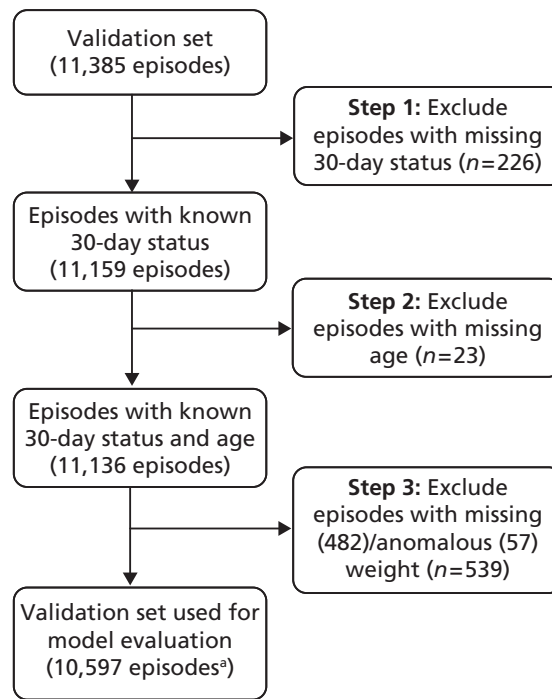
The performance of the model described in *Chapter 7* (see *Final parameter selection and model performance in the development set*) across the spectrum of predicted risk is shown in *Figure 22*. The AUC for this model was calculated as 0.76 (95% CI 0.74 to 0.79) within the test set, compared with 0.78 (95% CI 0.76 to 0.79) in the development set. The Hosmer–Lemeshow chi-squared statistic was 25.1 ( $p=0.002$ ), compared with 9.2 in the development set. The poor Hosmer–Lemeshow chi-squared statistic in the test set is due to an excess of deaths in the first (6 observed vs 1.6 predicted) and fourth (22 observed vs 13 predicted) deciles of risk. We note that there were three deaths from Rastelli procedures in the test set (out of 62 procedures), all in the first decile of risk, compared with no deaths from 125 Rastelli procedures in the development set. Overall, there were 335 observed deaths in the test set compared with 329 predicted.

In exploring the performance of this risk model and features of the case mix across different institutions, pre and post 2007, we identified an error in the data provided to us. Essentially, the algorithm employed by CCAD to determine the 'specific procedure' classification for each record had not been implemented for a tranche of records deriving from one centre and covering the period 2007 onwards, leaving 1474 episodes (4% of episodes over both sets) with an erroneous classification of 'not a specific procedure'. This arose as the centre concerned had resubmitted these data en bloc in 2010 and CCAD had not rerun the algorithm.

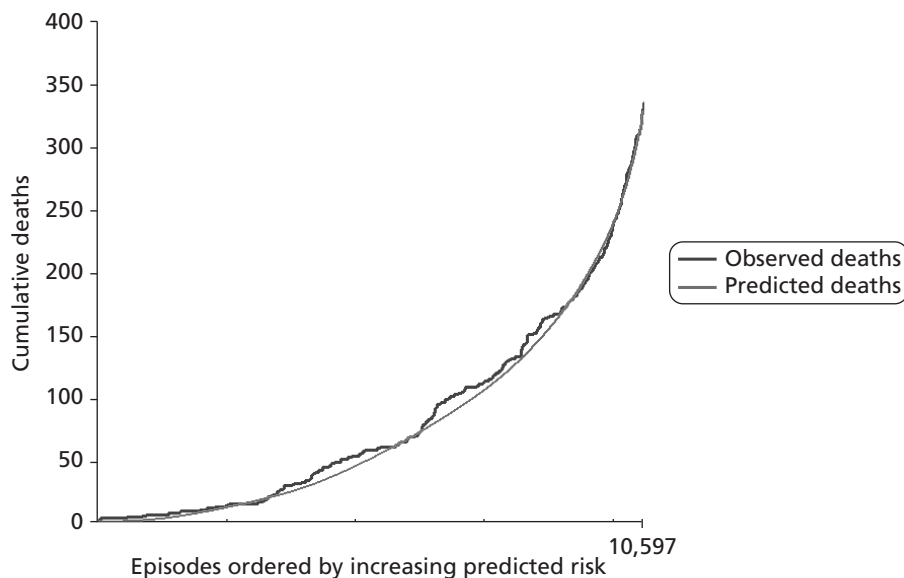
Given that we had already 'opened' the test set of data at this point, we could not revisit any of the analysis that led to the choice of factors to include in the logistic regression model. However, we decided that, if CCAD were able to provide us with corrected data (which they were), we would recalibrate the model on the corrected development data set. In addition to updating the specific procedure information for affected episodes, we reran the process of using specific procedure to identify episodes in which the patient had a functionally UVH but in which this was not recorded in the diagnostic fields [see *Chapter 5, Diagnosis classification scheme 2 (mapping International Paediatric Congenital and Cardiac codes to diagnostic groups defined by ventricle number and aortic obstruction)*]. The specific procedure field was also corrected for those episodes in the test data set that were affected.

From this point of the report, the performance of the model in the test set of data is presented following the recalibration of the model on the corrected development data. As the intent of this project is to provide units with a way of prospectively monitoring risk-adjusted outcomes, there is a focus on features of model performance in test data from 2007 (see *Chapter 7, Year of surgery*).

*Figure 23* shows the performance of the recalibrated model across the spectrum of predicted risk when assessed among the corrected test data from all years. *Figure 24* shows the receiver operating characteristic curve for this model in the test set across all years, for which the AUC is 0.77 (95% CI 0.74 to 0.79), similar to the value calculated for the development set (0.78, 95% CI 0.76 to 0.79). This confirms the good discrimination between high- and low-risk cases indicated by the appearance of the MADCAP curve, with a shallow climb of the stepped line indicating cumulative observed deaths at low predicted risk and a steeper climb at high predicted risk. The Hosmer–Lemeshow chi-squared statistic is 22.7, indicating that the discrepancies between observed and predicted mortality in deciles of predicted risk are statistically significant ( $p=0.004$ ). These discrepancies are evident in the portions of the MADCAP chart where the stepped line indicating cumulative observed deaths climbs at a higher or lower rate than the smooth line indicating



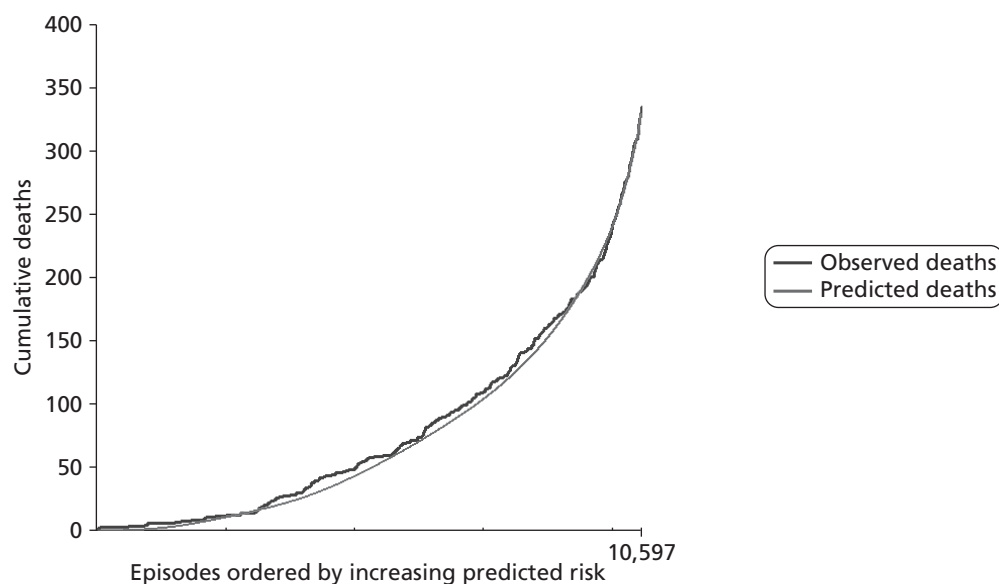
**FIGURE 21** The number of episodes with missing 30-day status and/or patient age and/or missing/anomalous patient weight that were removed from the test set (see text). a, Includes episodes with missing diagnostic information (n=392).



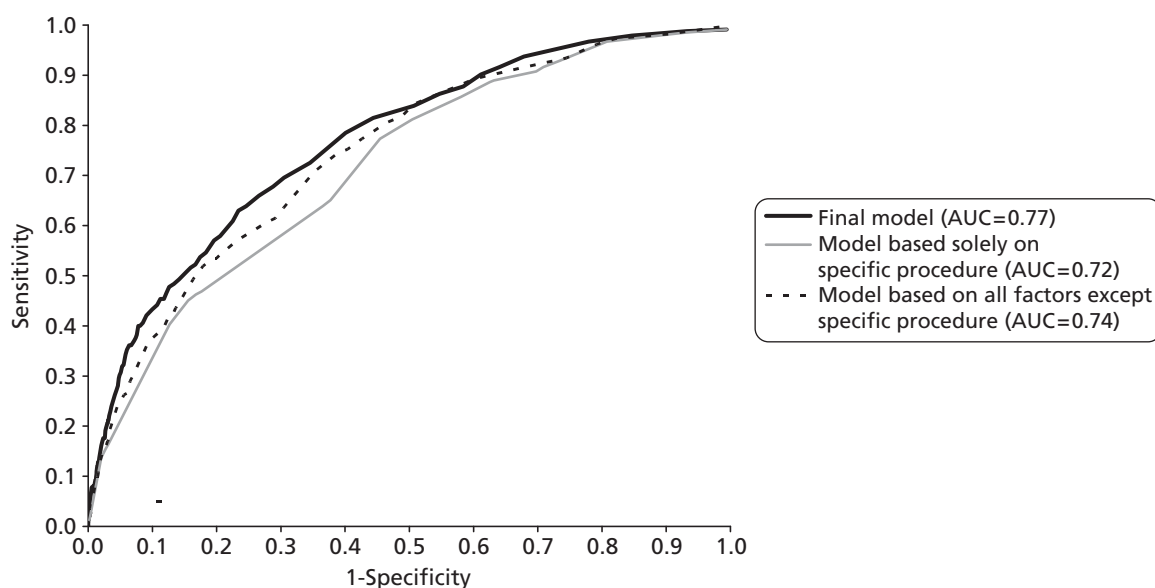
**FIGURE 22** Cumulative deaths in the test set plotted against episode number with episodes ordered by increasing risk, as predicted using the model set out in Chapter 7, *Final parameter selection and model performance in the development set* (n=10,597). Note that some data were corrected subsequent to this analysis (see text below and Figure 23).

cumulative predicted deaths. The extent to which these features of model performance affect the utility of the model in prospective use is discussed in *Institutional case mix and model performance* and Chapter 9, *Model limitations and cautions*. The overall number of observed deaths was 335 compared with 329.3 predicted.

In terms of the added value of including age, diagnosis, comorbidity and use of bypass, it is interesting to note that the performance of a model based solely on a specific procedure (AUC 0.72, 95% CI 0.69



**FIGURE 23** Cumulative deaths among the corrected test data from all years plotted against episode number with episodes ordered by increasing risk as predicted using the model set out in *Chapter 7, Final parameter selection and model performance in the development set* ( $n=10,597$ ).

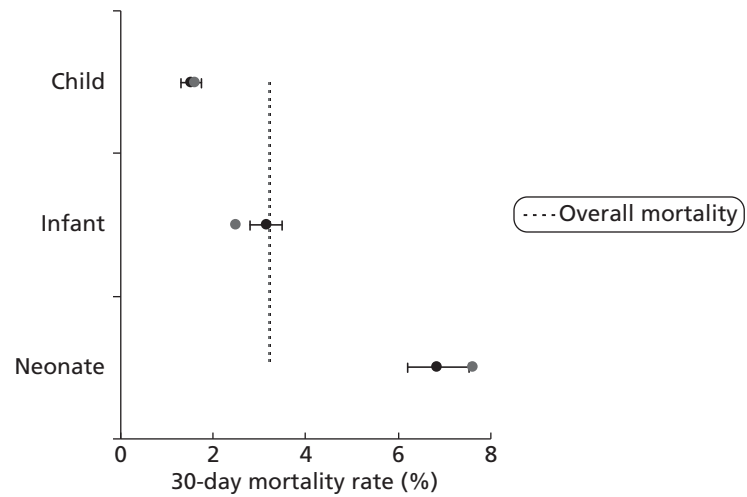


**FIGURE 24** Receiver operating characteristic curve for three models evaluated in the test set: the final risk model (AUC=0.77); a model based solely on a specific procedure (AUC=0.72); and a model based on all factors in the final model except a specific procedure (AUC=0.74) ( $n=10,597$ ).

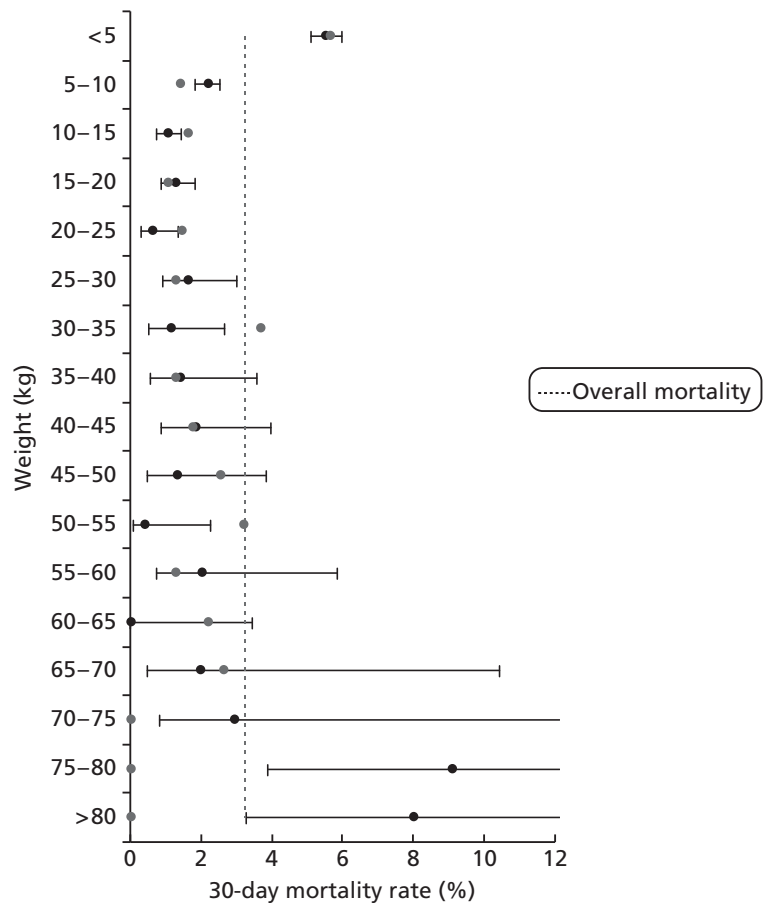
to 0.75) is worse than the performance of a model based on all factors other than a specific procedure (AUC 0.74, 95% CI 0.72 to 0.77) (see *Figure 24*).

### Features of the test set compared with the development set

*Figures 25–28* and *Table 8* show a comparison between the development and the test data sets in terms of the episode-level mortality rates associated with the factors included in the risk model. These show broad similarity in the univariate associations between individual factors and 30-day mortality in the two data sets as would be expected given the random selection process used. That said, there are some features worth noting, particularly the disparity in 30-day mortality rates associated with some individual specific procedure classifications and the higher 30-day mortality rate among neonates.

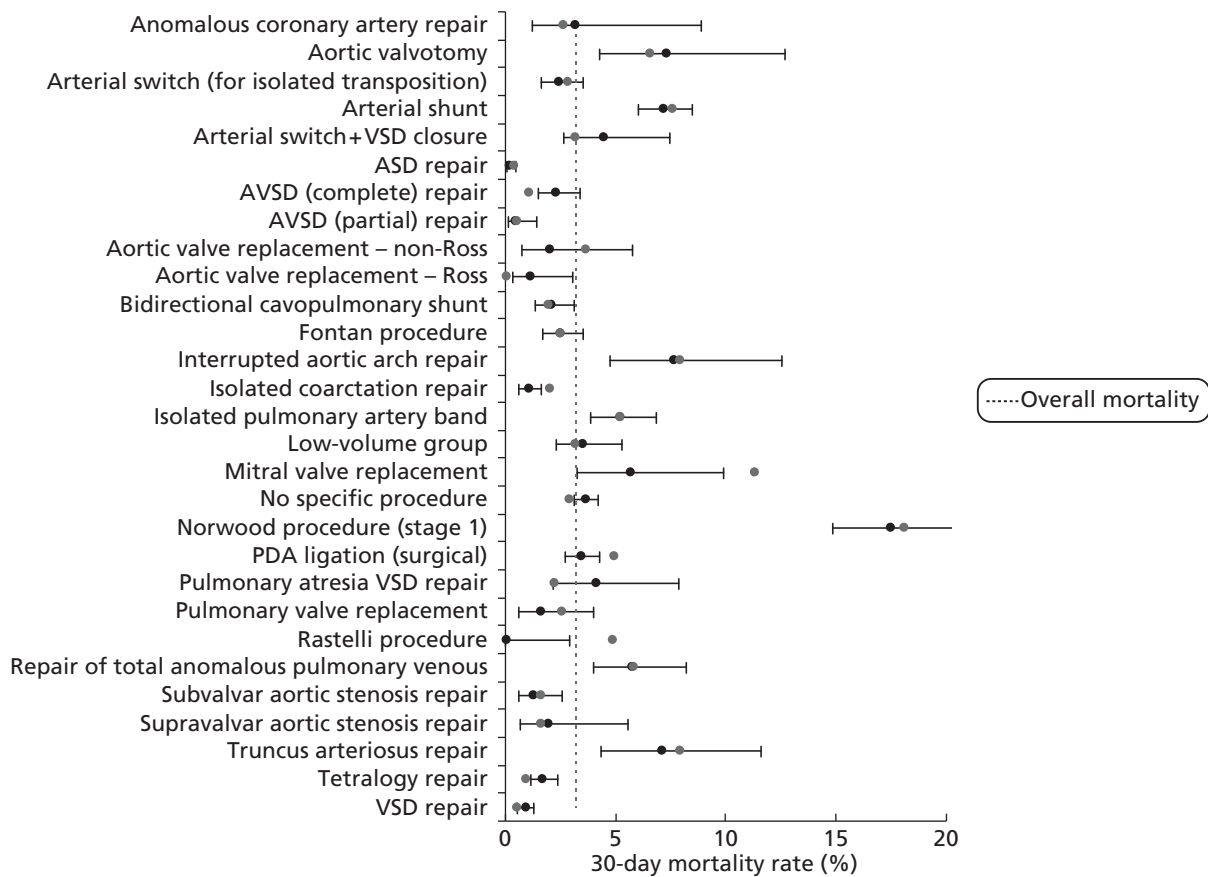


**FIGURE 25** Observed mortality rate by age band in the development set (black,  $n=26,447$ ) and the test set (grey,  $n=10,597$ ).

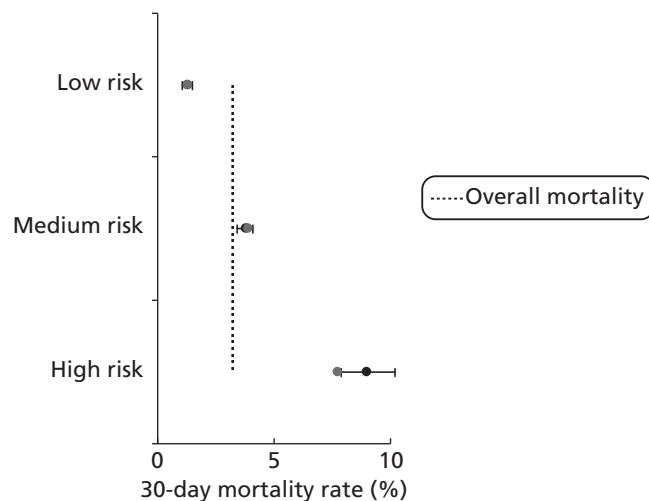


**FIGURE 26** Observed mortality rate by weight category in the development set (black,  $n=26,447$ ) and the test set (grey,  $n=10,597$ ).





**FIGURE 27** Observed mortality for specific procedures in the development set (black,  $n=26,447$ ) and the test set (grey,  $n=10,597$ ). ASD, atrial septal defect.



**FIGURE 28** Observed mortality by diagnosis risk group in the development set (black,  $n=26,447$ ) and the test set (grey,  $n=10,597$ ).

**TABLE 8** Mortality rates based on 30-day status for UVH attribute, procedure type, presence of non-Down syndrome comorbidity and epoch, taken from the development ( $n=26,447$ ) and test ( $n=10,597$ ) sets at episode level

	Development set 30-day mortality rate (%)	Development set 95% CI <sup>a</sup> (%)	Test set 30-day mortality rate (%)
Not UVH	2.6	2.4 to 2.8	2.5
UVH	6.8	6.1 to 7.7	6.8
Bypass	3.1	2.9 to 3.4	3.0
Non-bypass	3.5	3.1 to 4.0	3.6
No (non-Down syndrome) comorbidities	2.9	2.7 to 3.2	2.9
At least one (non-Down syndrome) comorbidity	5.5	4.7 to 6.4	4.9
Pre-2007 procedures	3.4	3.2 to 3.7	3.1
Procedures from 2007 onwards	2.9	2.5 to 3.2	3.3

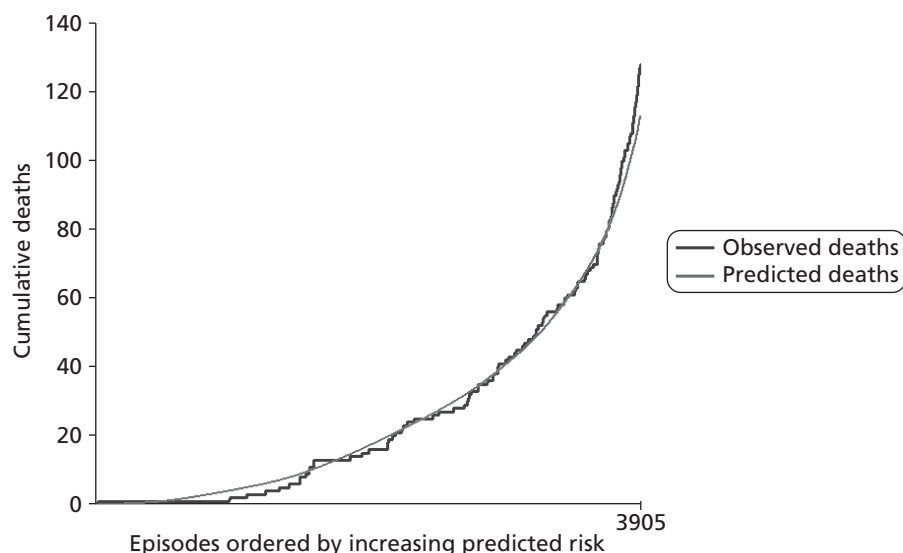
a CIs were calculated using Wilson’s approximation.

### Global performance of model from 2007 onwards

The risk model is intended for future use in routine monitoring. Given this and given the observed improvement in outcomes over time and our adjustment for this fact in the model (see *Chapter 7, Final parameter selection and model performance in the development set*), it is the performance of the model in the test set episodes that occurred during or after 2007 that is most informative concerning its fitness for purpose.

The MADCAP chart for these test episodes is shown in *Figure 29*.

Although the model has a higher value for the AUC among data from 2007 onwards than overall (0.81 vs 0.78), and so shows better discrimination among these episodes, the model underestimates risk at the very



**FIGURE 29** The performance of the model in the test set for all episodes occurring after 1 January 2007 ( $n=3905$ ).

high-risk end of the spectrum of predicted risk (very right-hand side of *Figure 29*) and is, as a result, less accurate overall in these more recent data than in the full development set. The overall number of observed deaths is 128 compared with 113 predicted. We considered the potential impact of Copas shrinkage on the performance of the model in the test set post 2007. We observe an underprediction of mortality at high risk and an overprediction at low risk, which is the opposite of that expected from Copas. Additionally, the inaccuracies observed are present only in the recent (post 2007) cohort rather than across the entire test set. Thus, we believe that the differing performance of the model in the test and development sets is driven by other effects, such as the markedly higher mortality rate among neonates in the randomly selected validation set (see *Figure 25*).

To explore the performance of the model in recent data in more detail, we plotted MADCAP charts in a number of important subgroups. We show here those we consider most relevant. *Figure 30* shows the performance of the model in neonates, infants and children, which shows that the model generally underestimates risk in the neonatal group. It is worth remembering here that the mortality rate in neonates was markedly higher in the test set than in the development set (see *Figure 25*). This artefact of the random selection of development and test sets underlines the value of having a quarantined test set to mimic as closely as possible prospective use.

The performance of the model within four mutually exclusive diagnostic groups is shown in *Figure 31*. There is underestimation of risk in higher-risk UVH patients, likely to correspond to neonates. That aside, the generally good performance across these groups is encouraging.

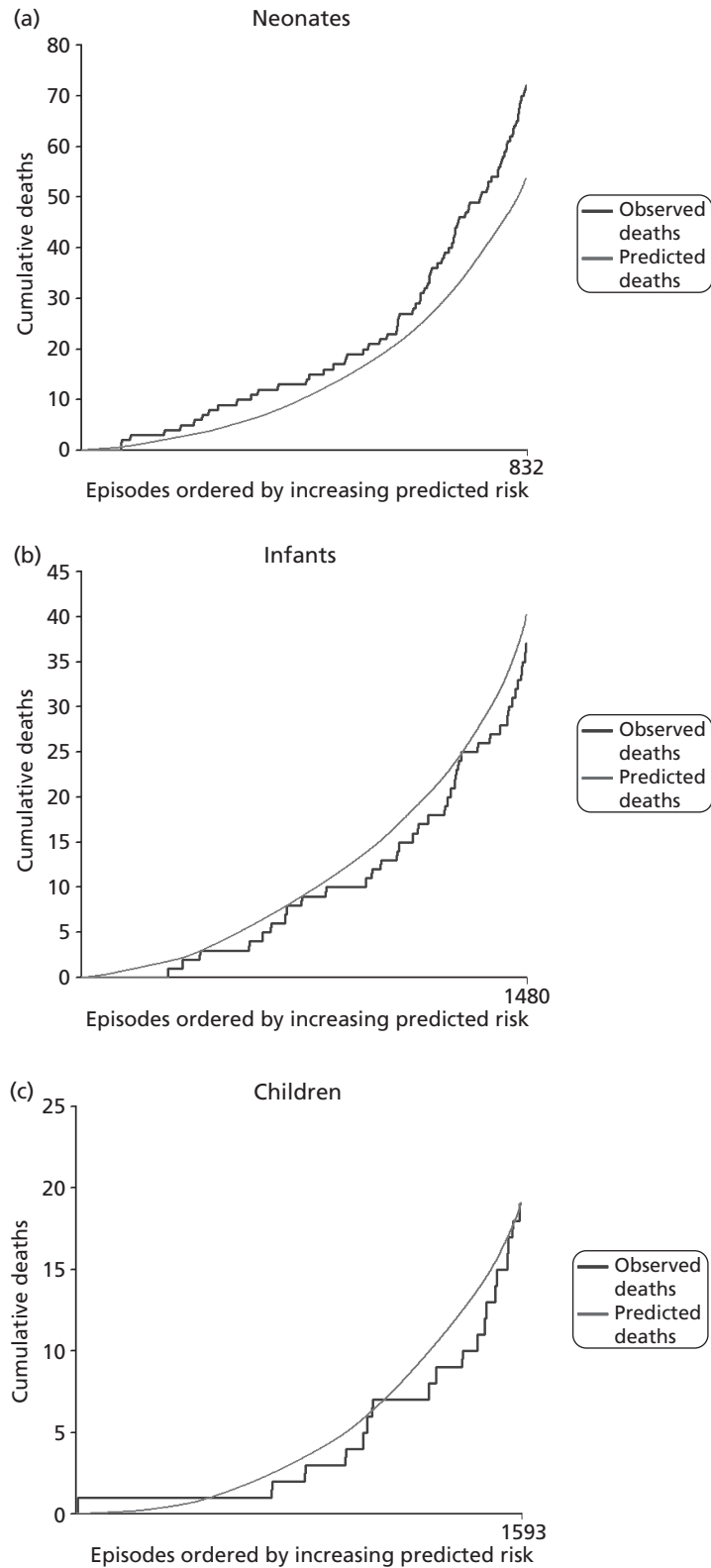
*Figure 32* shows the performance by procedure type: bypass or non-bypass. The model performs well in both subgroups, but shows more discrimination in the bypass procedures.

### Institutional case mix and model performance

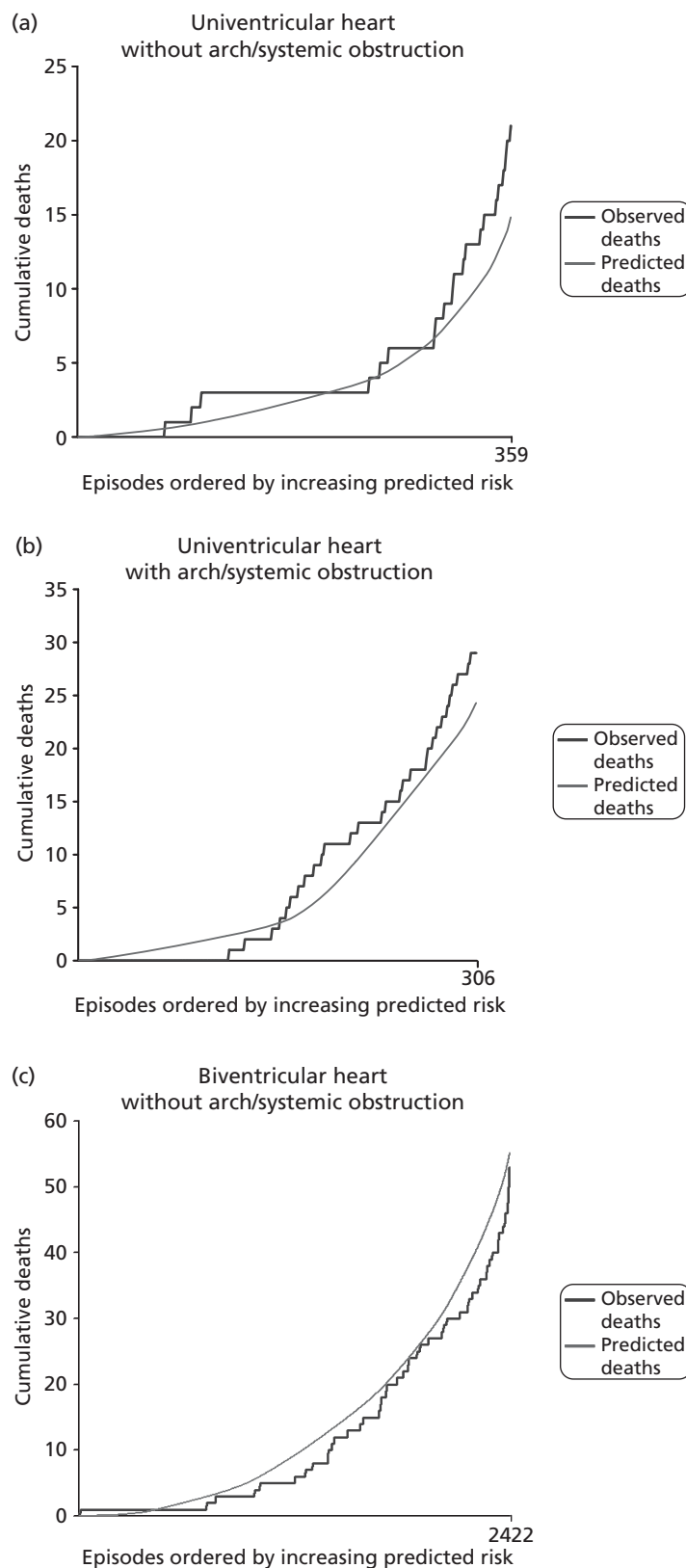
In addition to understanding the performance of the model across the spectrum of predicted risk and within key subgroups of patients, to assess the fitness for purpose of the model it is important to understand the distribution of case mix and, in particular, differences in case mix between institutions. Essentially, it is important to understand whether or not and under what circumstances differences in case mix and the differential performance of the risk model in different subgroups would combine to give an artefactual impression of better or worse risk-adjusted outcomes at one unit compared with another.

*Figure 33* shows the cumulative distributions of predicted risk within the development and test sets after 1 January 2007. The distribution of predicted risk in the two data sets is very similar and shows that just over 30% of episodes have a predicted risk of 30-day mortality of  $\leq 1\%$  and 80% have a predicted risk of  $\leq 4\%$ , but that around 5% of episodes have a predicted risk of  $> 10\%$ .

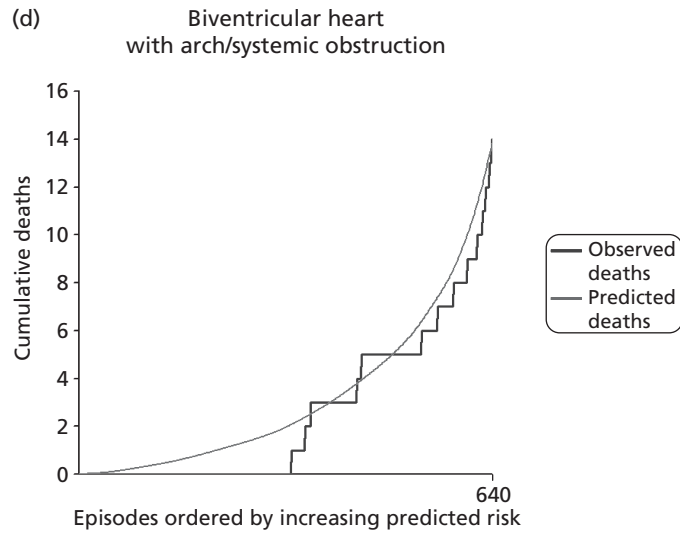
The differences in case mix between centres, in terms of predicted 30-day mortality risk, using the model across the development and test sets after 1 January 2007, are shown in *Figure 34*. Note that the institution numbers given in this plot are not in the same order as the institution letters in *Chapter 4*. Each line shows the case mix for a single institution. The steeper lines indicate institutions with a higher proportion of relatively low-risk episodes, as estimated by the model. Many of the institutions have quite similar profiles of predicted risk, but there are some marked differences. For instance, at institution 5, 12% of episodes have a predicted risk of  $> 10\%$ , whereas at institution 7 this proportion is  $< 1\%$ .



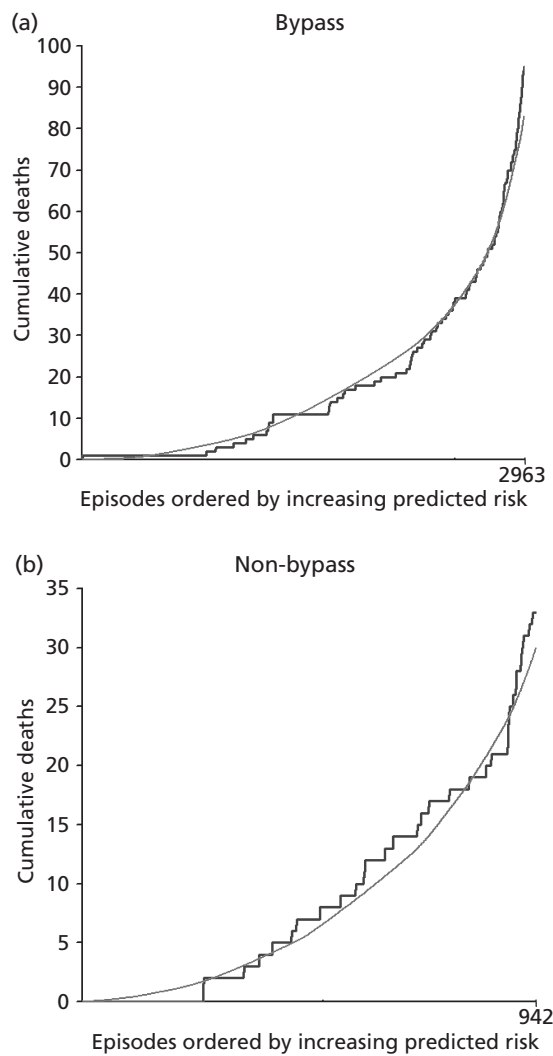
**FIGURE 30** The performance of the model in age band subgroups in the test set for data after 2007 ( $n=832$  for neonates;  $n=1480$  for infants;  $n=1593$  for children).



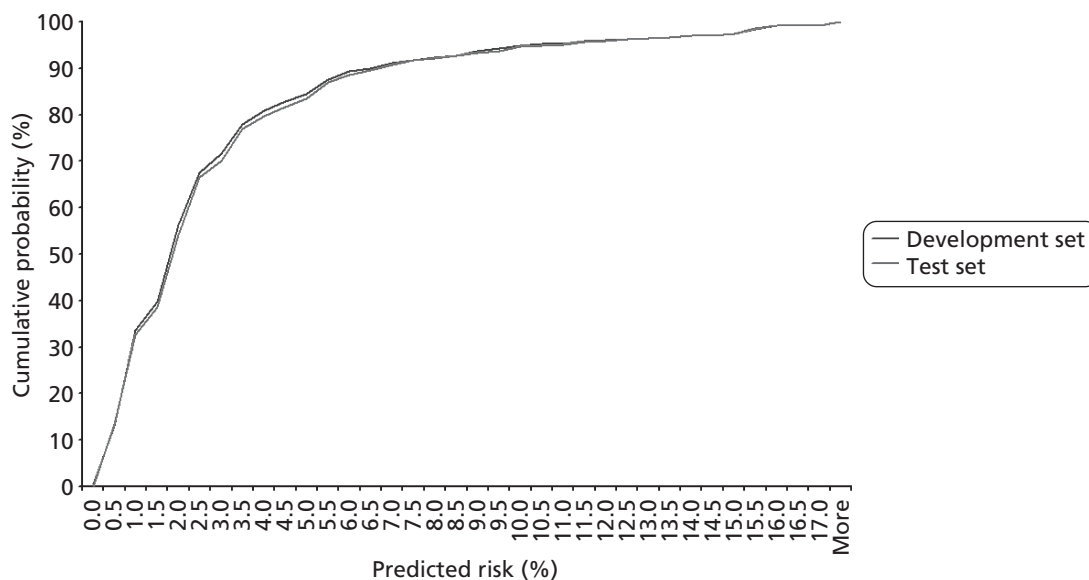
**FIGURE 31** The performance of the model in diagnostic group subgroups in the test set for data after 2007 ( $n=359$  for UVH without arch/systemic obstruction;  $n=306$  for UVH with arch/systemic obstruction;  $n=2422$  for BVH without arch/systemic obstruction; and  $n=640$  for BVH with arch/systemic obstruction).



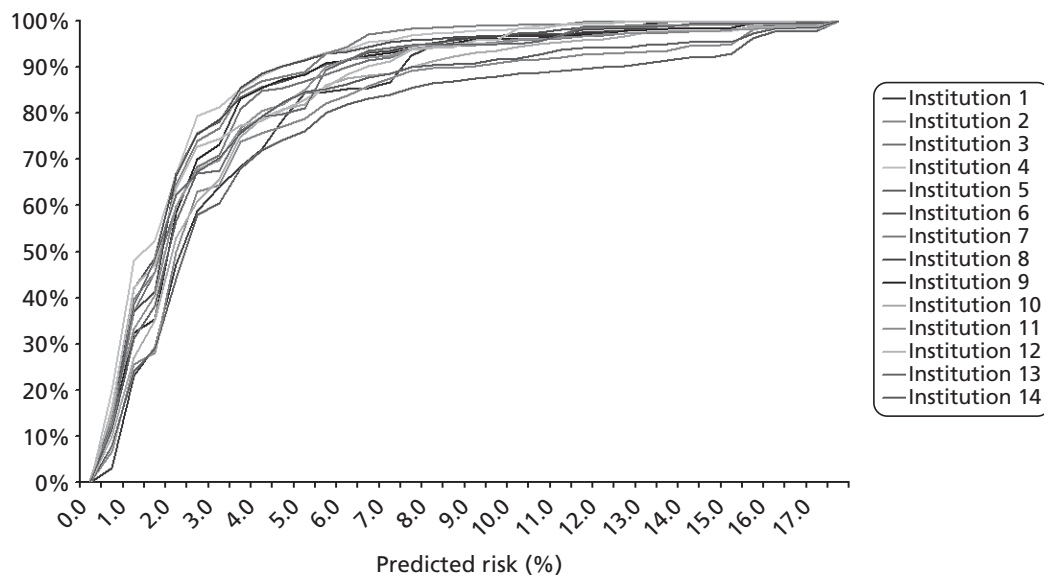
**FIGURE 31** The performance of the model in diagnostic group subgroups in the test set for data after 2007 ( $n=359$  for UVH without arch/systemic obstruction;  $n=306$  for UVH with arch/systemic obstruction;  $n=2422$  for BVH without arch/systemic obstruction; and  $n=640$  for BVH with arch/systemic obstruction) (*continued*).



**FIGURE 32** The performance of the model in procedure type subgroups in the test set for data after 2007 ( $n=2963$  for bypass and  $n=942$  for non-bypass).



**FIGURE 33** Comparison of the case mix as determined by predicted risk between the development ( $n=26,447$ ) and the test ( $n=10,597$ ) sets.



**FIGURE 34** Comparison of the case mix as determined by predicted risk between the different institutions. There are four institutions within the data set that had a very low number of procedures (non-specialist centres that happened to perform a few procedures which were eligible for paediatric CCAD submission). Given the low numbers, these institutions are not included.

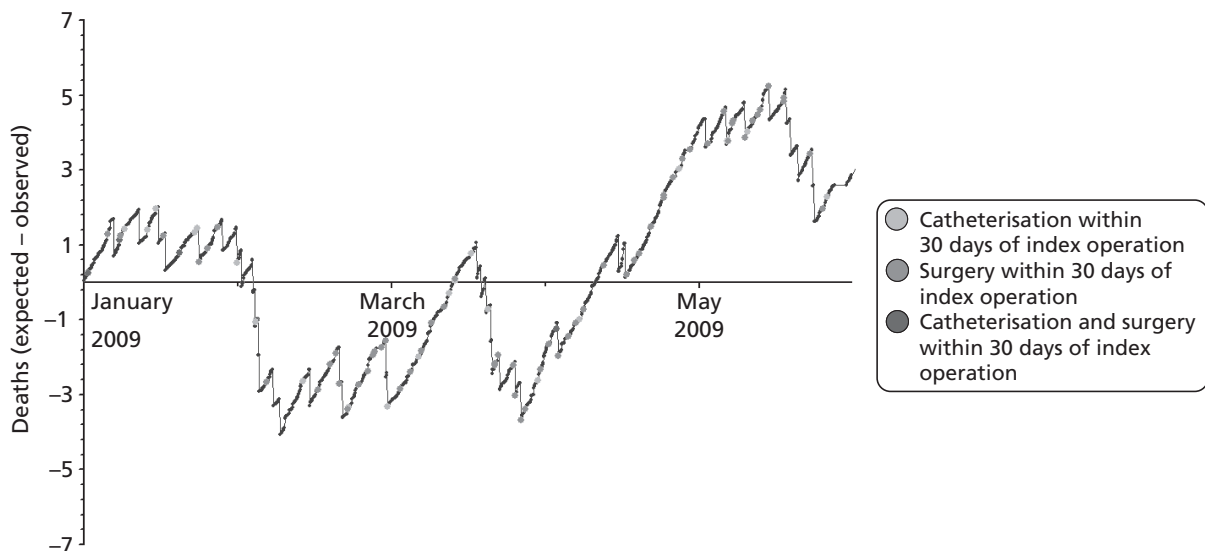
Given that the model underestimated risk at the very high-risk end of the recent (since 2007) test data, risk adjustment using the model as currently parameterised will potentially give an unfair assessment of outcomes at those centres with a high proportion of high-risk cases. This is an important caveat to the interpretation of risk-adjusted outcomes within and across institutions that will need to be discussed with CCAD and the institutions as this work is taken forward.

## Proposed monitoring tool

Figure 35 shows an example of the graphical monitoring tool that we envisage the risk model feeding into. The figure shows a standard VLAD chart of the difference between the expected number of deaths over a series of episodes, determined by summing the estimated risk for each episode, and the observed number of deaths. The trace rises with each survival (more so for high-risk episodes) and falls with each death (more so for low-risk episodes). The data shown here are national data for consecutive cases during the first 5 months of 2009. It should be noted that, as we had access only to year and month of operation, the date of cases within each month has been generated arbitrarily.

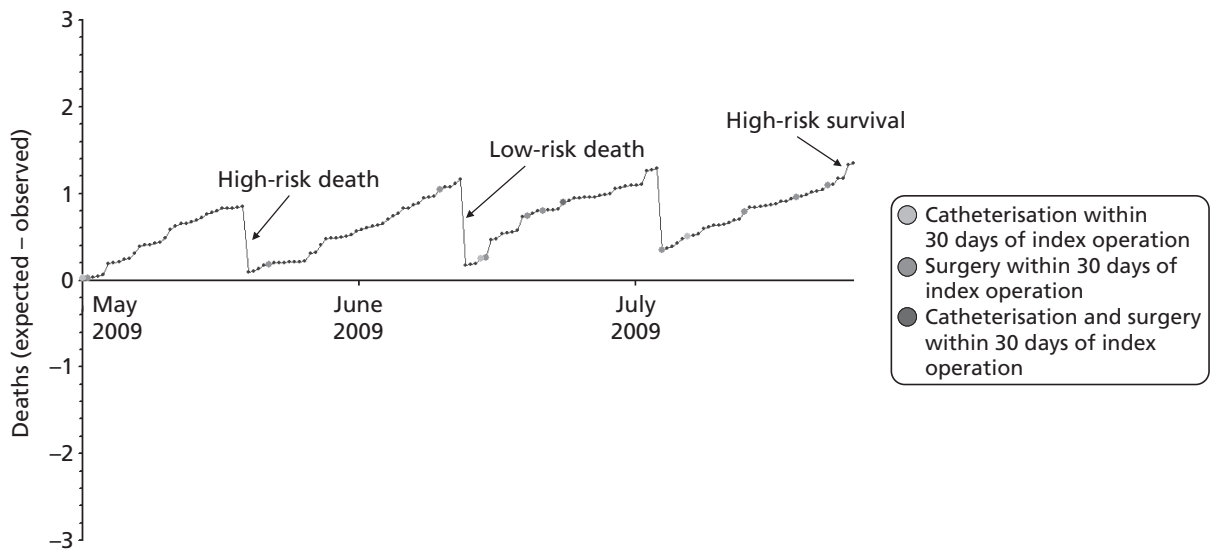
Shown along with the running tally of '(expected – observed) deaths' are episodes in which there was a surgical or catheter intervention within 30 days. This is added to discern between periods in which equivalent 30-day survival is achieved but with a higher reintervention rate and so arguably a higher level of postoperative morbidity.

Figure 36 shows a 'mock-up' of how the monitoring tool would be used at an institutional level to review risk-adjusted mortality and reinterventions on a regular basis.



**FIGURE 35** An example VLAD chart for 5 months of national data in 2009.





**FIGURE 36** An example VLAD chart for how the monitoring tool might be used within institutions on a rolling quarterly basis using a 'mocked-up' institution assuming a case load of 600 cases per year.



## Chapter 9 Discussion

### Model overview and attributes

We have developed a risk model for use in monitoring 30-day mortality in paediatric 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 an AUC of 0.77 when evaluated across the entire validation data set and an AUC of 0.81 for post-2007 data. This compares well with other published risk-adjustment tools from the same field of practice.<sup>24</sup> As we have shown, supplementing procedural information with diagnostic, age, weight and comorbidity characteristics increased the discriminatory performance of the risk model: the AUC across the entire validation data set was 0.72 when only specific procedure was included. The risk model we have developed is an empirical (such as the STS-EACTS score) rather than a consensus-based risk-adjustment tool (such as RACHS-1 or Aristotle). This offers the advantage of reflecting the data rather than subjective personal opinion. Developed using the UK national database, the model can be applied objectively to data already routinely collected by cardiac paediatric units. We believe that the risk model is fit for the purpose of routine monitoring of outcomes in paediatric cardiac surgery, and we have gone some way to disseminating our findings by meeting with key stakeholders to inform them of the model and its potential use for quality assurance.

It has been previously observed that procedure categories developed for use in risk adjustment may have incomplete coverage, leaving some operations excluded from outcome analyses.<sup>23–25</sup> The empirically based tool for analysing mortality published by STS and EACTS in 2009<sup>24</sup> increased the coverage of records by including 148 types of operation and using a Bayesian model to adjust for small denominators. As discussed in the methods, the specific procedure categories reported by CCAD online<sup>8</sup> are well established and accepted within the UK for benchmarking: these cover 83% of records in the data set. We therefore consider that these are the most appropriate markers of procedural complexity for a risk model designed for routine use in a UK context. Indeed, all choices of model inputs were made on the basis of their being both readily available and easily understood by clinicians 'on the ground'. We gathered additional information for use in mortality predictions by ascertaining cardiac diagnosis, which could be allocated to 97.1% of those records classed as 'not a specific procedure', the most common diagnosis (11.6%) in these records being 'acquired'. Therefore, the coverage of the data set in terms of attributed clinical information was very high, comparing well with previous attempts at risk adjustment in this context.

### Model limitations and cautions

The necessary pseudonymisation process had the negative effect of essentially blinding the analytical team to missing or anomalous data in the patient and hospital ID fields. This meant that we had to conduct extensive face validity exercises when trying to identify genuine groups of records pertaining to the same patient. This included analysis to check the internal consistency of patient histories in terms of the dates of operations and ages at each operation and also some manual checks of the clinical face validity of sequences of operations. We cannot be certain that the study team identified all incorrectly entered IDs. The close collaboration between the clinical and the analytical teams was essential here, as was having the involvement of CCAD, which could, when necessary, go back to CCAD'S own records for clarification. We note that using any audit database for research purposes requires some cleaning of the data and that collaboration with the CCAD team was essential for understanding the data and performing appropriate cleaning. As part of this ongoing collaboration, the analytical team shared the data cleaning procedures used with the CCAD team.

An important yet obvious caveat of this work is that the model pertains to short-term 30-day outcomes (not long-term outcomes) and is designed for the purpose of routine monitoring for quality assurance rather than bedside-type predictions for individual patients. The unit of analysis is a 30-day episode so that neither deaths nor survivors are counted twice, and this is a caution to be aware of when calculating mortality percentages. We note that 4.5% of episodes contain at least one surgical reintervention and 1.8% contain at least one catheter reintervention within 30 days in the development set.

The risk model is intended for future use in routine monitoring of risk-adjusted outcomes within UK paediatric cardiac centres. 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, in particular among neonates. This indicates that risk adjustment based on the current parameterisation of the model will potentially give an unduly negative impression of outcomes at those centres with a high proportion of high-risk cases. This is an important caveat to interpretation of risk-adjusted outcomes within and between centres that will need to be considered as the work is taken forward. Options that will be considered in our future work with CCAD and the centres include recalibration of the model over both development and test sets and then testing on data accrued since our project started and (in the longer term) reparameterisation using these new data to, potentially, improve stratification at high risk (for instance, specific comorbidities). Importantly, in our work with centres we will ensure that the case mix distribution is available to individual centres alongside risk-adjusted outcomes so that centres with a higher-risk case mix are openly identified and thus all are prewarned as to this issue.

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 centre 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 UK is currently the only country that displays unit-specific outcomes of procedures online,<sup>8</sup> there has been considerable debate of this issue in the professional journals, with the suggestion that programme-level reporting of unit-specific outcomes across a range of domains may evolve internationally over the coming years.<sup>51,52</sup> Centre effects were purposefully left out of the model. Our principal aim was to develop, for prospective use, a risk-adjustment system to account for patient-level factors, rather than create the most accurate statistical model of historical 30-day mortality. Importantly, it would not be possible from analysis of data alone to tease out whether any centre-to-centre differences in mortality rate are due to differences in case mix not captured by the data set, or to genuine, historical differences in the care processes delivered at surgical units. We are currently exploring the robustness and utility of the model across centres by working directly with centres and CCAD in a pilot study. Given some of the sensitivities that surround centre and 'volume' effects, we feel that routine monitoring of outcomes should start locally within individual clinical teams.

There is a need for a rolling programme of reparameterisation for a model in routine use to account for anticipated improvements in outcomes over time<sup>11</sup> and indeed any other evolving trends. Alongside a reparameterisation, 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 deprivation score may allow these factors to be reconsidered.

## Implications for clinical practice

The interactions and engagement between our research group, clinical units and CCAD are an important component of taking the risk model forward for use in routine monitoring for the purposes of quality improvement. The active engagement and genuine collaboration of the analytical team with CCAD and clinical co-applicants and collaborators was essential in the development of this risk model and will be equally essential in any implementation of this model within current practice.

### Dissemination activity

Up to June 2012 we have disseminated our work as:

1. An introductory talk about the planned risk model at the annual CCAD stakeholder meeting held at the Royal College of Surgeons in February 2011. This is an open meeting held for members of the professional community, data managers, commissioners and parent representatives to receive updates related to the national audit of paediatric cardiac procedures, and for individuals from those groups to raise and discuss issues.
2. An oral presentation at an academic Operational Research Society meeting (OR53) in Nottingham in September 2011, the title of which was 'The development of a mortality risk model to adjust for case mix in paediatric cardiac surgery'.
3. An oral presentation at the Multi Institutional Database Meeting in Cambridge in October 2011, the title of which was 'First steps to the development of a mortality risk model to adjust for case mix in paediatric cardiac surgery'. This is a collaborative meeting for those involved in audit databases related to children's heart disease from Europe, UK and North America. It was an opportunity for UK professionals to review and comment on our project.
4. An article concerning the work carried out with diagnostic information with the title 'Use of diagnostic information submitted to the United Kingdom Central Cardiac Audit Database (CCAD): development of categorisation and allocation algorithms' has been published online in the journal *Cardiology in the Young*.<sup>53</sup>
5. An article describing development and validation of the risk model with the title 'Development of a diagnosis- and procedure-based risk model for 30-day outcome after paediatric cardiac surgery' has been published online in the *Journal of Thoracic and Cardiovascular Surgery*.<sup>54</sup>

### Links with national clinical audit (Central Cardiac Audit Database)

One of the co-investigators joined the CCAD steering committee over the course of the project. This has further improved the link between CCAD and the project team, which was already robust given the presence of the senior strategist for the UK cardiac audits as a co-investigator. The CCAD steering committee has a strong interest in the work given their commitment to audit and quality assurance. Furthermore, the work will assist with implementation of one of the recommendations of the *Safe and Sustainable review*<sup>10</sup> to CCAD concerning the introduction of audit methods that offer meaningful and timely assessment of unit performance.

The project and further clinical use of the risk model is a regular agenda item for CCAD steering committee meetings. Over the course of the study, discussions took place planning the future use of the risk model and a kit for its application in routine monitoring of risk adjusted-outcomes to be provided directly to the clinical units, rather than held by CCAD. A significant advantage of this is the potential ability for clinical units to view their risk-adjusted outcomes within 1–2 months of the procedures being performed rather than waiting at least a year for the CCAD audit process (data collection, validation and analysis) to be completed, as is the case with the CCAD funnel plots.

The study team presented further work and next steps at the annual CCAD stakeholder meeting held at the Royal College of Surgeons in February 2012. This represented a further opportunity to discuss application of the risk-adjustment model for routine monitoring of outcomes in the UK cardiac centres.

### Pilot project of variable life-adjusted display

The aim of the project from the outset has been to assess whether or not a model can be developed that is fit for the purpose of risk adjustment of routine monitoring for paediatric cardiac surgery. The model has now been developed and tested by the study team and we believe that it is fit for purpose. The level of engagement of the research team with the professional community has been good, with those who have seen the information about the risk model at the dissemination events expressing interest in the use of the model within their centres to generate VLAD charts<sup>21</sup> for the purposes of routine monitoring. The

intention of the study team going forward is to perform a pilot study to assess the feasibility and logistics of applying the work in this way.

The research team visited the data manager from Evelina Children's Hospital, London, who has not been involved in the research project, in July 2011 to demonstrate the model and its intended use. This was a consultation exercise to explore the feasibility of using the model in routine monitoring, and whether or not the factors selected were transparent and user friendly for a data manager, who will potentially be engaged in the application of the model to this end going forward. The feedback was very positive: the view was that the factors included in the model were those that are already routinely collected and stored at the present time. Applying the model to these data does require some data manipulation and the project team is currently preparing a spreadsheet program that will provide a simple way of doing this.

We have had expressions of interest in the use of risk-adjusted VLAD from the CCAD steering committee meeting and from paediatric cardiothoracic surgeons and paediatric cardiologists based at Southampton General Hospital, Bristol Children's Hospital, Evelina Children's Hospital, Royal Brompton and Harefield Hospitals, Glenfield Hospital Leicester, Birmingham Children's Hospital and Yorkhill Hospital in Glasgow.

It may be that, in the course of working with CCAD and the centres on piloting the use of the risk model and monitoring tool, it is decided that it would be useful to recalibrate the model, particularly if the underestimation of risk among very high-risk episodes seen in the test set after 2007 is considered problematic in new data. Ultimately, given that the aim of monitoring outcomes is to improve outcomes, risk models can (and hopefully do) become out of date and it is likely that, if adopted, this model will need to be recalibrated periodically. Should the model start to be used routinely within units and confidence be gained in its use, dissemination via the CCAD public portal could follow.

## Implications and recommendations for research

The immediate research application of the risk-adjustment model is in the production of VLAD charts. There exists a strong level of engagement and intent to participate in this pilot project from Evelina Children's Hospital and Yorkhill Hospital in Glasgow. The pilot project will hopefully enable VLAD charts to be developed that are useful and meaningful to clinicians. Given the importance of case mix severity in monitoring outcome in paediatric cardiac surgery, it would also be advantageous to explore methods to describe and display information about centre-specific case mix complexity alongside VLAD plots.

The National Institute for Cardiovascular Outcomes Research (NICOR) and CCAD assisted with analytical work underpinning this project over the 5 months between November 2011 and March 2012. This pilot project of risk-adjusted VLAD incorporated data from Great Ormond Street Children's Hospital, Evelina Children's Hospital and Yorkhill Hospital in Glasgow. Although not part of the planned pilot study, future collaboration with qualitative researchers to identify barriers and facilitators of using risk-monitoring tools in clinical practice could be a fruitful area of research.

As a second step, the risk model provides the means to describe case mix severity both nationally and locally at the level of the centres. This information, which has not previously been available, will permit some exploration of trends in case mix severity over time and local practice patterns in terms of, for example, referral of higher-risk cases to particular centres. This type of future research analysis may be of importance when planning health services and would certainly have been of interest to the *Safe and Sustainable* review had it been available at that time.

A longer-term aim of our study was groundwork that could inform subsequent analyses of long-term outcome by diagnosis in CCAD as this may provide valuable information in terms of national audit of outcomes; the latter is a larger project beyond the scope of the current study. Such future analyses would carry the following advantages: diagnosis-based outcomes provide more relevant information to patients

and carers; many CHD patients have multiple procedures during their lives, each of which carries an individual risk; diagnosis-based analyses allow study of the variation in the strategies adopted in different centres for the surgical management of a particular diagnosis; and diagnosis-based analyses provide an appropriate way of assessing a service given that the aim of paediatric cardiac surgery is to contribute to good long-term outcomes for a group of patients with a CHD diagnosis.

We selected terms to label the primary diagnosis groupings, such as 'ventricular septal defect' and 'hypoplastic left heart syndrome', that will be recognisable both to lay persons who may have an interest in this topic and to professionals. This choice was in keeping with the philosophy of CCAD, which shares information about outcomes with the public; we considered that more technical medically focused schemes for describing and allocating primary diagnoses would be less suitable for this purpose.

The groundwork we have carried out has the potential for future exploitation in terms of identifying the incidence of particular cardiac diagnoses in CCAD, which is in effect a national registry of patients who have undergone procedures. Future research in this area may provide useful information to commissioners and other CHD stakeholders. For example, we can now say that the proportion of patients with a UVH in CCAD, which captures all UK children who have undergone an intervention since 2000, is 9.3%. A limitation of this information is the inclusion criterion children undergoing surgical or catheter intervention only, which excludes those children who were unfit for intervention from the outset.





## Chapter 10 Conclusion

A risk model has been developed for monitoring short-term surgical outcomes following paediatric cardiac surgery that, for the first time, makes use of diagnostic information as well as procedural data. The model shows good discrimination between groups of patients with high and low mortality and reasonable accuracy.

Given that the model underestimated risk at the very high-risk end of the recent (since 2007) test data, risk adjustment using the model as currently calibrated will potentially give an unduly negative impression of outcomes at those centres with a high proportion of high-risk cases. This is an important caveat to the interpretation of risk-adjusted outcomes within and across institutions that will need to be discussed with CCAD and the institutions as this work is taken forward. The risk model will require ongoing maintenance and refinement over time, including periodic reparameterisation and/or recalibration to ensure that it remains valid and useful. Concurrent display of centre-specific case mix distributions along with knowledge of model performance may be a useful way of providing context to the model output. The primary intent is for the model to be used locally by clinical teams along with VLAD techniques to monitor their own outcomes over time, with a view to conducting timely internal reviews should there be a worsening of risk-adjusted outcomes and identifying, learning from and sustaining improvements. We stress that this model (and any model) can only partially adjust for risk and should not, on its own, be used for judgments on performance.

The risk model has the potential for use in routine monitoring of risk-adjusted outcomes of paediatric cardiac centres using VLAD charts, which will enable centres to assess trends in outcome promptly. The work performed during development of the risk model in terms of grouping patients by diagnosis rather than procedure in the national database has potential future use for evaluations using units of the underpinning 'condition' rather than individual procedure performed.



# Acknowledgements

We acknowledge with thanks the assistance of Catherine Bull, Nagarajan Muthialu and Rodney Franklin with the diagnostic mapping schemes and the medical coding work and Brian Reddy for his early work on the CCAD data set at the start of this project. We thank John Gibbs for his work with the CCAD project and his advice during the course of this study. We thank Steven Gallivan for his work in operational research that underpins aspects of this study. We acknowledge the contribution of the CCAD steering committee, involved stakeholders and patient groups to the creation and maintenance of CCAD, which is a world-class audit database.

## Contribution of authors

**Christina Pagel** (Senior Research Fellow, Operational Research) cleaned the data, developed and tested the risk model and contributed to writing the report.

**Kate Brown** (Consultant, Cardiac Intensive Care) contributed a clinical perspective to development of the risk model and contributed to writing the report.

**Sonya Crowe** (Research Associate, Operational Research) cleaned the data, developed and tested the risk model and contributed to writing the report.

**Martin Utley** (Professor of Operational Research) supervised the data cleaning and the development and testing of the risk model and contributed to writing the report.

**David Cunningham** (Senior Strategist, NICOR) provided expertise with respect to the data set (particularly *Chapters 3 and 4*) and contributed to writing the report.

**Victor T Tsang** (Consultant Cardiothoracic Surgeon) contributed a clinical perspective to development of the risk model and contributed to writing the report.

## Publications

1. Brown KL, Crowe S, Pagel C, Bull C, Muthialu N, Gibbs J, *et al.* Use of diagnostic information submitted to the United Kingdom Central Cardiac Audit Database: development of categorisation and allocation algorithms [published online ahead of print 2 October 2012]. *Cardiol Young* 2012. doi:<http://dx.doi.org/10.1017/S1047951112001369>. URL: <http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=8704423> (accessed 16 December 2012).
2. Crowe S, Brown KL, Pagel C, Muthialu N, Cunningham D, Gibbs J, *et al.* Development of a diagnosis- and procedure-based risk model for 30-day outcome after pediatric cardiac surgery [published online ahead of print 23 July 2012]. *J Thoracic Cardiovasc Surg* 2012. URL: [www.jtcvsonline.org/article/S0022-5223%2812%2900701-5/abstract](http://www.jtcvsonline.org/article/S0022-5223%2812%2900701-5/abstract) (accessed 16 December 2012).



## References

1. Knowles R, Griebisch I, Dezateux C, Brown J, Bull C, Wren C. Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2005;**9**(44).
2. Billett J, Cowie MR, Gatzoulis MA, Vonder Muhll IF, Majeed A. Comorbidity, healthcare utilisation and process of care measures in patients with congenital heart disease in the UK: cross-sectional, population-based study with case-control analysis. *Heart* 2008;**94**:1194–9. <http://dx.doi.org/10.1136/hrt.2007.122671>
3. McCrindle BW, Williams RV, Mitchell PD, Hsu DT, Paridon SM, Atz AM, *et al.* Relationship of patient and medical characteristics to health status in children and adolescents after the Fontan procedure. *Circulation* 2006;**113**:1123–9. <http://dx.doi.org/10.1161/CIRCULATIONAHA.105.576660>
4. Shillingford AJ, Glanzman MM, Ittenbach RF, Clancy RR, Gaynor JW, Wernovsky G. Inattention, hyperactivity, and school performance in a population of school-age children with complex congenital heart disease. *Pediatrics* 2008;**121**:e759–67. <http://dx.doi.org/10.1542/peds.2007-1066>
5. Office for National Statistics. Report: death registrations in England and Wales, 2006: causes. *Health Stat Q* 2007;**34**.
6. Office for National Statistics. Report: death registrations in England and Wales, 2005: causes. *Health Stat Q* 2006;**33**.
7. Billett J, Majeed A, Gatzoulis M, Cowie M. Trends in hospital admissions, in-hospital case fatality and population mortality from congenital heart disease in England, 1994 to 2004. *Heart* 2008;**94**:342–8. <http://dx.doi.org/10.1136/hrt.2006.113787>
8. Office for National Statistics. Report: death registrations in England and Wales, 2007: causes. *Health Stat Q* 2008;**35**.
9. Central Cardiac Audit Database. *Central Cardiac Audit Database: paediatric analysis home page. Congenital heart disease website*; 2011. URL: [www.ccad.org.uk](http://www.ccad.org.uk) (accessed 29 January 2011).
10. Knowles RL, Griebisch I, Bull C, Brown J, Wren C, Dezateux C. Quality of life and congenital heart defects: comparing parent and professional values. *Arch Dis Child* 2007;**92**:388–93. <http://dx.doi.org/10.1136/adc.2005.075606>
11. NHS. *Safe and sustainable: Children's Congenital Cardiac Services: the future of children's heart services*; 2011. URL: [www.specialisedservices.nhs.uk/safe\\_sustainable/childrens-congenital-cardiac-services](http://www.specialisedservices.nhs.uk/safe_sustainable/childrens-congenital-cardiac-services) (accessed 26 February 2011).
12. 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. <http://dx.doi.org/10.1016/j.athoracsur.2008.10.065>
13. Stark J, Gallivan S, Lovegrove J, Hamilton JRL, Monro JL, Pollock JCS, *et al.* Mortality rates after surgery for congenital heart defects in children and surgeons' performance. *Lancet* 2000;**355**:1004–7. [http://dx.doi.org/10.1016/S0140-6736\(00\)90001-1](http://dx.doi.org/10.1016/S0140-6736(00)90001-1)
14. Spiegelhalter DJ. Mortality and volume of cases in paediatric cardiac surgery: retrospective study based on routinely collected data. *BMJ* 2002;**324**:261–3. <http://dx.doi.org/10.1136/bmj.324.7332.261>
15. Spiegelhalter DJ. Risk stratification for open heart surgery. *BMJ* 1992;**305**:1500. <http://dx.doi.org/10.1136/bmj.305.6867.1500-a>

16. Spiegelhalter D, Grigg O, Kinsman R, Treasure T. Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care* 2003;**15**:7–13. <http://dx.doi.org/10.1093/intqhc/15.1.7>
17. Gibbs JL, Monro JL, Cunningham D, Rickards A. Survival after surgery or therapeutic catheterisation for congenital heart disease in children in the United Kingdom: analysis of the central cardiac audit database for 2000–1. *BMJ* 2004;**328**:611. <http://dx.doi.org/10.1136/bmj.38027.613403.F6>
18. 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. [http://dx.doi.org/10.1016/S0140-6736\(01\)05404-6](http://dx.doi.org/10.1016/S0140-6736(01)05404-6)
19. 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**:I3–12.
20. Bridgewater B, Grayson AD, Brooks N, Grotte G, Fabri BM, Au J, *et al.* Has the publication of cardiac surgery outcome data been associated with changes in practice in northwest England: an analysis of 25,730 patients undergoing CABG surgery under 30 surgeons over eight years. *Heart* 2007;**93**:744–8. <http://dx.doi.org/10.1136/hrt.2006.106393>
21. Sherlaw-Johnson C, Lovegrove J, Treasure T, Gallivan S. Likely variations in perioperative mortality associated with cardiac surgery: when does high mortality reflect bad practice? *Heart* 2000;**84**:79–82. <http://dx.doi.org/10.1136/heart.84.1.79>
22. Lovegrove J, Valencia O, Treasure T, Sherlaw-Johnson C, Gallivan S. Monitoring the results of cardiac surgery by variable life-adjusted display. *Lancet* 1997;**350**:1128–30. [http://dx.doi.org/10.1016/S0140-6736\(97\)06507-0](http://dx.doi.org/10.1016/S0140-6736(97)06507-0)
23. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg* 1999;**16**:9–13. [http://dx.doi.org/10.1016/S1010-7940\(99\)00134-7](http://dx.doi.org/10.1016/S1010-7940(99)00134-7)
24. Kang N, Cole T, Tsang V, Elliott M, de Leval M. Risk stratification in paediatric open-heart surgery. *Eur J Cardiothorac Surg* Jul 2004;**26**:3–11.
25. 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. <http://dx.doi.org/10.1016/j.jtcvs.2009.03.071>
26. 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**:110–18. <http://dx.doi.org/10.1067/mtc.2002.119064>
27. 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. <http://dx.doi.org/10.1053/j.pcsu.2004.02.011>
28. Kang N, Tsang VT, Gallivan S, Sherlaw-Johnson C, Cole TJ, Elliott MJ, *et al.* Quality assurance in congenital heart surgery. *Eur J Cardiothorac Surg* May 2006;**29**:693–7; discussion 697–8. <http://dx.doi.org/10.1016/j.ejcts.2006.01.034>
29. Kang N, Tsang VT, Elliott MJ, de Leval MR, Cole TJ. Does the Aristotle score predict outcome in congenital heart surgery? *Eur J Cardiothorac Surg* 2006;**29**:986–8. <http://dx.doi.org/10.1016/j.ejcts.2006.01.066>
30. Jacobs ML, Jacobs JP, Franklin RC, 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. <http://dx.doi.org/10.1017/S1047951108002813>

31. Jenkins KJ, Beekman IJH, Bergersen LJ, Everett AD, Forbes TJ, Franklin RC, *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. <http://dx.doi.org/10.1017/S1047951108002825>
32. Clarke DR, Breen LS, Jacobs ML, Franklin RC, Tobota Z, Maruszewski B, *et al.* Verification of data in congenital cardiac surgery. *Cardiol Young* 2008;**18**(Suppl. 2):177–87. <http://dx.doi.org/10.1017/S1047951108002862>
33. Franklin RC, 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. <http://dx.doi.org/10.1017/S1047951108002795>
34. Franklin RC. The European Paediatric Cardiac Code Long List: structure and function. *Cardiol Young* 2000;**10**(Suppl. 1):27–146. <http://dx.doi.org/10.1017/S1047951100007770>
35. Bergersen L, Everett AD, Giroud JM, Martin GR, Franklin RC, Béland MJ, *et al.* Report from the International Society for Nomenclature of Paediatric and Congenital Heart Disease: cardiovascular catheterisation for congenital and paediatric cardiac disease (part 1 – procedural nomenclature). *Cardiol Young* 2011;**21**:252–9. <http://dx.doi.org/10.1017/S104795111000185X>
36. Jacobs JP, Anderson RH, Weinberg PM, Walters HL III, 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. <http://dx.doi.org/10.1017/S1047951107001138>
37. Brown KL, Ridout DA, Goldman AP, Hoskote A, Penny DJ. Risk factors for long intensive care unit stay after cardiopulmonary bypass in children. *Crit Care Med* 2003;**31**:28–33. <http://dx.doi.org/10.1097/00003246-200301000-00004>
38. 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. <http://dx.doi.org/10.1016/j.ejcts.2005.12.057>
39. Wellesley D, Boyd P, Dolk H, Pattenden S. An aetiological classification of birth defects for epidemiological research. *J Med Genet* 2005;**42**:54–7. <http://dx.doi.org/10.1136/jmg.2004.023309>
40. 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. <http://dx.doi.org/10.1136/hrt.2005.078097>
41. Jacobs JP, Jacobs ML, Mavroudis C, Backer CL, Lacour-Gayet FG, Tchervenkov CI, *et al.* Nomenclature and databases for the surgical treatment of congenital cardiac disease – an updated primer and an analysis of opportunities for improvement. *Cardiol Young* 2008;**18**(Suppl. 2):38–62. <http://dx.doi.org/10.1017/S1047951108003028>
42. Report of the New England Regional Infant Cardiac Program. *Pediatrics* 1980;**65**:375–461.
43. Ferencz C, Rubin JD, McCarter RJ, Brenner JI, Neill CA, Perry LW, *et al.* Congenital heart disease: prevalence at livebirth. The Baltimore–Washington Infant Study. *Am J Epidemiol* 1985;**121**:31–6.
44. Moller JH, Moodie DS, Blee M, Norton JB, Nouri S. Symptomatic heart disease in infants: comparison of three studies performed during 1969–1987. *Pediatr Cardiol* 1995;**16**:216–22. <http://dx.doi.org/10.1007/BF00795710>
45. Extracorporeal Life Support Organisation. *Registry of the Extracorporeal Life Support Organisation*. Ann Arbor, MI: Extracorporeal Life Support Organisation; 2011.
46. Wren C, Richmond S, Donaldson L. Temporal variability in birth prevalence of cardiovascular malformations. *Heart* 2000;**83**:414–19. <http://dx.doi.org/10.1136/heart.83.4.414>

47. Wren C, O'Sullivan JJ. Survival with congenital heart disease and need for follow up in adult life. *Heart* 2001;**85**:438–43. <http://dx.doi.org/10.1136/heart.85.4.438>
48. Wren C, Reinhardt Z, Khawaja K. Twenty-year trends in diagnosis of life-threatening neonatal cardiovascular malformations. *Arch Dis Child Fetal Neonatal Ed* 2008;**93**:F33–5. <http://dx.doi.org/10.1136/adc.2007.119032>
49. Clancy RR, McGaurn SA, Wernovsky G, Spray TL, Norwood WI, Jacobs ML, *et al*. Preoperative risk-of-death prediction model in heart surgery with deep hypothermic circulatory arrest in the neonate. *J Thorac Cardiovasc Surg* 2000;**119**:347–57. [http://dx.doi.org/10.1016/S0022-5223\(00\)70191-7](http://dx.doi.org/10.1016/S0022-5223(00)70191-7)
50. Harrell F. *Regression modelling strategies*. New York, NY: Springer; 2001.
51. Shahian DM, Edwards FH, Jacobs JP, Prager RL, Normand SL, Shewan CM, *et al*. Public reporting of cardiac surgery performance: part 1 – history, rationale and consequences. *Ann Thorac Surg* 2011: **92**(Suppl. 3):S2–11. <http://dx.doi.org/10.1016/j.athoracsur.2011.06.100>
52. Shahian DM, Edwards FH, Jacobs JP, Prager RL, Normand SL, Shewan CM, *et al*. Public reporting of cardiac surgery performance: part 2 – implementation. *Ann Thorac Surg* 2011: **92**(Suppl. 3):S12–23. <http://dx.doi.org/10.1016/j.athoracsur.2011.06.101>
53. Brown KL, Crowe S, Pagel C, Bull C, Muthialu N, Gibbs J, *et al*. Use of diagnostic information submitted to the United Kingdom Central Cardiac Audit Database: development of categorisation and allocation algorithms [published online ahead of print 2 October 2012]. *Cardiol Young* 2012. URL: <http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=8704423> (accessed 16 December 2012). <http://dx.doi.org/10.1017/S1047951112001369>
54. Crowe S, Brown KL, Pagel C, Muthialu N, Cunningham D, Gibbs J, *et al*. Development of a diagnosis- and procedure-based risk model for 30-day outcome after pediatric cardiac surgery [published online ahead of print 23 July 2012]. *J Thoracic Cardiovasc Surg* 2012. URL: [www.jtcvsonline.org/article/S0022-5223%2812%2900701-5/abstract](http://www.jtcvsonline.org/article/S0022-5223%2812%2900701-5/abstract) (accessed 16 December 2012).
55. Copas JB. Using regression models for prediction: shrinkage and regression to the mean. *Stat Methods Med Res* 1997;**6**:167–83. <http://dx.doi.org/10.1191/096228097667367976>



## Appendix 1 Dealing with duplicate data

Figure 37 shows the protocol that was used to decide which record from a set of duplicate records should be retained within the data set. Note that the numbers given do not include cases in which there were more than two duplicates of a single record. These were dealt with in the same manner as shown here.

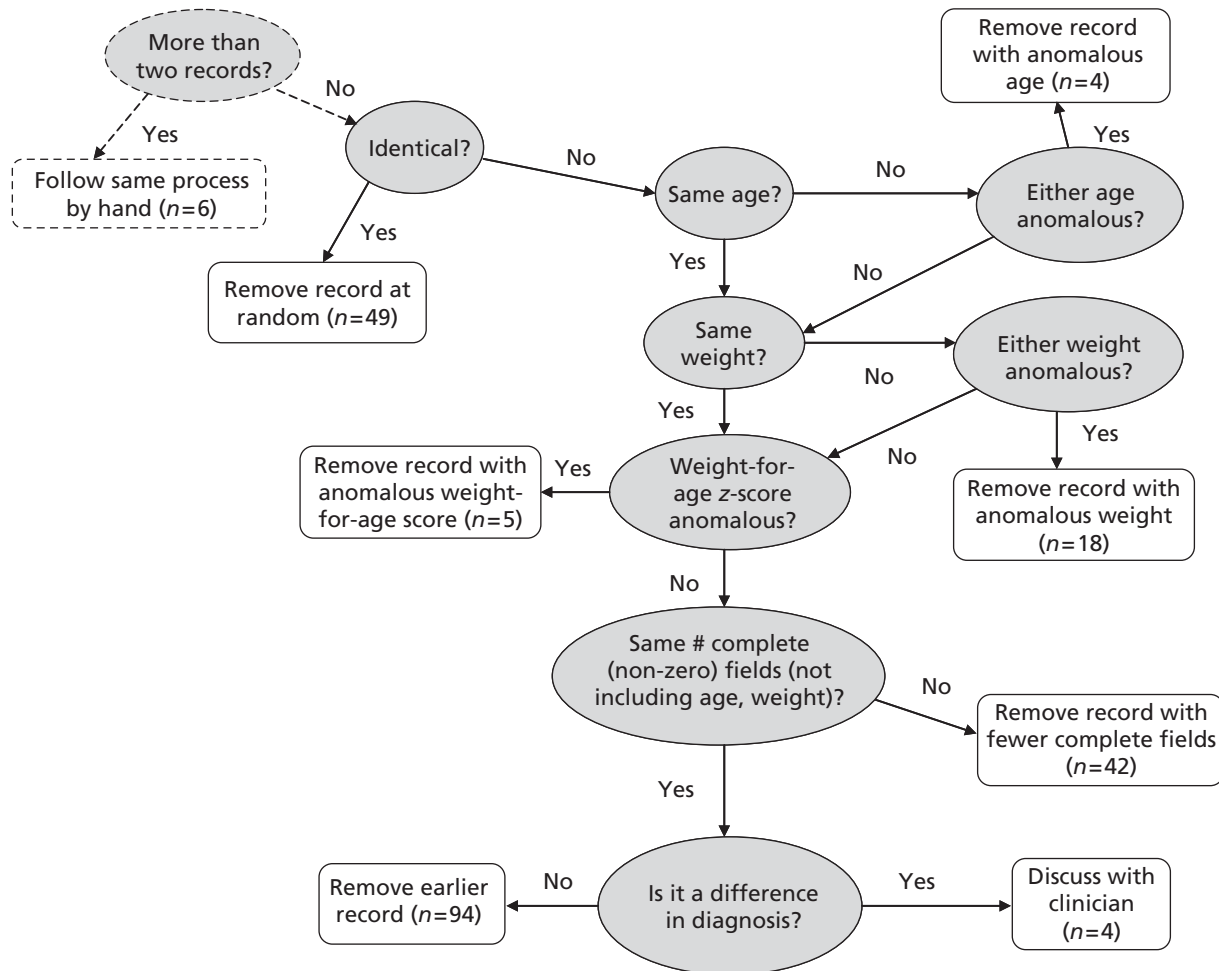


FIGURE 37 The process used to decide which record to retain from a set of duplicate records.



## Appendix 2 Moving from International Paediatric Congenital and Cardiac codes to diagnostic groups

In this appendix we provide the full mappings from the CCAD diagnostic and comorbidity codes to the corresponding groups used within our analysis (*Chapter 4*). At the end of this appendix we also give the protocol for moving from a set of groups for a single record to an overall primary diagnostic group (scheme 1) or a set of diagnostic indicators (scheme 2) for a record.

**TABLE 9** Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1

CCAD official diagnosis codes	Diagnosis group in scheme 1
010100. Normal heart	Normal
010101. Tetralogy of Fallot	Fallot/DORV: Fallot type
010102. TGA (concordant AV and discordant VA connections) and IVS	TGA (concordant AV and discordant VA connections) and IVS
010103. Congenitally corrected TGA (discordant AV and VA connections)	Miscellaneous congenital
010104. DORV	Miscellaneous congenital
010106. Pulmonary atresia + VSD (including Fallot type)	Pulmonary atresia + VSD (including Fallot type)
010107. Pulmonary atresia + IVS	Pulmonary atresia (including pulmonary atresia + IVS)
010109. HLHS	HLHS
010114. Double inlet AV connection (double inlet ventricle)	Functionally UVH
010116. Partially anomalous pulmonary venous connections: Scimitar syndrome	Miscellaneous congenital
010117. DORV: Fallot type (subaortic or doubly committed VSD and pulmonary stenosis)	Fallot/DORV: Fallot type
010118. DORV: transposition type (subpulmonary VSD)	TGA + VSD/DORV: TGA type
010119. DORV: with non-committed VSD	Miscellaneous congenital
010120. AV septal defect and tetralogy of Fallot	AVSD
010122. Functionally UVH	Functionally UVH
010124. DORV: with IVS	Functionally UVH
010125. Pulmonary atresia + VSD + systemic-to-pulmonary collateral artery(ies) [MAPCA(s)]	Pulmonary atresia + VSD (including Fallot type)
010139. Cardiac abnormality	Miscellaneous congenital
010140. DORV: subaortic or doubly committed VSD without pulmonary stenosis ('VSD type')	Miscellaneous congenital
010160. Vascular abnormality	Miscellaneous congenital
010300. Usual atrial arrangement (atrial situs solitus)	Normal
010306. Abnormal atrial arrangement	Miscellaneous congenital
010309. AV and/or VA connections abnormal	Miscellaneous congenital
010310. Normal atrial arrangement (situs), AV and VA connections	Normal
010403. Double inlet right ventricle	Functionally UVH
010404. Double inlet right ventricle	Functionally UVH

continued

**TABLE 9** Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (*continued*)

CCAD official diagnosis codes	Diagnosis group in scheme 1
010500. Concordant VA connections	Normal
010501. Discordant VA connections (TGA)	TGA + VSD/DORV: TGA type
010503. Double outlet left ventricle	Miscellaneous congenital
010510. Concordant VA connections with parallel great arteries (anatomically corrected malposition)	Miscellaneous congenital
020102. Dextrocardia: heart predominantly in right hemithorax	Miscellaneous congenital
020109. Position/orientation of heart abnormal	Miscellaneous congenital
020305. Solitary ventricle of indeterminate morphology	Functionally UVH
030103. Total mirror imagery (situs inversus)	Miscellaneous congenital
030104. Right isomerism ('asplenia')	Miscellaneous congenital
030105. Left isomerism ('polysplenia')	Miscellaneous congenital
030109. Position or morphology of thoraco-abdominal organs abnormal	Comorbidity
030305. Tracheobronchial anomaly	Comorbidity
030703. Spleen absent (asplenia)	Comorbidity
030704. Multiple spleens (polysplenia)	Comorbidity
040100. Superior caval vein abnormality	Miscellaneous congenital
040101. Left superior caval vein persisting to coronary sinus	Miscellaneous congenital
040200. Hepatic vein abnormality	Miscellaneous congenital
040300. Inferior caval vein abnormality	Miscellaneous congenital
040310. Inferior caval vein interruption (absent suprarenal segment) with azygos continuation	Miscellaneous congenital
040400. Coronary sinus abnormality	Miscellaneous congenital
040500. Systemic vein abnormality: congenital	Miscellaneous congenital
040600. Totally anomalous pulmonary venous connection: supracardiac	TAPVC
040701. Partially anomalous pulmonary venous connection(s)	Interatrial communication ('ASD')
040800. Pulmonary vein abnormality	Miscellaneous congenital
040805. Totally anomalous pulmonary venous connection	TAPVC
040806. Obstructed pulmonary venous connection(s)	TAPVC
040810. Totally anomalous pulmonary venous connection: intracardiac	TAPVC
040820. Totally anomalous pulmonary venous connection: infracardiac	TAPVC
040830. Totally anomalous pulmonary venous connection: mixed	TAPVC
040891. Pulmonary vein stenosis	Acquired
050100. Right atrial abnormality	Miscellaneous congenital
050200. Left atrial abnormality	Miscellaneous congenital
050201. Cor triatriatum (divided left atrium)	Miscellaneous congenital
050202. Supravalvar mitral ring	Mitral valve abnormality (including supravalvar, subvalvar)
050300. Atrial septum abnormality	Miscellaneous congenital

**TABLE 9** Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (*continued*)

CCAD official diagnosis codes	Diagnosis group in scheme 1
050301. Patent foramen ovale	Miscellaneous congenital
050310. Intact atrial septum (no interatrial communication)	Normal
050401. Interatrial communication ('ASD')	Interatrial communication ('ASD')
050402. ASD within oval fossa (secundum)	Interatrial communication ('ASD')
050403. Spontaneous closure of ASD within oval fossa (secundum)	Interatrial communication ('ASD')
050500. Sinus venosus defect (ASD)	Interatrial communication ('ASD')
050503. Interatrial communication (ASD) through coronary sinus orifice	Interatrial communication ('ASD')
050601. Common atrium (virtual absence of atrial septum)	AVSD
060100. Tricuspid valvar abnormality	Tricuspid valve abnormality (including Ebstein's)
060101. Tricuspid atresia	Functionally UVH
060103. Tricuspid valvar dysplasia	Tricuspid valve abnormality (including Ebstein's)
060109. Straddling tricuspid valve	Miscellaneous congenital
060125. Tricuspid regurgitation: congenital	Tricuspid valve abnormality (including Ebstein's)
060134. Ebstein's malformation of tricuspid valve	Tricuspid valve abnormality (including Ebstein's)
060191. Tricuspid regurgitation	Tricuspid valve abnormality (including Ebstein's)
060192. Tricuspid stenosis	Tricuspid valve abnormality (including Ebstein's)
060200. Mitral valvar abnormality	Mitral valve abnormality (including supralvalvar, subvalvar)
060201. Mitral atresia	HLHS
060207. Mitral valvar stenosis: congenital	Mitral valve abnormality (including supralvalvar, subvalvar)
060209. Straddling mitral valve	Miscellaneous congenital
060212. Mitral subvalvar apparatus abnormality	Mitral valve abnormality (including supralvalvar, subvalvar)
060213. Mitral subvalvar stenosis	Mitral valve abnormality (including supralvalvar, subvalvar)
060225. Mitral regurgitation: congenital	Mitral valve abnormality (including supralvalvar, subvalvar)
060235. Mitral valvar prolapse	Mitral valve abnormality (including supralvalvar, subvalvar)
060236. True cleft of mitral leaflet (without AVSD)	Mitral valve abnormality (including supralvalvar, subvalvar)
060256. Parachute malformation of mitral valve	Mitral valve abnormality (including supralvalvar, subvalvar)
060291. Mitral regurgitation	Mitral valve abnormality (including supralvalvar, subvalvar)
060292. Mitral stenosis	Mitral valve abnormality (including supralvalvar, subvalvar)

continued

TABLE 9 Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (continued)

CCAD official diagnosis codes	Diagnosis group in scheme 1
060293. Mitral valve stenosis	Mitral valve abnormality (including supralvalvar, subvalvar)
060501. AVSD AV valvar abnormality	AVSD
060506. AVSD AV valvar regurgitation	AVSD
060600. AVSD	AVSD
060601. AVSD: isolated atrial component (primum ASD) (partial)	AVSD
060608. AVSD: isolated ventricular component	AVSD
060609. AVSD: atrial and ventricular components with common AV orifice (complete)	AVSD
060610. AVSD: atrial and (restrictive) ventricular components + separate AV valves ('intermediate')	AVSD
060726. AVSD with ventricular imbalance	Functionally UVH
070100. Right ventricular abnormality	Miscellaneous congenital
070110. Arrhythmogenic right ventricular cardiomyopathy	Acquired
070111. Right ventricular dysfunction	Acquired
070114. Right ventricular aneurysm	Miscellaneous congenital
070200. Right ventricular hypoplasia	Miscellaneous congenital
070301. Double-chambered right ventricle	Miscellaneous congenital
070501. Right ventricular outflow tract obstruction	Miscellaneous congenital
070530. Subpulmonary stenosis	Pulmonary stenosis
070600. Left ventricular abnormality	Miscellaneous congenital
070610. Left ventricular dysfunction	Acquired
070613. Left ventricular aneurysm	Miscellaneous congenital
070700. Left ventricular hypoplasia	Miscellaneous congenital
070841. Functionally UVH	Functionally UVH
070842. Functionally UVH	Functionally UVH
070850. Ventricular myocardial non-compaction cardiomyopathy	Acquired
070900. Subaortic stenosis	Subaortic stenosis (isolated)
070901. Left ventricular outflow tract obstruction	Miscellaneous congenital
070903. Subaortic stenosis due to fibromuscular shelf	Subaortic stenosis (isolated)
070931. Aortic abnormality	Miscellaneous congenital
071000. VSD	VSD
071001. Perimembranous VSD	VSD
071012. VSD + malaligned outlet septum	VSD
071101. Muscular VSD	VSD
071200. Subarterial VSD	VSD
071201. Doubly committed subarterial VSD	VSD
071402. Communication between left ventricle + right atrium (Gerbode defect)	VSD
071405. Inlet VSD	VSD

**TABLE 9** Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (*continued*)

CCAD official diagnosis codes	Diagnosis group in scheme 1
071501. Tiny VSD (Maladie de Roger)	VSD
071504. Multiple VSDs	VSD
071505. Single VSD	VSD
071601. Spontaneous closure of VSD	Normal
072000. Ventricular septal abnormality	Miscellaneous congenital
072001. Aneurysm of membranous septum	Normal
072100. IVS	Normal
090101. Common arterial trunk (truncus arteriosus)	Common arterial trunk (truncus arteriosus)
090200. Truncal valvar abnormality	Common arterial trunk (truncus arteriosus)
090203. Truncal valvar regurgitation	Common arterial trunk (truncus arteriosus)
090401. Aortopulmonary window	Miscellaneous congenital
090500. Pulmonary valvar abnormality	Pulmonary stenosis
090501. Pulmonary valvar stenosis	Pulmonary stenosis
090504. Pulmonary valvar stenosis: congenital	Pulmonary stenosis
090511. Pulmonary atresia	Pulmonary atresia (including pulmonary atresia + IVS)
090512. Pulmonary atresia: imperforate valve	Pulmonary atresia (including pulmonary atresia + IVS)
090515. Pulmonary valvar atresia: acquired	Acquired
090522. Pulmonary regurgitation: congenital	Acquired
090525. Absent pulmonary valve syndrome: Fallot type	Fallot/DORV: Fallot type
090591. Pulmonary regurgitation	Acquired
090592. Pulmonary stenosis	Pulmonary stenosis
090700. Pulmonary trunk (main pulmonary artery) abnormality	Acquired
090711. Pulmonary trunk hypoplasia	Acquired
090713. Supravalvar pulmonary trunk stenosis	Acquired
090726. Solitary arterial trunk (absent intrapericardial pulmonary arteries)	Pulmonary atresia + VSD (including Fallot type)
090801. Major systemic-to-pulmonary collateral artery(ies) [MAPCA(s)]	Pulmonary atresia + VSD (including Fallot type)
090906. Pulmonary arterial sling	Miscellaneous congenital
090908. Pulmonary artery from ascending aorta (hemitruncus)	Miscellaneous congenital
091000. Pulmonary arterial abnormality	Acquired
091001. Pulmonary arterial stenosis	Acquired
091006. Peripheral pulmonary arterial stenoses: at/beyond hilar bifurcation	Acquired
091007. Central pulmonary arterial stenosis: proximal to hilar bifurcation	Acquired
091010. Discontinuous (non-confluent) pulmonary arteries	Miscellaneous congenital

*continued*

TABLE 9 Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (continued)

CCAD official diagnosis codes	Diagnosis group in scheme 1
091011. Pulmonary arterial hypoplasia	Acquired
091025. Right pulmonary arterial stenosis	Acquired
091026. Left pulmonary arterial stenosis	Acquired
091044. Pulmonary arterial aneurysm	Acquired
091500. Aortic valvar abnormality	Miscellaneous congenital
091501. Aortic valvar stenosis: congenital	Aortic valve stenosis (isolated)
091503. Aortic atresia	HLHS
091507. Aortic regurgitation: congenital	Aortic regurgitation
091512. Eccentric opening of tricuspid aortic valve	Aortic valve stenosis (isolated)
091513. Aortic valvar stenosis	Aortic valve stenosis (isolated)
091522. Bicuspid aortic valve	Miscellaneous congenital
091530. Aortic valvar prolapse	Aortic regurgitation
091591. Aortic regurgitation	Aortic regurgitation
091592. Aortic stenosis	Aortic valve stenosis (isolated)
091600. Supravalvar aortic stenosis	Miscellaneous congenital
091602. Ascending aorta hypoplasia	Miscellaneous congenital
091605. Ascending aorta dilatation associated with Marfan syndrome	Acquired
091609. Ascending aorta dilatation	Acquired
091610. Ascending aorta abnormality	Miscellaneous congenital
091613. Aortic root dilatation	Acquired
091701. Aortoventricular tunnel	Miscellaneous congenital
091702. Aorto-left ventricular tunnel	Miscellaneous congenital
091801. Aortic sinus of Valsalva aneurysm	Miscellaneous congenital
091901. Arteriovenous fistula (malformation)	Miscellaneous congenital
091905. Pulmonary arteriovenous fistula (malformation)	Miscellaneous congenital
092020. Distal systemic arterial abnormality	Miscellaneous congenital
092025. Systemic-to-pulmonary collateral arter(ies) [MAPCA(s)] stenosis(es)	Pulmonary atresia + VSD (including Fallot type)
092700. Arterial duct (ductus arteriosus) abnormality	PDA
092721. Patent arterial duct (PDA)	PDA
092800. Aortic arch abnormality	Miscellaneous congenital
092809. Double aortic arch	Miscellaneous congenital
092815. Right aortic arch	Miscellaneous congenital
092816. Descending aorta dilatation	Acquired
092901. Aortic coarctation	Aortic arch obstruction +/- VSD/ASD
092911. Aortic arch hypoplasia (tubular)	Aortic arch obstruction +/- VSD/ASD
092916. Descending abdominal aorta hypoplasia (middle aortic syndrome)	Miscellaneous congenital
092931. Interrupted aortic arch	Interrupted aortic arch



**TABLE 9** Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (*continued*)

CCAD official diagnosis codes	Diagnosis group in scheme 1
093000. Aortic arch branch abnormality	Miscellaneous congenital
093002. Aberrant origin right subclavian artery	Miscellaneous congenital
093004. Aberrant origin left subclavian artery	Miscellaneous congenital
093100. Vascular ring	Miscellaneous congenital
094101. Anomalous origin of coronary artery from pulmonary artery	Miscellaneous congenital
094200. Coronary artery: anomalous aortic origin or course	Miscellaneous congenital
094305. Intramural proximal coronary arterial course	Miscellaneous congenital
094501. Coronary fistula	Miscellaneous congenital
094511. Coronary fistulas from right ventricle ('sinusoidal')	Miscellaneous congenital
094600. Coronary arterial abnormality	Miscellaneous congenital
094601. Coronary arterial aneurysm(s)	Acquired
094606. Right ventricle-dependent coronary circulation	Miscellaneous congenital
100100. Pericardial abnormality	Miscellaneous congenital
100301. Heart tumour	Acquired
100501. Acute rheumatic fever	Acquired
100521. Rheumatic fever with cardiac involvement	Acquired
100530. Rheumatic valvar disease	Acquired
100531. Rheumatic mitral valvar disease	Acquired
100533. Rheumatic aortic valvar disease	Acquired
100601. Infective endocarditis	Acquired
100620. Heart abscess	Acquired
100641. Bacterial endocarditis	Acquired
100664. Postprocedural endocarditis	Acquired
100701. Infectious myocarditis	Acquired
100703. Viral myocarditis	Acquired
100705. Drug-induced heart muscle disease	Acquired
100708. Trypanosomal myocarditis (Chagas' disease)	Acquired
100740. Myocardial failure in end-stage CHD	Acquired
100742. Heart muscle disease in cardiac rejection	Acquired
100761. Nutritional heart muscle disease	Acquired
100771. Heart muscle disease in infant of diabetic mother	Acquired
100781. Heart muscle disease in collagen vascular/connective tissue disorder	Acquired
100800. Pericarditis	Acquired
100801. Infectious pericarditis	Acquired
100803. Viral pericarditis	Acquired
100804. Bacterial pericarditis	Acquired
100809. Constrictive pericarditis	Acquired

*continued*

TABLE 9 Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (continued)

CCAD official diagnosis codes	Diagnosis group in scheme 1
100813. Cardiac tamponade	Acquired
100829. Pericardial abnormality: acquired	Acquired
100831. Pericardial effusion	Acquired
100901. Kawasaki disease	Acquired
100902. Kawasaki disease with aneurysm(s) or dilated coronary vessels	Acquired
100908. Kawasaki disease without cardiac involvement	Acquired
100910. Acquired coronary arterial disease	Acquired
100930. Ischaemic heart disease	Acquired
101001. Cardiomyopathy	Acquired
101011. Idiopathic restrictive cardiomyopathy	Acquired
101012. Endocardial fibroelastosis	Acquired
101013. Infiltrative cardiomyopathy	Acquired
101020. Hypertrophic cardiomyopathy	Acquired
101025. Dilated cardiomyopathy	Acquired
101201. Innocent murmur	Normal
101239. Failure to thrive	Acquired
101242. Chest pain	Empty/unknown
101301. Pulmonary arterial hypertension	Acquired
101302. Primary pulmonary hypertension	Acquired
101306. Pulmonary vascular disease	Acquired
101308. Irreversible pulmonary vascular disease due to CHD (Eisenmenger syndrome)	Acquired
101320. Secondary pulmonary hypertension	Acquired
101321. Pulmonary hypertension due to left-to-right shunt	Acquired
101350. Pulmonary arterial disease: acquired	Acquired
101351. Pulmonary embolism	Acquired
101400. Secondary systemic hypertension	Comorbidity
101401. Systemic hypertension	Acquired
101402. Primary (essential) systemic hypertension	Comorbidity
101404. Systemic hypertension due to aortic arch obstruction	Acquired
101440. Ascending aorta dilatation: acquired	Acquired
101442. Ascending aortic aneurysm	Acquired
101443. Descending aortic aneurysm	Acquired
101444. Abdominal aortic aneurysm	Comorbidity
101445. Rupture of thoracic aortic aneurysm	Comorbidity
101446. Rupture of abdominal aortic aneurysm	Comorbidity
101450. Aortic aneurysm	Acquired
101451. Aortic dissection	Acquired

**TABLE 9** Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (*continued*)

CCAD official diagnosis codes	Diagnosis group in scheme 1
101452. Ascending aorta dissection and propagation beyond arch (DeBakey type I)	Acquired
101453. Ascending aorta dissection not beyond arch (DeBakey type II/Stanford type A)	Acquired
101454. Descending aorta dissection and distal propagation (DeBakey type III/Stanford type B)	Comorbidity
101460. Systemic arteritis	Comorbidity
101470. Abnormality of aorta: acquired	Acquired
101472. Recoarctation of aorta	Acquired
101477. Supravalvar aortic stenosis: acquired	Acquired
101480. Arterial duct (ductus arteriosus) abnormality: acquired	Acquired
101500. Neonatal disorder	Empty/unknown
101501. Persistent pulmonary hypertension of the newborn (persistent fetal circulation)	Acquired
101505. Necrotising enterocolitis	Comorbidity
101510. Transient myocardial ischaemia	Acquired
101512. Meconium aspiration	Comorbidity
101600. Right ventricular abnormality: acquired	Acquired
101608. Right ventricular congestive heart failure	Acquired
101616. Right ventricular outflow tract obstruction: acquired	Acquired
101640. Left ventricular abnormality: acquired	Acquired
101646. Recurrent left ventricular outflow tract obstruction	Acquired
101647. Left ventricular failure	Acquired
101660. Abnormality associated with ventricular septum: acquired	Acquired
101662. Postmyocardial infarct VSD	Acquired
101681. Narrowing of constructed intraventricular tunnel: acquired	Acquired
101682. Subaortic stenosis in complex heart disease: acquired	Acquired
101683. Subpulmonary stenosis in complex heart disease: acquired	Acquired
101700. Symptom/sign of heart disease	Empty/unknown
101703. Cyanosis	Empty/unknown
101705. Heart failure	Acquired
101712. Cyanotic spells	Acquired
101713. Palpitations	Acquired
101740. Atrial septum abnormality: acquired	Acquired
101800. Myocardial infarction	Acquired
101801. Acute myocardial infarction	Acquired
101901. Dyslipidaemia	Empty/unknown
102000. No preprocedural risk factors	Normal
102002. Preprocedural shock	Comorbidity

continued

TABLE 9 Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (continued)

CCAD official diagnosis codes	Diagnosis group in scheme 1
102003. Preprocedural arrhythmia	Comorbidity
102007. Preprocedural renal failure (creatinine > 176 µmol/l)	Comorbidity
102009. Preprocedural septicaemia	Comorbidity
102012. Preprocedural neurological impairment	Comorbidity
102014. Preprocedural mechanical ventilatory support	Comorbidity
102015. Preprocedural mechanical circulatory support	Comorbidity
102016. Preprocedural pulmonary hypertension	Comorbidity
102017. Preprocedural tracheostomy	Comorbidity
102018. Preprocedural seizures	Comorbidity
102019. Preprocedural risk factor	Empty/unknown
102202. Premature birth	Comorbidity
102203. Infant of diabetic mother	Comorbidity
102205. Premature birth 32–35 weeks	Comorbidity
102206. Premature birth < 32 weeks	Comorbidity
102300. Hereditary/non-cardiac abnormality not apparent	Empty/unknown
102301. Family history of congenital heart lesion	Empty/unknown
102302. Maternal systemic lupus erythematosus	Empty/unknown
102303. Family history of disorder with cardiac involvement	Empty/unknown
102304. Hereditary disorder associated with heart disease	Comorbidity
102400. Pulmonary venous abnormality: acquired	Acquired
103000. Systemic vein abnormality: acquired	Acquired
103009. Systemic vein obstruction	Acquired
103101. Superior caval vein abnormality: acquired	Acquired
103121. Inferior caval vein abnormality: acquired	Acquired
103200. Heart valvar abnormality: acquired	Acquired
103201. Tricuspid valvar abnormality: acquired	Acquired
103300. Prosthetic valve failure	Acquired
103301. Mitral valvar abnormality: acquired	Acquired
103302. Mitral stenosis: acquired	Acquired
103303. Mitral valvar stenosis: recurrent	Acquired
103304. Mitral regurgitation: acquired	Acquired
103306. Mitral regurgitation: recurrent	Acquired
103444. Left AV valvar regurgitation: acquired	Acquired
103460. AV valvar abnormality in AVSD: acquired	Acquired
103501. Pulmonary valvar abnormality: acquired	Acquired
103502. Pulmonary valvar stenosis: acquired	Acquired
103503. Pulmonary valvar stenosis: recurrent	Acquired
103504. Pulmonary regurgitation: acquired	Acquired

**TABLE 9** Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (*continued*)

CCAD official diagnosis codes	Diagnosis group in scheme 1
103601. Aortic valvar abnormality: acquired	Acquired
103602. Aortic valvar stenosis: acquired	Acquired
103603. Aortic valvar stenosis: recurrent	Acquired
103604. Aortic regurgitation: acquired	Acquired
103606. Aortic regurgitation: recurrent	Acquired
103701. Truncal valvar abnormality: acquired	Acquired
104001. Syncope	Empty/unknown
104030. Hypotension	Empty/unknown
105101. Common arterial trunk (truncus) abnormality: acquired	Acquired
109001. Traumatic injury of heart	Acquired
110000. Arrhythmia	Acquired
110021. Cardiac arrest	Acquired
110100. Supraventricular tachycardia	Acquired
110101. Supraventricular rhythm disturbance	Acquired
110203. Sinus node dysfunction (including sick sinus)	Acquired
110204. Sinus bradycardia	Acquired
110207. Sinus tachycardia	Acquired
110305. Paroxysmal atrial tachycardia	Acquired
110307. Atrial flutter	Acquired
110308. Atrial fibrillation	Acquired
110312. Focal atrial tachycardia: ectopic (automatic)	Acquired
110313. Macro-re-entrant atrial tachycardia (including atrial flutter)	Acquired
110321. Premature atrial beats (complexes/contractions)	Acquired
110400. Rhythm disturbance at level of AV junction	Acquired
110407. AV junctional (nodal) tachycardia	Acquired
110411. AV nodal re-entry tachycardia	Acquired
110500. Ventricular rhythm disturbance	Acquired
110506. Ventricular tachycardia	Acquired
110509. Ventricular flutter	Acquired
110510. Ventricular fibrillation	Acquired
110521. Premature ventricular beats (complexes/contractions)	Acquired
110550. Non-sustained ventricular tachycardia	Acquired
110600. Conduction disturbance	Acquired
110601. Sinoatrial block	Acquired
110602. First-degree AV block	Acquired
110603. Second-degree AV block	Acquired
110607. Complete AV block (third degree)	Acquired

*continued*

TABLE 9 Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (continued)

CCAD official diagnosis codes	Diagnosis group in scheme 1
110610. Acquired complete AV block	Acquired
110616. Congenital complete heart block	Acquired
110623. Complete R bundle branch block	Acquired
110624. Complete L bundle branch block	Acquired
110701. AV reciprocating (re-entry) tachycardia: manifest pre-excitation in sinus rhythm (Wolff–Parkinson–White syndrome)	Acquired
110706. Accessory pathway: retrograde conduction only (concealed: no pre-excitation sinus rhythm)	Acquired
110711. Manifest accessory pathway	Acquired
110714. Permanent junctional reciprocating tachycardia	Acquired
110722. AV reciprocating (re-entry) tachycardia: orthodromic	Acquired
110723. AV re-entry (reciprocating) tachycardia: antidromic (typically wide QRS)	Acquired
111100. Pacemaker dysfunction: complication necessitating replacement	Acquired
111101. Pacemaker dysfunction: complication	Acquired
111103. Pacemaker battery exhaustion: end of life	Acquired
111117. Pacemaker: implantable cardioverter and defibrillator loss of capture	Acquired
111140. Pacemaker lead dysfunction: complication	Acquired
111159. Pacemaker generator site local complication	Acquired
111160. Implantable cardioverter and defibrillator dysfunction: complication	Acquired
111201. Prolonged QT interval	Acquired
111229. Long QT syndrome	Acquired
112000. Electrocardiogram abnormality	Acquired
130010. Electrocardiogram	Empty/unknown
130021. Chest radiography	Empty/unknown
130023. Computerised tomographic scan of chest	Empty/unknown
130024. Cardiovascular magnetic resonance imaging	Empty/unknown
130100. Echocardiographic examination	Empty/unknown
130102. Transthoracic echocardiographic examination	Empty/unknown
130103. Transoesophageal echocardiographic examination	Empty/unknown
130501. Diagnostic cardiovascular catheterisation procedure	Procedure
140101. Chromosomal anomaly	Comorbidity
140102. Trisomy 21: Down syndrome	Comorbidity
140104. Trisomy 13: Patau syndrome	Comorbidity
140105. 45XO: Turner syndrome	Comorbidity
140121. 22q11 microdeletion (CATCH 22)	Comorbidity
140200. Syndrome/association with cardiac involvement	Comorbidity
140206. DiGeorge sequence	Comorbidity
140217. Marfan syndrome	Comorbidity
140219. Noonan syndrome	Comorbidity

**TABLE 9** Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (*continued*)

CCAD official diagnosis codes	Diagnosis group in scheme 1
140228. Tuberous sclerosis	Comorbidity
140230. Williams syndrome (infantile hypercalcaemia)	Comorbidity
140232. Fetal rubella syndrome	Comorbidity
140266. Alagille syndrome: arteriohepatic dysplasia	Comorbidity
140300. Non-cardiac abnormality associated with heart disease	Comorbidity
140304. Non-cardiothoracic vascular abnormality	Comorbidity
140305. Psychomotor developmental delay	Comorbidity
140306. Cystic fibrosis	Comorbidity
140307. Congenital diaphragmatic hernia	Comorbidity
140308. Tracheo-oesophageal fistula	Comorbidity
140310. Omphalocele	Comorbidity
140311. Duodenal stenosis/atresia	Comorbidity
140323. Renal abnormality	Comorbidity
140329. Thoracic/mediastinal abnormality	Comorbidity
140333. Microcephaly	Comorbidity
140349. Tracheobronchial malacia	Comorbidity
140404. Pectus carinatum	Comorbidity
140405. Pectus excavatum	Comorbidity
140409. Kyphoscoliosis	Comorbidity
140412. Cleft lip or palate	Comorbidity
140414. Anterior chest wall (pectus) deformity	Comorbidity
140501. Maternal teratogen associated with CHD	Comorbidity
140601. Multiple congenital malformations	Comorbidity
150401. Postprocedural superior caval vein complication	Acquired
150405. Postprocedural inferior caval vein complication	Acquired
150415. Postprocedural femoral vein complication	Empty/unknown
150434. Postprocedural major vein complication	Empty/unknown
150501. Postprocedural pulmonary vein complication	Acquired
150503. Pulmonary vein obstruction	Acquired
151010. Right atrial abnormality: acquired	Acquired
151011. Postprocedural right atrial complication	Acquired
151013. Obstruction of right atrial conduit (total cavopulmonary connection)	Acquired
151020. Left atrial abnormality: acquired	Acquired
151021. Postprocedural left atrial complication	Acquired
151061. Postprocedural atrial septum complication	Acquired
151063. Residual interatrial communication ('ASD')	Acquired
151066. Ineffective balloon atrial septostomy	Acquired

continued

TABLE 9 Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (continued)

CCAD official diagnosis codes	Diagnosis group in scheme 1
151100. Postprocedural tricuspid valvar complication	Acquired
151103. Residual tricuspid regurgitation	Acquired
151108. Tricuspid valvar prosthesis complication	Acquired
151200. Postprocedural mitral valvar complication	Acquired
151201. Residual mitral valvar stenosis	Acquired
151203. Residual mitral regurgitation	Acquired
151209. Mitral valvar prosthesis complication	Acquired
151302. Residual common AV valvar regurgitation	Acquired
151400. Postprocedural right AV valvar complication	Acquired
151500. Postprocedural left AV valvar complication	Acquired
151600. Postprocedural AV septal defect complication	Acquired
151602. Residual ventricular component of AVSD	Acquired
152001. Postprocedural right ventricular complication	Acquired
152021. Postprocedural right ventricular outflow tract complication	Acquired
152023. Residual right ventricular outflow tract obstruction	Acquired
152025. Aneurysm of right ventricular outflow tract patch	Acquired
152075. Residual subaortic stenosis in complex heart disease	Acquired
152076. Residual subpulmonary stenosis in complex heart disease	Acquired
152101. Postprocedural left ventricular complication	Acquired
152121. Postprocedural left ventricular outflow tract complication	Acquired
152202. Residual VSD	Acquired
152503. Residual truncal regurgitation	Acquired
153000. Postprocedural pulmonary valvar complication	Acquired
153001. Residual pulmonary valvar stenosis	Acquired
153003. Residual pulmonary regurgitation	Acquired
153008. Pulmonary valvar prosthesis complication	Acquired
153201. Postprocedural pulmonary trunk complication	Acquired
153221. Postprocedural right pulmonary artery complication	Acquired
153223. Residual right pulmonary artery stenosis	Acquired
153241. Postprocedural left pulmonary artery complication	Acquired
153243. Residual left pulmonary artery stenosis	Acquired
153500. Postprocedural aortic valvar complication	Acquired
153501. Residual aortic valvar stenosis	Acquired
153503. Residual aortic regurgitation	Acquired
153508. Aortic valvar prosthesis complication	Acquired
153601. Postprocedural ascending aorta complication	Acquired
153701. Postprocedural descending aorta complication	Acquired
153705. Residual aortic coarctation	Acquired



**TABLE 9** Mapping from the CCAD official diagnostic codes to our diagnostic scheme 1 (*continued*)

CCAD official diagnosis codes	Diagnosis group in scheme 1
153707. Postprocedural aneurysm of aorta at coarctation site	Acquired
153901. Postprocedural arterial duct complication	Acquired
153902. Residual arterial duct (PDA) patency	Acquired
153950. Postprocedural systemic-to-pulmonary collateral artery complication	Acquired
154100. Postprocedural coronary arterial complication	Acquired
154113. Cardiac transplant-associated coronary allograft vasculopathy	Acquired
155500. Cardiac conduit complication	Acquired
155516. Cardiac conduit failure	Acquired
155600. Systemic-to-pulmonary arterial shunt complication	Acquired
155601. Systemic-to-pulmonary arterial shunt partial obstruction	Acquired
155602. Systemic-to-pulmonary arterial shunt complete obstruction	Acquired
158300. Pericardial effusion requiring drainage	Acquired
159500. Complication after heart or lung transplant	Empty/unknown
159564. Post-lung transplant obliterative bronchiolitis	Empty/unknown
159566. Lung disease in lung transplant rejection	Empty/unknown
160101. Pneumothorax	Acquired
160104. Pleural effusion	Acquired
160107. Chylothorax	Acquired
160111. Empyema	Acquired
160121. Pleural disease: benign	Empty/unknown
160122. Pleural disease: malignant	Empty/unknown
160301. Lung disease: benign	Empty/unknown
160302. Lower respiratory tract infection	Empty/unknown
160305. Lung disease	Comorbidity
160321. Lung disease: malignant	Empty/unknown
160511. Mediastinal disease: benign	Empty/unknown
160512. Mediastinal disease: malignant	Empty/unknown
161001. Tracheal stenosis	Comorbidity
161009. Tracheal disease	Comorbidity
161509. Diaphragm disease	Empty/unknown
162001. Oesophageal disease: benign	Empty/unknown
162002. Oesophageal disease: malignant	Empty/unknown
163001. Respiratory failure	Comorbidity
Q19051. Status: post procedure	Empty/unknown
Q19067. Diagnosis uncertain	Empty/unknown

ASD, atrial septal defect; AV, atrioventricular; DORV, double-outlet right ventricle; IVS, intact ventricular septum; MAPCA, major aortopulmonary collateral arteries; TAPVC, totally anomalous pulmonary venous connection; VA, ventriculoarterial.

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
010100. Normal heart	Normal	0	0	0	0	0
010101. Tetralogy of Fallot	Congenital	1	0	0	0	0
010102. TGA (concordant AV and discordant VA connections) and IVS	Congenital	1	0	0	0	0
010103. Congenitally corrected TGA (discordant AV and VA connections)	Congenital	1	0	0	0	0
010104. DORV	Congenital	1	0	0	0	0
010106. Pulmonary atresia + VSD (including Fallot type)	Congenital	1	0	0	0	0
010107. Pulmonary atresia + IVS	Congenital	0	1	0	0	0
010109. HLHS	Congenital	0	1	1	0	0
010114. Double inlet AV connection (double inlet ventricle)	Congenital	0	1	0	0	0
010116. Partially anomalous pulmonary venous connections: Scimitar syndrome	Congenital	1	0	0	0	0
010117. DORV: Fallot type (subaortic or doubly committed VSD and pulmonary stenosis)	Congenital	1	0	0	0	0
010118. DORV: transposition type (subpulmonary VSD)	Congenital	1	0	0	0	0
010119. DORV: with non-committed VSD	Congenital	0	1	0	0	0
010120. AV septal defect and tetralogy of Fallot	Congenital	1	0	0	0	0
010122. Functionally UVH	Congenital	0	1	0	0	0
010124. DORV: with IVS	Congenital	0	1	0	0	0
010125. Pulmonary atresia + VSD + systemic-to-pulmonary collateral artery(ies) [MAPCA(s)]	Congenital	1	0	0	0	0
010139. Cardiac abnormality	Congenital	1	0	0	0	0
010140. DORV: subaortic or doubly committed VSD without pulmonary stenosis ('VSD type')	Congenital	1	0	0	0	0
010160. Vascular abnormality	Congenital	1	0	0	0	0

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
010300. Usual atrial arrangement (atrial situs solitus)	Normal	0	0	0	0	0
010306. Abnormal atrial arrangement	Congenital	0	0	0	0	1
010309. AV and/or VA connections abnormal	Congenital	0	1	0	0	0
010310. Normal atrial arrangement (situs), AV and VA connections	Normal	0	0	0	0	0
010403. Double inlet right ventricle	Congenital	0	1	0	0	0
010404. Double inlet left ventricle	Congenital	0	1	0	0	0
010500. Concordant VA connections	Normal	0	0	0	0	0
010501. Discordant VA connections (TGA)	Congenital	0	0	0	0	1
010503. Double outlet left ventricle	Congenital	1	0	0	0	0
010510. Concordant VA connections with parallel great arteries (anatomically corrected malposition)	Congenital	1	0	0	0	0
020102. Dextrocardia: heart predominantly in right hemithorax	Congenital	0	0	0	0	1
020109. Position/orientation of heart abnormal	Congenital	0	0	0	0	1
020305. Solitary ventricle of indeterminate morphology	Congenital	0	1	0	0	0
030103. Total mirror imagery (situs inversus)	Congenital	0	0	0	0	1
030104. Right isomerism ('asplenia')	Congenital	0	0	0	0	1
030105. Left isomerism ('polysplenia')	Congenital	0	0	0	0	1
030109. Position or morphology of thoraco-abdominal organs abnormal	Comorbidity	0	0	0	0	0
030305. Tracheobronchial anomaly	Comorbidity	0	0	0	0	0
030703. Spleen absent (asplenia)	Comorbidity	0	0	0	0	0

continued

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
030704. Multiple spleens (polysplenia)	Comorbidity	0	0	0	0	0
040100. Superior caval vein abnormality	Congenital	0	0	0	0	1
040101. Left superior caval vein persisting to coronary sinus	Congenital	0	0	0	0	1
040200. Hepatic vein abnormality	Congenital	0	0	0	0	1
040300. Inferior caval vein abnormality	Congenital	0	0	0	0	1
040310. Inferior caval vein interruption (absent suprarenal segment) with azygos continuation	Congenital	0	0	0	0	1
040400. Coronary sinus abnormality	Congenital	1	0	0	0	0
040500. Systemic vein abnormality: congenital	Congenital	0	0	0	0	1
040600. Totally anomalous pulmonary venous connection: supracardiac	Congenital	1	0	0	0	0
040701. Partially anomalous pulmonary venous connection(s)	Congenital	1	0	0	0	0
040800. Pulmonary vein abnormality	Congenital	1	0	0	0	0
040805. Totally anomalous pulmonary venous connection	Congenital	1	0	0	0	0
040806. Obstructed pulmonary venous connection(s)	Congenital	1	0	0	0	0
040810. Totally anomalous pulmonary venous connection: intracardiac	Congenital	1	0	0	0	0
040820. Totally anomalous pulmonary venous connection: infracardiac	Congenital	1	0	0	0	0
040830. Totally anomalous pulmonary venous connection: mixed	Congenital	1	0	0	0	0
040891. Pulmonary vein stenosis	Acquired	1	0	0	0	0
050100. Right atrial abnormality	Congenital	1	0	0	0	0
050200. Left atrial abnormality	Congenital	1	0	0	0	0

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
050201. Cor triatriatum (divided left atrium)	Congenital	1	0	0	0	0
050202. Supravalvar mitral ring	Congenital	1	0	0	0	0
050300. Atrial septum abnormality	Congenital	1	0	0	0	0
050301. Patent foramen ovale	Congenital	1	0	0	0	0
050310. Intact atrial septum (no interatrial communication)	Normal	0	0	0	0	0
050401. Interatrial communication ('ASD')	Congenital	1	0	0	0	0
050402. ASD within oval fossa (secundum)	Congenital	1	0	0	0	0
050403. Spontaneous closure of ASD within oval fossa (secundum)	Congenital	1	0	0	0	0
050500. Sinus venosus defect (ASD)	Congenital	1	0	0	0	0
050503. Interatrial communication (ASD) through coronary sinus orifice	Congenital	1	0	0	0	0
050601. Common atrium (virtual absence of atrial septum)	Congenital	1	0	0	0	0
060100. Tricuspid valvar abnormality	Congenital	1	0	0	0	0
060101. Tricuspid atresia	Congenital	0	1	0	0	0
060103. Tricuspid valvar dysplasia	Congenital	1	0	0	0	0
060109. Straddling tricuspid valve	Congenital	0	0	0	0	1
060125. Tricuspid regurgitation: congenital	Congenital	1	0	0	0	0
060134. Ebstein's malformation of tricuspid valve	Congenital	1	0	0	0	0
060191. Tricuspid regurgitation	Congenital	1	0	0	0	0
060192. Tricuspid stenosis	Congenital	1	0	0	0	0
060200. Mitral valvar abnormality	Congenital	1	0	0	0	0
060201. Mitral atresia	Congenital	0	1	1	0	0

continued

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
060207. Mitral valvar stenosis: congenital	Congenital	1	0	0	0	0
060209. Straddling mitral valve	Congenital	0	0	0	0	1
060212. Mitral subvalvar apparatus abnormality	Congenital	1	0	0	0	0
060213. Mitral subvalvar stenosis	Congenital	1	0	0	0	0
060225. Mitral regurgitation: congenital	Congenital	1	0	0	0	0
060235. Mitral valvar prolapse	Congenital	1	0	0	0	0
060236. True cleft of mitral leaflet (without AVSD)	Congenital	1	0	0	0	0
060256. Parachute malformation of mitral valve	Congenital	1	0	0	0	0
060291. Mitral regurgitation	Congenital	1	0	0	0	0
060292. Mitral stenosis	Congenital	1	0	0	0	0
060293. Mitral valve stenosis	Congenital	1	0	0	0	0
060501. AVSD AV valvar abnormality	Congenital	1	0	0	0	0
060506. AVSD AV valvar regurgitation	Congenital	1	0	0	0	0
060600. AVSD	Congenital	1	0	0	0	0
060601. AVSD: isolated atrial component (primum ASD) (partial)	Congenital	1	0	0	0	0
060608. AVSD: isolated ventricular component	Congenital	1	0	0	0	0
060609. AVSD: atrial and ventricular components with common AV orifice (complete)	Congenital	1	0	0	0	0
060610. AVSD: atrial and (restrictive) ventricular components + separate AV valves ('intermediate')	Congenital	1	0	0	0	0
060726. AVSD with ventricular imbalance	Congenital	0	1	0	0	0
070100. Right ventricular abnormality	Congenital	1	0	0	0	0
070110. Arrhythmogenic right ventricular cardiomyopathy	Acquired	0	0	0	1	0

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
070111. Right ventricular dysfunction	Acquired	0	0	0	1	0
070114. Right ventricular aneurysm	Congenital	1	0	0	0	0
070200. Right ventricular hypoplasia	Congenital	0	0	0	0	1
070301. Double-chambered right ventricle	Congenital	1	0	0	0	0
070501. Right ventricular outflow tract obstruction	Congenital	1	0	0	0	0
070530. Subpulmonary stenosis	Congenital	1	0	0	0	0
070600. Left ventricular abnormality	Congenital	1	0	0	0	0
070610. Left ventricular dysfunction	Acquired	0	0	0	1	0
070613. Left ventricular aneurysm	Congenital	1	0	0	0	0
070700. Left ventricular hypoplasia	Congenital	0	0	0	0	1
070841. Functionally UVH	Congenital	0	1	0	0	0
070842. Functionally UVH	Congenital	0	1	0	0	0
070850. Ventricular myocardial non-compaction cardiomyopathy	Acquired	0	0	0	1	0
070900. Subaortic stenosis	Congenital	1	0	1	0	0
070901. Left ventricular outflow tract obstruction	Congenital	1	0	1	0	0
070903. Subaortic stenosis due to fibromuscular shelf	Congenital	1	0	1	0	0
070931. Aortic abnormality	Congenital	1	0	0	0	0
071000. VSD	Congenital	1	0	0	0	0
071001. Perimembranous VSD	Congenital	1	0	0	0	0
071012. VSD + malaligned outlet septum	Congenital	1	0	0	0	0
071101. Muscular VSD	Congenital	1	0	0	0	0
071200. Subarterial VSD	Congenital	1	0	0	0	0
071201. Doubly committed subarterial VSD	Congenital	1	0	0	0	0

continued

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
071402. Communication between left ventricle and right atrium (Gerbode defect)	Congenital	1	0	0	0	0
071405. Inlet VSD	Congenital	1	0	0	0	0
071501. Tiny VSD (Maladie de Roger)	Congenital	1	0	0	0	0
071504. Multiple VSDs	Congenital	1	0	0	0	0
071505. Single VSD	Congenital	1	0	0	0	0
071601. Spontaneous closure of VSD	Normal	0	0	0	0	0
072000. Ventricular septal abnormality	Congenital	1	0	0	0	0
072001. Aneurysm of membranous septum	Normal	0	0	0	0	0
072100. IVS	Normal	0	0	0	0	1
090101. Common arterial trunk (truncus arteriosus)	Congenital	1	0	0	0	0
090200. Truncal valvar abnormality	Congenital	1	0	0	0	0
090203. Truncal valvar regurgitation	Congenital	1	0	0	0	0
090401. Aortopulmonary window	Congenital	1	0	0	0	0
090500. Pulmonary valvar abnormality	Congenital	1	0	0	0	0
090501. Pulmonary valvar stenosis	Congenital	1	0	0	0	0
090504. Pulmonary valvar stenosis: congenital	Congenital	1	0	0	0	0
090511. Pulmonary atresia	Congenital	1	0	0	0	0
090512. Pulmonary atresia: imperforate valve	Congenital	1	0	0	0	0
090515. Pulmonary valvar atresia: acquired	Acquired	0	0	0	1	0
090522. Pulmonary regurgitation: congenital	Acquired	1	0	0	0	0
090525. Absent pulmonary valve syndrome: Fallot type	Congenital	1	0	0	0	0
090591. Pulmonary regurgitation	Acquired	1	0	0	0	0
090592. Pulmonary stenosis	Congenital	1	0	0	0	0



TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
090700. Pulmonary trunk (main pulmonary artery) abnormality	Acquired	1	0	0	0	0
090711. Pulmonary trunk hypoplasia	Acquired	1	0	0	0	0
090713. Supravalvar pulmonary trunk stenosis	Acquired	1	0	0	0	0
090726. Solitary arterial trunk (absent intrapericardial pulmonary arteries)	Congenital	1	0	0	0	0
090801. Major systemic-to-pulmonary collateral artery(ies) [MAPCA(s)]	Congenital	1	0	0	0	0
090906. Pulmonary arterial sling	Congenital	1	0	0	0	0
090908. Pulmonary artery from ascending aorta (hemitruncus)	Congenital	1	0	0	0	0
091000. Pulmonary arterial abnormality	Acquired	1	0	0	0	0
091001. Pulmonary arterial stenosis	Acquired	1	0	0	0	0
091006. Peripheral pulmonary arterial stenoses: at/beyond hilar bifurcation	Acquired	1	0	0	0	0
091007. Central pulmonary arterial stenosis: proximal to hilar bifurcation	Acquired	1	0	0	0	0
091010. Discontinuous (non-confluent) pulmonary arteries	Congenital	1	0	0	0	0
091011. Pulmonary arterial hypoplasia	Acquired	1	0	0	0	0
091025. Right pulmonary arterial stenosis	Acquired	1	0	0	0	0
091026. Left pulmonary arterial stenosis	Acquired	1	0	0	0	0
091044. Pulmonary arterial aneurysm	Acquired	1	0	0	0	0
091500. Aortic valvar abnormality	Congenital	1	0	1	0	0
091501. Aortic valvar stenosis: congenital	Congenital	1	0	1	0	0
091503. Aortic atresia	Congenital	0	1	1	0	0
091507. Aortic regurgitation: congenital	Congenital	1	0	0	0	0

continued

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
091512. Eccentric opening of tricuspid aortic valve	Congenital	1	0	1	0	0
091513. Aortic valvar stenosis	Congenital	1	0	1	0	0
091522. Bicuspid aortic valve	Congenital	1	0	0	0	0
091530. Aortic valvar prolapse	Congenital	1	0	0	0	0
091591. Aortic regurgitation	Congenital	1	0	0	0	0
091592. Aortic stenosis	Congenital	1	0	1	0	0
091600. Supravalvar aortic stenosis	Congenital	1	0	1	0	0
091602. Ascending aorta hypoplasia	Congenital	0	0	1	0	1
091605. Ascending aorta dilatation associated with Marfan syndrome	Acquired	1	0	0	0	0
091609. Ascending aorta dilatation	Acquired	1	0	0	0	0
091610. Ascending aorta abnormality	Congenital	1	0	0	0	0
091613. Aortic root dilatation	Acquired	0	0	0	1	0
091701. Aortoventricular tunnel	Congenital	1	0	0	0	0
091702. Aorto-left ventricular tunnel	Congenital	1	0	0	0	0
091801. Aortic sinus of Valsalva aneurysm	Congenital	1	0	0	0	0
091901. Arteriovenous fistula (malformation)	Congenital	1	0	0	0	0
091905. Pulmonary arteriovenous fistula (malformation)	Congenital	1	0	0	0	0
092020. Distal systemic arterial abnormality	Congenital	1	0	0	0	0
092025. Systemic-to-pulmonary collateral arter(ies) [MAPCA(s)] stenosis(es)	Congenital	1	0	0	0	0
092700. Arterial duct (ductus arteriosus) abnormality	Congenital	1	0	0	0	0
092721. Patent arterial duct (PDA)	Congenital	1	0	0	0	0

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
092800. Aortic arch abnormality	Congenital	1	0	0	0	0
092809. Double aortic arch	Congenital	1	0	0	0	0
092815. Right aortic arch	Congenital	1	0	0	0	0
092816. Descending aorta dilatation	Acquired	1	0	0	0	0
092901. Aortic coarctation	Congenital	1	0	1	0	0
092911. Aortic arch hypoplasia (tubular)	Congenital	1	0	1	0	0
092916. Descending abdominal aorta hypoplasia (middle aortic syndrome)	Congenital	1	0	0	0	0
092931. Interrupted aortic arch	Congenital	1	0	1	0	0
093000. Aortic arch branch abnormality	Congenital	1	0	0	0	0
093002. Aberrant origin right subclavian artery	Congenital	0	0	0	0	1
093004. Aberrant origin left subclavian artery	Congenital	0	0	0	0	1
093100. Vascular ring	Congenital	1	0	0	0	0
094101. Anomalous origin of coronary artery from pulmonary artery	Congenital	1	0	0	0	0
094200. Coronary artery: anomalous aortic origin or course	Congenital	1	0	0	0	0
094305. Intramural proximal coronary arterial course	Congenital	1	0	0	0	0
094501. Coronary fistula	Congenital	1	0	0	0	0
094511. Coronary fistulas from right ventricle ('sinusoidal')	Congenital	1	0	0	0	0
094600. Coronary arterial abnormality	Congenital	1	0	0	0	0
094601. Coronary arterial aneurysm(s)	Acquired	1	0	0	0	0
094606. Right ventricle-dependent coronary circulation	Congenital	1	0	0	0	0
100100. Pericardial abnormality	Congenital	1	0	0	0	0
100301. Heart tumour	Acquired	0	0	0	1	0

continued

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
100501. Acute rheumatic fever	Acquired	0	0	0	1	0
100521. Rheumatic fever with cardiac involvement	Acquired	0	0	0	1	0
100530. Rheumatic valvar disease	Acquired	0	0	0	1	0
100531. Rheumatic mitral valvar disease	Acquired	0	0	0	1	0
100533. Rheumatic aortic valvar disease	Acquired	0	0	0	1	0
100601. Infective endocarditis	Acquired	0	0	0	1	0
100620. Heart abscess	Acquired	0	0	0	1	0
100641. Bacterial endocarditis	Acquired	0	0	0	1	0
100664. Postprocedural endocarditis	Acquired	0	0	0	1	0
100701. Infectious myocarditis	Acquired	0	0	0	1	0
100703. Viral myocarditis	Acquired	0	0	0	1	0
100705. Drug-induced heart muscle disease	Acquired	0	0	0	1	0
100708. Trypanosomal myocarditis (Chagas' disease)	Acquired	0	0	0	1	0
100740. Myocardial failure in end-stage CHD	Acquired	0	0	0	1	0
100742. Heart muscle disease in cardiac rejection	Acquired	0	0	0	1	0
100761. Nutritional heart muscle disease	Acquired	0	0	0	1	0
100771. Heart muscle disease in infant of diabetic mother	Acquired	0	0	0	1	0
100781. Heart muscle disease in collagen vascular/connective tissue disorder	Acquired	0	0	0	1	0
100800. Pericarditis	Acquired	0	0	0	1	0
100801. Infectious pericarditis	Acquired	0	0	0	1	0
100803. Viral pericarditis	Acquired	0	0	0	1	0
100804. Bacterial pericarditis	Acquired	0	0	0	1	0
100809. Constrictive pericarditis	Acquired	0	0	0	1	0

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
100813. Cardiac tamponade	Acquired	0	0	0	1	0
100829. Pericardial abnormality: acquired	Acquired	0	0	0	1	0
100831. Pericardial effusion	Acquired	0	0	0	1	0
100901. Kawasaki disease	Acquired	0	0	0	1	0
100902. Kawasaki disease with aneurysm(s) or dilated coronary vessels	Acquired	0	0	0	1	0
100908. Kawasaki disease without cardiac involvement	Acquired	0	0	0	1	0
100910. Acquired coronary arterial disease	Acquired	0	0	0	1	0
100930. Ischaemic heart disease	Acquired	0	0	0	1	0
101001. Cardiomyopathy	Acquired	0	0	0	1	0
101011. Idiopathic restrictive cardiomyopathy	Acquired	0	0	0	1	0
101012. Endocardial fibroelastosis	Acquired	0	0	0	1	0
101013. Infiltrative cardiomyopathy	Acquired	0	0	0	1	0
101020. Hypertrophic cardiomyopathy	Acquired	0	0	0	1	0
101025. Dilated cardiomyopathy	Acquired	0	0	0	1	0
101201. Innocent murmur	Normal	0	0	0	0	0
101239. Failure to thrive	Acquired	0	0	0	1	0
101242. Chest pain	Empty/ unknown	0	0	0	0	0
101301. Pulmonary arterial hypertension	Acquired	0	0	0	1	0
101302. Primary pulmonary hypertension	Acquired	0	0	0	1	0
101306. Pulmonary vascular disease	Acquired	0	0	0	1	0
101308. Irreversible pulmonary vascular disease due to CHD (Eisenmenger syndrome)	Acquired	0	0	0	1	0
101320. Secondary pulmonary hypertension	Acquired	0	0	0	1	0

continued

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
101321. Pulmonary hypertension due to left-to-right shunt	Acquired	0	0	0	1	0
101350. Pulmonary arterial disease: acquired	Acquired	0	0	0	1	0
101351. Pulmonary embolism	Acquired	0	0	0	1	0
101400. Secondary systemic hypertension	Comorbidity	0	0	0	0	0
101401. Systemic hypertension	Acquired	0	0	0	1	0
101402. Primary (essential) systemic hypertension	Comorbidity	0	0	0	0	0
101404. Systemic hypertension due to aortic arch obstruction	Acquired	0	0	0	1	0
101440. Ascending aorta dilatation: acquired	Acquired	0	0	0	1	0
101442. Ascending aortic aneurysm	Acquired	0	0	0	1	0
101443. Descending aortic aneurysm	Acquired	0	0	0	1	0
101444. Abdominal aortic aneurysm	Comorbidity	0	0	0	0	0
101445. Rupture of thoracic aortic aneurysm	Comorbidity	0	0	0	0	0
101446. Rupture of abdominal aortic aneurysm	Comorbidity	0	0	0	0	0
101450. Aortic aneurysm	Acquired	0	0	0	1	0
101451. Aortic dissection	Acquired	0	0	0	1	0
101452. Ascending aorta dissection and propagation beyond arch (DeBakey type I)	Acquired	0	0	0	1	0
101453. Ascending aorta dissection not beyond arch (DeBakey type II/Stanford type A)	Acquired	0	0	0	1	0
101454. Descending aorta dissection and distal propagation (DeBakey type III/Stanford type B)	Comorbidity	0	0	0	0	0
101460. Systemic arteritis	Comorbidity	0	0	0	0	0
101470. Abnormality of aorta: acquired	Acquired	0	0	0	1	0

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
101472. Recoarctation of aorta	Acquired	0	0	0	1	0
101477. Supravalvar aortic stenosis: acquired	Acquired	0	0	0	1	0
101480. Arterial duct (ductus arteriosus) abnormality: acquired	Acquired	0	0	0	1	0
101500. Neonatal disorder	Empty/unknown	0	0	0	0	0
101501. Persistent pulmonary hypertension of the newborn (persistent fetal circulation)	Acquired	0	0	0	1	0
101505. Necrotising enterocolitis	Comorbidity	0	0	0	0	0
101510. Transient myocardial ischaemia	Acquired	0	0	0	1	0
101512. Meconium aspiration	Comorbidity	0	0	0	0	0
101600. Right ventricular abnormality: acquired	Acquired	0	0	0	1	0
101608. Right ventricular congestive heart failure	Acquired	0	0	0	1	0
101616. Right ventricular outflow tract obstruction: acquired	Acquired	0	0	0	1	0
101640. Left ventricular abnormality: acquired	Acquired	0	0	0	1	0
101646. Recurrent left ventricular outflow tract obstruction	Acquired	0	0	0	1	0
101647. Left ventricular failure	Acquired	0	0	0	1	0
101660. Abnormality associated with ventricular septum: acquired	Acquired	0	0	0	1	0
101662. Postmyocardial infarct VSD	Acquired	0	0	0	1	0
101681. Narrowing of constructed intraventricular tunnel: acquired	Acquired	0	0	0	1	0
101682. Subaortic stenosis in complex heart disease: acquired	Acquired	0	0	0	1	0

continued

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
101683. Subpulmonary stenosis in complex heart disease: acquired	Acquired	0	0	0	1	0
101700. Symptom/sign of heart disease	Empty/unknown	0	0	0	0	0
101703. Cyanosis	Empty/unknown	0	0	0	0	0
101705. Heart failure	Acquired	0	0	0	1	0
101712. Cyanotic spells	Acquired	0	0	0	1	0
101713. Palpitations	Acquired	0	0	0	1	0
101740. Atrial septum abnormality: acquired	Acquired	0	0	0	1	0
101800. Myocardial infarction	Acquired	0	0	0	1	0
101801. Acute myocardial infarction	Acquired	0	0	0	1	0
101901. Dyslipidaemia	Empty/unknown	0	0	0	0	0
102000. No preprocedural risk factors	Normal	0	0	0	0	0
102002. Preprocedural shock	Comorbidity	0	0	0	0	0
102003. Preprocedural arrhythmia	Comorbidity	0	0	0	0	0
102007. Preprocedural renal failure (creatinine > 176 µmol/l)	Comorbidity	0	0	0	0	0
102009. Preprocedural septicaemia	Comorbidity	0	0	0	0	0
102012. Preprocedural neurological impairment	Comorbidity	0	0	0	0	0
102014. Preprocedural mechanical ventilatory support	Comorbidity	0	0	0	0	0
102015. Preprocedural mechanical circulatory support	Comorbidity	0	0	0	0	0
102016. Preprocedural pulmonary hypertension	Comorbidity	0	0	0	0	0
102017. Preprocedural tracheostomy	Comorbidity	0	0	0	0	0
102018. Preprocedural seizures	Comorbidity	0	0	0	0	0
102019. Preprocedural risk factor	Empty/unknown	0	0	0	0	0
102202. Premature birth	Comorbidity	0	0	0	0	0



TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
102203. Infant of diabetic mother	Comorbidity	0	0	0	0	0
102205. Premature birth 32–35 weeks	Comorbidity	0	0	0	0	0
102206. Premature birth < 32 weeks	Comorbidity	0	0	0	0	0
102300. Hereditary/non-cardiac abnormality not apparent	Empty/unknown	0	0	0	0	0
102301. Family history of congenital heart lesion	Empty/unknown	0	0	0	0	0
102302. Maternal systemic lupus erythematosus	Empty/unknown	0	0	0	0	0
102303. Family history of disorder with cardiac involvement	Empty/unknown	0	0	0	0	0
102304. Hereditary disorder associated with heart disease	Comorbidity	0	0	0	0	0
102400. Pulmonary venous abnormality: acquired	Acquired	0	0	0	1	0
103000. Systemic vein abnormality: acquired	Acquired	0	0	0	1	0
103009. Systemic vein obstruction	Acquired	0	0	0	1	0
103101. Superior caval vein abnormality: acquired	Acquired	0	0	0	1	0
103121. Inferior caval vein abnormality: acquired	Acquired	0	0	0	1	0
103200. Heart valvar abnormality: acquired	Acquired	0	0	0	1	0
103201. Tricuspid valvar abnormality: acquired	Acquired	0	0	0	1	0
103300. Prosthetic valve failure	Acquired	0	0	0	1	0
103301. Mitral valvar abnormality: acquired	Acquired	0	0	0	1	0
103302. Mitral stenosis: acquired	Acquired	0	0	0	1	0
103303. Mitral valvar stenosis: recurrent	Acquired	0	0	0	1	0
103304. Mitral regurgitation: acquired	Acquired	0	0	0	1	0

continued

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
103306. Mitral regurgitation: recurrent	Acquired	0	0	0	1	0
103444. Left AV valvar regurgitation: acquired	Acquired	0	0	0	1	0
103460. AV valvar abnormality in AVSD: acquired	Acquired	0	0	0	1	0
103501. Pulmonary valvar abnormality: acquired	Acquired	0	0	0	1	0
103502. Pulmonary valvar stenosis: acquired	Acquired	0	0	0	1	0
103503. Pulmonary valvar stenosis: recurrent	Acquired	0	0	0	1	0
103504. Pulmonary regurgitation: acquired	Acquired	0	0	0	1	0
103601. Aortic valvar abnormality: acquired	Acquired	0	0	0	1	0
103602. Aortic valvar stenosis: acquired	Acquired	0	0	0	1	0
103603. Aortic valvar stenosis: recurrent	Acquired	0	0	0	1	0
103604. Aortic regurgitation: acquired	Acquired	0	0	0	1	0
103606. Aortic regurgitation: recurrent	Acquired	0	0	0	1	0
103701. Truncal valvar abnormality: acquired	Acquired	0	0	0	1	0
104001. Syncope	Empty/ unknown	0	0	0	0	0
104030. Hypotension	Empty/ unknown	0	0	0	0	0
105101. Common arterial trunk (truncus) abnormality: acquired	Acquired	0	0	0	1	0
109001. Traumatic injury of heart	Acquired	0	0	0	1	0
110000. Arrhythmia	Acquired	0	0	0	1	0
110021. Cardiac arrest	Acquired	0	0	0	1	0
110100. Supraventricular tachycardia	Acquired	0	0	0	1	0
110101. Supraventricular rhythm disturbance	Acquired	0	0	0	1	0
110203. Sinus node dysfunction (including sick sinus)	Acquired	0	0	0	1	0
110204. Sinus bradycardia	Acquired	0	0	0	1	0

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
110207. Sinus tachycardia	Acquired	0	0	0	1	0
110305. Paroxysmal atrial tachycardia	Acquired	0	0	0	1	0
110307. Atrial flutter	Acquired	0	0	0	1	0
110308. Atrial fibrillation	Acquired	0	0	0	1	0
110312. Focal atrial tachycardia: ectopic (automatic)	Acquired	0	0	0	1	0
110313. Macro-re-entrant atrial tachycardia (including atrial flutter)	Acquired	0	0	0	1	0
110321. Premature atrial beats (complexes/contractions)	Acquired	0	0	0	1	0
110400. Rhythm disturbance at level of AV junction	Acquired	0	0	0	1	0
110407. AV junctional (nodal) tachycardia	Acquired	0	0	0	1	0
110411. AV nodal re-entry tachycardia	Acquired	0	0	0	1	0
110500. Ventricular rhythm disturbance	Acquired	0	0	0	1	0
110506. Ventricular tachycardia	Acquired	0	0	0	1	0
110509. Ventricular flutter	Acquired	0	0	0	1	0
110510. Ventricular fibrillation	Acquired	0	0	0	1	0
110521. Premature ventricular beats (complexes/contractions)	Acquired	0	0	0	1	0
110550. Non-sustained ventricular tachycardia	Acquired	0	0	0	1	0
110600. Conduction disturbance	Acquired	0	0	0	1	0
110601. Sinoatrial block	Acquired	0	0	0	1	0
110602. First-degree AV block	Acquired	0	0	0	1	0
110603. Second-degree AV block	Acquired	0	0	0	1	0
110607. Complete AV block (third degree)	Acquired	0	0	0	1	0
110610. Acquired complete AV block	Acquired	0	0	0	1	0

continued

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
110616. Congenital complete heart block	Acquired	0	0	0	1	0
110623. Complete R bundle branch block	Acquired	0	0	0	1	0
110624. Complete L bundle branch block	Acquired	0	0	0	1	0
110701. AV reciprocating (re-entry) tachycardia: manifest pre-excitation in sinus rhythm (Wolff–Parkinson–White syndrome)	Acquired	0	0	0	1	0
110706. Accessory pathway: retrograde conduction only (concealed: no pre-excitation sinus rhythm)	Acquired	0	0	0	1	0
110711. Manifest accessory pathway	Acquired	0	0	0	1	0
110714. Permanent junctional reciprocating tachycardia	Acquired	0	0	0	1	0
110722. AV reciprocating (re-entry) tachycardia: orthodromic	Acquired	0	0	0	1	0
110723. AV re-entry (reciprocating) tachycardia: antidromic (typically wide QRS)	Acquired	0	0	0	1	0
111100. Pacemaker dysfunction/complication necessitating replacement	Acquired	0	0	0	1	0
111101. Pacemaker dysfunction/complication	Acquired	0	0	0	1	0
111103. Pacemaker battery exhaustion: end of life	Acquired	0	0	0	1	0
111117. Pacemaker: implantable cardioverter and defibrillator loss of capture	Acquired	0	0	0	1	0
111140. Pacemaker lead dysfunction/complication	Acquired	0	0	0	1	0
111159. Pacemaker generator site local complication	Acquired	0	0	0	1	0
111160. Implantable cardioverter and defibrillator dysfunction/complication	Acquired	0	0	0	1	0

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
111201. Prolonged QT interval	Acquired	0	0	0	1	0
111229. Long QT syndrome	Acquired	0	0	0	1	0
112000. Electrocardiogram abnormality	Acquired	0	0	0	1	0
130010. Electrocardiogram	Empty/ unknown	0	0	0	0	0
130021. Chest radiography	Empty/ unknown	0	0	0	0	0
130023. Computerised tomographic scan of chest	Empty/ unknown	0	0	0	0	0
130024. Cardiovascular magnetic resonance imaging	Empty/ unknown	0	0	0	0	0
130100. Echocardiographic examination	Empty/ unknown	0	0	0	0	0
130102. Transthoracic echocardiographic examination	Empty/ unknown	0	0	0	0	0
130103. Transoesophageal echocardiographic examination	Empty/ unknown	0	0	0	0	0
130501. Diagnostic cardiovascular catheterisation procedure	Procedure	0	0	0	0	0
140101. Chromosomal anomaly	Comorbidity	0	0	0	0	0
140102. Trisomy 21: Down syndrome	Comorbidity	0	0	0	0	0
140104. Trisomy 13: Patau syndrome	Comorbidity	0	0	0	0	0
140105. 45XO: Turner syndrome	Comorbidity	0	0	0	0	0
140121. 22q11 microdeletion (CATCH 22)	Comorbidity	0	0	0	0	0
140200. Syndrome/association with cardiac involvement	Comorbidity	0	0	0	0	0
140206. DiGeorge sequence	Comorbidity	0	0	0	0	0
140217. Marfan syndrome	Comorbidity	0	0	0	0	0
140219. Noonan syndrome	Comorbidity	0	0	0	0	0
140228. Tuberous sclerosis	Comorbidity	0	0	0	0	0

continued

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
140230. Williams syndrome (infantile hypercalcaemia)	Comorbidity	0	0	0	0	0
140232. Fetal rubella syndrome	Comorbidity	0	0	0	0	0
140266. Alagille syndrome: arteriohepatic dysplasia	Comorbidity	0	0	0	0	0
140300. Non-cardiac abnormality associated with heart disease	Comorbidity	0	0	0	0	0
140304. Non-cardiothoracic/vascular abnormality	Comorbidity	0	0	0	0	0
140305. Psychomotor developmental delay	Comorbidity	0	0	0	0	0
140306. Cystic fibrosis	Comorbidity	0	0	0	0	0
140307. Congenital diaphragmatic hernia	Comorbidity	0	0	0	0	0
140308. Tracheo-oesophageal fistula	Comorbidity	0	0	0	0	0
140310. Omphalocele	Comorbidity	0	0	0	0	0
140311. Duodenal stenosis/atresia	Comorbidity	0	0	0	0	0
140323. Renal abnormality	Comorbidity	0	0	0	0	0
140329. Thoracic/mediastinal abnormality	Comorbidity	0	0	0	0	0
140333. Microcephaly	Comorbidity	0	0	0	0	0
140349. Tracheobronchial malacia	Comorbidity	0	0	0	0	0
140404. Pectus carinatum	Comorbidity	0	0	0	0	0
140405. Pectus excavatum	Comorbidity	0	0	0	0	0
140409. Kyphoscoliosis	Comorbidity	0	0	0	0	0
140412. Cleft lip or palate	Comorbidity	0	0	0	0	0
140414. Anterior chest wall (pectus) deformity	Comorbidity	0	0	0	0	0
140501. Maternal teratogen associated with CHD	Comorbidity	0	0	0	0	0
140601. Multiple congenital malformations	Comorbidity	0	0	0	0	0
150401. Postprocedural superior caval vein complication	Acquired	0	0	0	1	0

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
150405. Postprocedural inferior caval vein complication	Acquired	0	0	0	1	0
150415. Postprocedural femoral vein complication	Empty/unknown	0	0	0	0	0
150434. Postprocedural major vein complication	Empty/unknown	0	0	0	0	0
150501. Postprocedural pulmonary vein complication	Acquired	0	0	0	1	0
150503. Pulmonary vein obstruction	Acquired	0	0	0	1	0
151010. Right atrial abnormality: acquired	Acquired	0	0	0	1	0
151011. Postprocedural right atrial complication	Acquired	0	0	0	1	0
151013. Obstruction of right atrial conduit (total cavopulmonary connection)	Acquired	0	0	0	1	0
151020. Left atrial abnormality: acquired	Acquired	0	0	0	1	0
151021. Postprocedural left atrial complication	Acquired	0	0	0	1	0
151061. Postprocedural atrial septum complication	Acquired	0	0	0	1	0
151063. Residual interatrial communication ('ASD')	Acquired	0	0	0	1	0
151066. Ineffective balloon atrial septostomy	Acquired	0	0	0	1	0
151100. Postprocedural tricuspid valvar complication	Acquired	0	0	0	1	0
151103. Residual tricuspid regurgitation	Acquired	0	0	0	1	0
151108. Tricuspid valvar prosthesis complication	Acquired	0	0	0	1	0
151200. Postprocedural mitral valvar complication	Acquired	0	0	0	1	0
151201. Residual mitral valvar stenosis	Acquired	0	0	0	1	0
151203. Residual mitral regurgitation	Acquired	0	0	0	1	0
151209. Mitral valvar prosthesis complication	Acquired	0	0	0	1	0

continued

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
151302. Residual common AV valvar regurgitation	Acquired	0	0	0	1	0
151400. Postprocedural right AV valvar complication	Acquired	0	0	0	1	0
151500. Postprocedural left AV valvar complication	Acquired	0	0	0	1	0
151600. Postprocedural AV septal defect complication	Acquired	0	0	0	1	0
151602. Residual ventricular component of AVSD	Acquired	0	0	0	1	0
152001. Postprocedural right ventricular complication	Acquired	0	0	0	1	0
152021. Postprocedural right ventricular outflow tract complication	Acquired	0	0	0	1	0
152023. Residual right ventricular outflow tract obstruction	Acquired	0	0	0	1	0
152025. Aneurysm of right ventricular outflow tract patch	Acquired	0	0	0	1	0
152075. Residual subaortic stenosis in complex heart disease	Acquired	0	0	0	1	0
152076. Residual subpulmonary stenosis in complex heart disease	Acquired	0	0	0	1	0
152101. Postprocedural left ventricular complication	Acquired	0	0	0	1	0
152121. Postprocedural left ventricular outflow tract complication	Acquired	0	0	0	1	0
152202. Residual VSD	Acquired	0	0	0	1	0
152503. Residual truncal regurgitation	Acquired	0	0	0	1	0
153000. Postprocedural pulmonary valvar complication	Acquired	0	0	0	1	0
153001. Residual pulmonary valvar stenosis	Acquired	0	0	0	1	0
153003. Residual pulmonary regurgitation	Acquired	0	0	0	1	0
153008. Pulmonary valvar prosthesis complication	Acquired	0	0	0	1	0



TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
153201. Postprocedural pulmonary trunk complication	Acquired	0	0	0	1	0
153221. Postprocedural right pulmonary artery complication	Acquired	0	0	0	1	0
153223. Residual right pulmonary artery stenosis	Acquired	0	0	0	1	0
153241. Postprocedural left pulmonary artery complication	Acquired	0	0	0	1	0
153243. Residual left pulmonary artery stenosis	Acquired	0	0	0	1	0
153500. Postprocedural aortic valvar complication	Acquired	0	0	0	1	0
153501. Residual aortic valvar stenosis	Acquired	0	0	0	1	0
153503. Residual aortic regurgitation	Acquired	0	0	0	1	0
153508. Aortic valvar prosthesis complication	Acquired	0	0	0	1	0
153601. Postprocedural ascending aorta complication	Acquired	0	0	0	1	0
153701. Postprocedural descending aorta complication	Acquired	0	0	0	1	0
153705. Residual aortic coarctation	Acquired	0	0	0	1	0
153707. Postprocedural aneurysm of aorta at coarctation site	Acquired	0	0	0	1	0
153901. Postprocedural arterial duct complication	Acquired	0	0	0	1	0
153902. Residual arterial duct (PDA) patency	Acquired	0	0	0	1	0
153950. Postprocedural systemic-to-pulmonary collateral artery complication	Acquired	0	0	0	1	0
154100. Postprocedural coronary arterial complication	Acquired	0	0	0	1	0
154113. Cardiac transplant-associated coronary allograft vasculopathy	Acquired	0	0	0	1	0
155500. Cardiac conduit complication	Acquired	0	0	0	1	0

continued

TABLE 10 Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (continued)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
155516. Cardiac conduit failure	Acquired	0	0	0	1	0
155600. Systemic-to-pulmonary arterial shunt complication	Acquired	0	0	0	1	0
155601. Systemic-to-pulmonary arterial shunt partial obstruction	Acquired	0	0	0	1	0
155602. Systemic-to-pulmonary arterial shunt complete obstruction	Acquired	0	0	0	1	0
158300. Pericardial effusion requiring drainage	Acquired	0	0	0	1	0
159500. Complication after heart or lung transplant	Empty/ unknown	0	0	0	0	0
159564. Post-lung transplant obliterative bronchiolitis	Empty/ unknown	0	0	0	0	0
159566. Lung disease in lung transplant rejection	Empty/ unknown	0	0	0	0	0
160101. Pneumothorax	Acquired	0	0	0	1	0
160104. Pleural effusion	Acquired	0	0	0	1	0
160107. Chylothorax	Acquired	0	0	0	1	0
160111. Empyema	Acquired	0	0	0	1	0
160121. Pleural disease: benign	Empty/ unknown	0	0	0	0	0
160122. Pleural disease: malignant	Empty/ unknown	0	0	0	0	0
160301. Lung disease: benign	Empty/ unknown	0	0	0	0	0
160302. Lower respiratory tract infection	Empty/ unknown	0	0	0	0	0
160305. Lung disease	Comorbidity	0	0	0	0	0
160321. Lung disease: malignant	Empty/ unknown	0	0	0	0	0
160511. Mediastinal disease: benign	Empty/ unknown	0	0	0	0	0
160512. Mediastinal disease: malignant	Empty/ unknown	0	0	0	0	0
161001. Tracheal stenosis	Comorbidity	0	0	0	0	0
161009. Tracheal disease	Comorbidity	0	0	0	0	0
161509. Diaphragm disease	Empty/ unknown	0	0	0	0	0
162001. Oesophageal disease: benign	Empty/ unknown	0	0	0	0	0

**TABLE 10** Mapping from the CCAD diagnosis codes to diagnosis scheme 2 (*continued*)

CCAD official diagnosis code	Diagnosis category	Definite BVH	Definite UVH	Arch obstruction?	Acquired?	Ambiguous congenital
162002. Oesophageal disease: malignant	Empty/unknown	0	0	0	0	0
163001. Respiratory failure	Comorbidity	0	0	0	0	0
Q19051. Status post procedure	Empty/unknown	0	0	0	0	0
Q19067. Diagnosis uncertain	Empty/unknown	0	0	0	0	0

ASD, atrial septal defect; AV, atrioventricular; DORV, double-outlet right ventricle; IVS, intact ventricular septum; MAPCA, major aortopulmonary collateral arteries; VA, ventriculoarterial.

Some of the information entered into the diagnostic fields in the CCAD data set was judged by the clinical members of the team to indicate the presence of a comorbidity. This comorbidity information was used along with any information in the comorbidity field to populate our risk factor 'presence of a non-Down syndrome comorbidity'. The mapping from those codes to comorbidity groups is shown in *Table 11* (note that codes for which no comorbidity is indicated are not shown).

*Table 12* gives the mapping from CCAD official comorbidity codes to our broad comorbidity groups described in *Chapter 4* (see *Defining comorbidity*).

**TABLE 11** Mapping from the CCAD diagnosis codes to comorbidity groupings

CCAD official diagnosis code	Comorbidity: premature	Comorbidity: Down syndrome	Comorbidity: congenital non-Down syndrome	Comorbidity: acquired comorbidity
030109. Position or morphology of thoraco-abdominal organs abnormal	0	0	1	0
030305. Tracheobronchial anomaly	0	0	1	0
030703. Spleen absent (asplenia)	0	0	1	0
030704. Multiple spleens (polysplenia)	0	0	1	0
101400. Secondary systemic hypertension	0	0	0	1
101402. Primary (essential) systemic hypertension	0	0	0	1
101444. Abdominal aortic aneurysm	0	0	0	1
101445. Rupture of thoracic aortic aneurysm	0	0	0	1
101446. Rupture of abdominal aortic aneurysm	0	0	0	1
101454. Descending aorta dissection and distal propagation (DeBakey type III/Stanford type B)	0	0	0	1
101460. Systemic arteritis	0	0	0	1

continued

TABLE 11 Mapping from the CCAD diagnosis codes to comorbidity groupings (continued)

CCAD official diagnosis code	Comorbidity: premature	Comorbidity: Down syndrome	Comorbidity: congenital non-Down syndrome	Comorbidity: acquired comorbidity
101505. Necrotising enterocolitis	0	0	0	1
101512. Meconium aspiration	0	0	0	1
102002. Preprocedural shock	0	0	0	1
102003. Preprocedural arrhythmia	0	0	0	1
102007. Preprocedural renal failure (creatinine > 176 µmol/l)	0	0	0	1
102009. Preprocedural septicaemia	0	0	0	1
102012. Preprocedural neurological impairment	0	0	0	1
102014. Preprocedural mechanical ventilatory support	0	0	0	1
102015. Preprocedural mechanical circulatory support	0	0	0	1
102016. Preprocedural pulmonary hypertension	0	0	0	1
102017. Preprocedural tracheostomy	0	0	0	1
102018. Preprocedural seizures	0	0	0	1
102202. Premature birth	1	0	0	0
102203. Infant of diabetic mother	0	0	0	1
102205. Premature birth 32–35 weeks	1	0	0	0
102206. Premature birth < 32 weeks	1	0	0	0
102304. Hereditary disorder associated with heart disease	0	0	1	0
140101. Chromosomal anomaly	0	0	1	0
140102. Trisomy 21: Down syndrome	0	1	0	0
140104. Trisomy 13: Patau syndrome	0	0	1	0
140105. 45XO: Turner syndrome	0	0	1	0
140121. 22q11 microdeletion (CATCH 22)	0	0	1	0
140200. Syndrome/association with cardiac involvement	0	0	1	0
140206. DiGeorge sequence	0	0	1	0
140217. Marfan syndrome	0	0	1	0
140219. Noonan syndrome	0	0	1	0
140228. Tuberous sclerosis	0	0	1	0
140230. Williams syndrome (infantile hypercalcaemia)	0	0	1	0
140232. Fetal rubella syndrome	0	0	1	0
140266. Alagille syndrome: arteriohepatic dysplasia	0	0	1	0

**TABLE 11** Mapping from the CCAD diagnosis codes to comorbidity groupings (*continued*)

CCAD official diagnosis code	Comorbidity: premature	Comorbidity: Down syndrome	Comorbidity: congenital non- Down syndrome	Comorbidity: acquired comorbidity
140300. Non-cardiac abnormality associated with heart disease	0	0	1	0
140304. Non-cardiothoracic/vascular abnormality	0	0	1	0
140305. Psychomotor developmental delay	0	0	0	1
140306. Cystic fibrosis	0	0	1	0
140307. Congenital diaphragmatic hernia	0	0	1	0
140308. Tracheo-oesophageal fistula	0	0	1	0
140310. Omphalocele	0	0	1	0
140311. Duodenal stenosis/atresia	0	0	1	0
140323. Renal abnormality	0	0	1	0
140329. Thoracic/mediastinal abnormality	0	0	1	0
140333. Microcephaly	0	0	1	0
140349. Tracheobronchial malacia	0	0	1	0
140404. Pectus carinatum	0	0	1	0
140405. Pectus excavatum	0	0	1	0
140409. Kyphoscoliosis	0	0	1	0
140412. Cleft lip or palate	0	0	1	0
140414. Anterior chest wall (pectus) deformity	0	0	1	0
140501. Maternal teratogen associated with CHD	0	0	1	0
140601. Multiple congenital malformations	0	0	1	0
160305. Lung disease	0	0	0	1
161001. Tracheal stenosis	0	0	1	0
161009. Tracheal disease	0	0	1	0
163001. Respiratory failure	0	0	0	1

TABLE 12 Mapping from the CCAD comorbidity codes to comorbidity groupings

CCAD official comorbidity code	Comorbidity group
100665. Preprocedural endocarditis	Acquired comorbidity
101505. Necrotising enterocolitis	Acquired comorbidity
101512. Meconium aspiration	Acquired comorbidity
102002. Preprocedural shock	Acquired comorbidity
102003. Preprocedural arrhythmia	Acquired comorbidity
102005. Preprocedural acidosis	Acquired comorbidity
102006. Preprocedural coagulation disorder	Acquired comorbidity
102007. Preprocedural renal failure	Acquired comorbidity
102008. Preprocedural renal failure requiring dialysis	Acquired comorbidity
102009. Preprocedural septicaemia	Acquired comorbidity
102012. Preprocedural neurological impairment	Acquired comorbidity
102014. Preprocedural mechanical ventilatory support	Acquired comorbidity
102015. Preprocedural mechanical circulatory support	Acquired comorbidity
102017. Preprocedural tracheostomy	Acquired comorbidity
110635. Preprocedural complete atrioventricular block	Acquired comorbidity
140305. Psychomotor developmental delay	Acquired comorbidity
030703. Spleen absent (asplenia)	Congenital non-Down syndrome
030704. Multiple spleens (polysplenia)	Congenital non-Down syndrome
102016. Preprocedural pulmonary hypertension	Congenital non-Down syndrome
102300. Hereditary/non-cardiac abnormality not apparent	Congenital non-Down syndrome
102304. Hereditary disorder associated with heart disease	Congenital non-Down syndrome
140101. Chromosomal anomaly	Congenital non-Down syndrome
140103. Trisomy 18: Edward syndrome	Congenital non-Down syndrome
140104. Trisomy 13: Patau syndrome	Congenital non-Down syndrome
140105. 45XO: Turner syndrome	Congenital non-Down syndrome
140121. 22q11 microdeletion (CATCH 22)	Congenital non-Down syndrome
140200. Syndrome/association with cardiac involvement	Congenital non-Down syndrome
140206. DiGeorge sequence	Congenital non-Down syndrome
140217. Marfan syndrome	Congenital non-Down syndrome
140219. Noonan syndrome	Congenital non-Down syndrome
140228. Tuberous sclerosis	Congenital non-Down syndrome
140230. Williams syndrome (infantile hypercalcaemia)	Congenital non-Down syndrome
140232. Fetal rubella syndrome	Congenital non-Down syndrome
140266. Alagille syndrome: arteriohepatic dysplasia	Congenital non-Down syndrome
140300. Non-cardiac abnormality associated with heart disease	Congenital non-Down syndrome
140304. Non-cardiothoracic/vascular abnormality (describe)	Congenital non-Down syndrome
140306. Cystic fibrosis	Congenital non-Down syndrome
140307. Diaphragmatic hernia	Congenital non-Down syndrome

**TABLE 12** Mapping from the CCAD comorbidity codes to comorbidity groupings (*continued*)

CCAD official comorbidity code	Comorbidity group
140308. Tracheo-oesophageal fistula	Congenital non-Down syndrome
140310. Omphalocele	Congenital non-Down syndrome
140311. Duodenal stenosis/atresia	Congenital non-Down syndrome
140323. Renal abnormality	Congenital non-Down syndrome
140329. Thoracic/mediastinal abnormality	Congenital non-Down syndrome
140333. Microcephaly	Congenital non-Down syndrome
140404. Pectus carinatum	Congenital non-Down syndrome
140405. Pectus excavatum	Congenital non-Down syndrome
140409. Kyphoscoliosis	Congenital non-Down syndrome
140412. Cleft lip/palate	Congenital non-Down syndrome
140414. Anterior chest wall (pectus) deformity	Congenital non-Down syndrome
140601. Multiple congenital malformations	Congenital non-Down syndrome
140102. Trisomy 21: Down syndrome	Down syndrome
000000. No preprocedural risk factors	Normal
102202. Premature birth	Premature
102019. Preprocedural risk factor	Unknown

## Protocol for diagnostic grouping at the level of a record

### *Scheme 1: mapping International Paediatric Congenital and Cardiac codes to primary diagnosis*

#### At the level of a diagnosis code (that is, an individual diagnosis field)

Each diagnosis code is assigned to one of the 27 categories in *Table 13*, which are in order of hierarchy from highest to lowest in terms of clinical severity.

#### At the level of a record

Once each diagnostic code (diagnosis fields 1–6) has been assigned to one of these groups, a hierarchy is used to allocate the record a single group; the record is assigned to the highest group on this list to which one of its codes corresponds.

### *Scheme 2: mapping Central Cardiac Audit Database codes/records to diagnostic groups containing univentricular heart/biventricular heart/arch obstruction information*

#### At the level of a diagnosis code (that is, an individual diagnosis field)

First, each diagnosis code is assigned to one of the categories in *Table 14*.

For those diagnosis codes that are in the congenital category, we seek four further pieces of information:

1. a diagnosis code is assigned a UVH flag if that code, in isolation, definitely indicates that a patient has a UVH
2. a diagnosis code is assigned a BVH flag if that code would definitely indicate, in the absence of any other significant codes, that a patient has a BVH

**TABLE 13** Hierarchy of the 27 groups in scheme 1<sup>a</sup>

1	HLHS
2	Functionally UVH
3	Common arterial trunk (truncus arteriosus)
4	TGA + VSD/double-outlet right ventricle (DORV): TGA type
5	Interrupted aortic arch
6	TGA (concordant atrioventricular and discordant ventriculoarterial connections) and IVS
7	Pulmonary atresia (including pulmonary atresia + IVS)
8	Pulmonary atresia + VSD (including Fallot type)
9	AVSD
10	Fallot/DORV: Fallot type
11	Aortic valve stenosis (isolated)
12	Tricuspid valve abnormality (including Ebstein's)
13	Mitral valve abnormality (including supralvar, subvalvar)
14	Totally anomalous pulmonary venous connection (TAPVC)
15	Aortic arch obstruction +/- VSD/ASD
16	Pulmonary stenosis
17	Subaortic stenosis (isolated)
18	Aortic regurgitation
19	VSD
20	Interatrial communication ('ASD')
21	PDA
22	Miscellaneous congenital
23	Acquired
24	Procedure
25	Comorbidity
26	Normal
27	Ungrouped diagnostic terms

ASD, atrial septal defect; IVS, intact ventricular septum.  
 a Categories shaded grey are non-congenital diagnoses.

**TABLE 14** Diagnosis categories in mapping scheme 2

1	Congenital
2	Acquired
3	Procedure
4	Comorbidity
5	Normal
6	Ungrouped diagnosis



3. a diagnosis code is called 'ambiguous congenital' if, in isolation, that code provides no information on whether that patient has either a UVH or a BVH
4. a diagnosis code is assigned an 'arch obstruction' flag if that code, in isolation, definitely indicates the presence of an arch obstruction or systemic arterial obstruction.

### At the level of a record

Once each diagnostic code (diagnosis fields 1–6) has been assigned a category and the four additional pieces of information, these are combined at record level in the following manner:

1. if at least one diagnosis code in a record has a congenital diagnostic code then the entire record is assigned a 'congenital' flag
2. if at least one diagnosis code in a record has an acquired diagnostic code then the entire record is assigned an 'acquired' flag
3. if at least one diagnosis code in a record has a comorbidity diagnostic code then the entire record is assigned a 'comorbid diagnosis' flag
4. if none of the diagnosis codes in a record is either acquired or congenital then the entire record is assigned a 'not a cardiac diagnosis' flag

(Note that the steps 5–8 will be relevant only to records that have a congenital diagnosis flag)

5. if any of the diagnosis codes in the record give positive information that the patient has a UVH (i.e. record has a UVH flag) then the entire record is assigned a UVH flag and the BVH flag is set to zero
6. if none of the diagnosis codes in the record gives positive information that the patient has a UVH and any of them *do* give positive information that the patient has a BVH (i.e. record has a BVH flag) then the record is assigned a BVH flag and the UVH flag is set to zero
7. if any of the diagnosis codes in the record are 'ambiguous congenital' then the record is also flagged as 'ambiguous congenital', regardless of which other flags have already been set
8. if any of the diagnosis codes in the record give positive information that the patient has an arch obstruction or systemic arterial obstruction (i.e. record has an arch obstruction flag) then the entire record is assigned an arch obstruction flag.

### Working back from specific procedures

There are some specific procedures which indicate that a patient must have a certain category of congenital problem.

#### *Univentricular heart*

For **all** records we look at the specific procedure for that record. The following three specific procedures are unambiguously associated with UVH patients: Norwood procedure (stage 1), Fontan procedure and bidirectional cavopulmonary shunt (Glenn). For any record that has any of these specific procedures, we flag the record as UVH and set the BVH flag to zero. We also flag that this assignment is based on specific procedure.

#### *Arch obstruction*

Similarly, if a record has not been positively identified as a patient with an arch obstruction then we look at the specific procedure for that record. The following seven specific procedures are unambiguously associated with patients who have an arch obstruction: Norwood procedure (stage 1), isolated coarctation repair, interrupted aortic arch repair, truncus and interruption repair, aortic valvotomy, subvalvar aortic stenosis repair and supra-aortic stenosis repair. For any record that has any of these specific procedures, we flag the record as an arch obstruction and also flag that this assignment is based on specific procedure.



## Appendix 3 Comparison of diagnostic groups and specific procedure

There was good face validity between primary diagnoses and the procedures associated with them at the record level in the development set, as shown in *Figure 38*, which is a tabular match between scheme 1 diagnoses and the 'specific CCAD procedures'. For example, the most commonly performed specific procedures in patients with a diagnosis of VSD were (grey and dark-grey cells) VSD repair, isolated pulmonary artery banding and 'not a specific procedure'; for HLHS the most common were Norwood 1, isolated pulmonary artery banding, arterial shunt, bidirectional cavopulmonary shunt, Fontan and 'not a specific procedure'. It is important to recall that these data relate to the first instance of a patient in the data set.

Each cell in this tabulation represents a possible combination of specific procedure and primary diagnosis category. Many combinations do not occur in our data set (white cells). Light-grey cells indicate that a small proportion (< 2%) of records with a given primary diagnosis category had that specific procedure. Grey (2–10%) and dark-grey (> 10%) cells indicate more commonly occurring combinations. Such tabulations were used to identify incongruous combinations of diagnosis category and procedure during the iterative development of the schemes. This final tabulation indicates where diagnostic information adds value to procedural information in characterising patients.

Procedure	No specific procedure type	Anomalous coronary artery repair	Aortic root replacement (not Ross)	Aortic valve replacement-Ross	Aortic valvotomy	Aortopulmonary window repair	Arterial shunt	Arterial switch (for isolated transposition)	Arterial switch+VSD closure	ASD repair	AVSD (complete) repair	AVSD (partial) repair	AVSD and tetralogy repair	Aortic valve replacement - non-Ross	Bidirectional cavopulmonary shunt	Cor triatriatum repair	Fontan procedure	Interrupted aortic arch repair	Isolated coarctation repair	Isolated pulmonary artery band	Mitral valve replacement	Multiple VSD Closure	Norwood procedure (stage 1)	PDA ligation (surgical)	Pulmonary atresia VSD repair	Pulmonary valve replacement	Rastelli procedure	Repair of total anomalous pulmonary venous drainage	Senning or Mustard procedure	Subvalvar aortic stenosis repair	Subvalvar aortic stenosis repair	Tetralogy repair	Tetralogy with absent pulmonary valve repair	Tricuspid valve replacement	Truncus and interruption repair	Truncus arteriosus repair	VSD repair							
<b>Scheme 1</b>																																												
Hypoplastic left heart syndrome																																												
Functionally UVH																																												
Common arterial trunk (truncus arteriosus)																																												
TGA+ VSD/DORV - TGA type																																												
Interrupted aortic arch																																												
TGA (concordant AV and discordant VA connections) and IVS																																												
Pulmonary atresia (including PA+IVS)																																												
Pulmonary atresia+VSD (including Fallot type)																																												
AVSD																																												
Aortic valve stenosis (isolated)																																												
Tricuspid valve abnormality (including Ebstein's)																																												
Mitral valve abnormality (including supralvalvar, subvalvar)																																												
TAPVC																																												
Aortic arch obstruction ± VSD/ASD																																												
Pulmonary stenosis																																												
Subaortic stenosis (isolated)																																												
Aortic regurgitation																																												
VSD																																												
Interatrial communication ('ASD')																																												
PDA																																												
Miscellaneous congenital																																												
Acquired																																												
Non-cardiac or uncoded diagnosis																																												
<b>Scheme 2</b>																																												
UVH+																																												
UVH-																																												
BVH+																																												
BVH-																																												
Acquired only																																												

**FIGURE 38** Comparison of specific procedure and diagnosis (see explanation in the text). ASD, atrial septal defect; DORV, double-outlet right ventricle; IVS, intact ventricular septum; PA, pulmonary atresia; TAPVC, totally anomalous pulmonary venous connection.

## Appendix 4 Final logistic regression risk model

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

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

In this appendix parameters  $i=1-39$  are tabulated 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.

We calculated the Copas shrinkage for this model to be an adequate 0.87 using the method detailed in Copas<sup>55</sup> 1997, *Stats Medicine*.

$i$	$X_i=1$ 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=atrial septal defect repair	-1.234
7	Specific procedure=AVSD (complete) repair	0.135
8	Specific procedure=AVSD (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

continued

<i>i</i>	$X_i=1$ if condition satisfied ( $X_i=0$ otherwise)	$B_i$
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 UVH	-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

The most up-to-date version of the PRAiS model specification will always be available from the UCL Clinical Operational Research Unit website: [www.ucl.ac.uk/operational-research/AnalysisTools/PRAiS](http://www.ucl.ac.uk/operational-research/AnalysisTools/PRAiS). Please check for updates before implementing.

## Appendix 5 Protocol

**HSR Protocol - project ref:** 09/2001/13

**Version:** 1

**Date:** 30/07/2010

**The Application of a Mortality Risk Model to Adjust for Case Mix in Paediatric Cardiac Surgery for the United Kingdom using the Central Cardiac Audit Database**

**Chief investigator** Mr Victor Tsang

**Sponsor** Great Ormond Street Hospital for Children NHS Trust

**Funder** HSR Programme

**NIHR Portfolio number**

**ISRCTN registration (if applicable)**

### The Application of a Mortality Risk Model to Adjust for Case Mix in Paediatric Cardiac Surgery for the United Kingdom using the Central Cardiac Audit Database

#### 1. Aims/Objectives:

##### AIMS:

To establish whether a risk model can be developed that is fit for the purpose of adjusting for case mix severity during routine monitoring of outcomes for paediatric cardiac surgery in the United Kingdom (UK).

##### OBJECTIVES:

Our objectives are as follows.

1. To test an existing risk model based on the RACHS-1 score and patient age, derived from outcomes at one centre (GOSH), to the CCAD data from all centres across the UK.
2. To understand the contribution that diagnostic information can make to risk estimation and monitoring of outcomes, establish whether information concerning co-morbidities can contribute to improved methods of risk estimation and, if indicated and possible, revise the existing risk model such that it is suitable for use at other centres and by CCAD.
3. To examine the implications of reporting mortality outcomes by diagnosis as well as by procedure category.
4. To disseminate our findings and any risk models and monitoring tools developed to UK centres and to the CCAD so that it can consider how best to share the information with stake holders including via its "public portal" web pages.

##### STATEMENT OF RESEARCH QUESTIONS:

It has been recognised for some time that it is important and valuable to monitor outcomes in cardiac surgery and that to do so fairly and effectively, one needs to risk-stratify the case load of each unit. Analytical methods for doing this are well advanced for adult cardiac surgery and the use of graphical techniques to

display risk-adjusted outcome charts is now a commonly used part of the quality assurance process. Currently no process for routine risk adjustment for paediatric cardiac surgery exists. Such methods have been suggested for paediatric cardiac surgery but as yet these have not been subjected to rigorous testing using large multi-site data sets and this is necessary before the methods are introduced into routine use. Support is sought for a multi disciplinary clinical group with expertise on outcomes in paediatric cardiac practice, to work with the team of mathematicians behind many of the techniques used in outcome monitoring for adult cardiac surgery. Working with the Central Cardiac Audit Database, we will use the data to investigate the robustness of risk scoring across different centres with a view to establishing the feasibility of routine outcome monitoring methods in paediatric cardiac surgery. This will include exploring the addition of diagnosis and co-morbidity information to existing risk models and the implications of reporting outcomes by both diagnosis and procedure. The proposed study is not a trial or an experiment based on formal hypothesis testing. Rather, the hypothesis that motivates the study is that the use of case-mix adjustment in the routine monitoring of outcomes for congenital heart surgery would improve quality assurance processes and the information available to surgeons and carers. The work will establish whether a risk-model can be found that is fit for this purpose.

## 2. Background:

### Progress in risk stratification for paediatric congenital heart surgery

It has long been accepted [1] that risk stratification of adult cardiac surgery patients is an essential part of the audit process which reduces the prospect of unfair assessment of outcomes attributed to a surgeon or team whose mortality rate is relatively high simply because it reflects patients that were inherently higher risk. In addition, various methods have been developed to help in the analysis and interpretation of risk stratified data [2–6].

Achieving something similar for paediatric congenital heart surgery is clearly desirable, but is challenging, due to the great diversity of the patient population in terms of the diagnoses, indications for surgery, the operation performed, age at operation and other factors [7]. At present, mandatory data submissions are requested every three months to CCAD from hospitals performing cardiac surgery in the UK, including details about the patient and the operation performed. CCAD later independently checks the patients' survival status with the Office of National Statistics and therefore patient identifiers including the NHS number are provided to CCAD, which holds exemption under section 60 of the Health and Social Care Act 2001.

Even before much had been published on risk scoring systems for congenital surgery, simple monitoring methods had been discussed [5]; however since then there have been advances in this field. With respect to the operation performed, there exists an internationally accepted classification scheme called RACHS-1 [8], by which 79 different types of operation are grouped into 6 categories ranked in order of increasing risk, as perceived by clinicians. The RACHS-1 scheme appears to be useful as a basis for forecasting risk [7], has been validated in a range of contexts and as of November 2009 has been cited in 214 scholarly articles. There is also a scoring system known as ARISTOTLE [9] for grading the complexity of operations, although it has been shown that the RACHS-1 scoring system is better suited for stratifying risk [10]. Recently two large databases in North America and Europe were used to develop an empirically based tool for monitoring mortality after congenital heart surgery, which appears to perform well [11]. However, this tool was developed using voluntary data submissions and is based on the survival status at hospital discharge, which may differ significantly to the outcome at 30 days.

Importantly, the literature and our own work to date focus on outcomes and risks according to the procedure(s) that patients have undergone, augmented by patient-specific information such as age and weight. Currently, no use is made of information concerning the nature of the heart defect nor of any co-morbidities that the patient has. Whilst in many areas of surgery there is a one-to-one mapping between diagnosis and surgical intervention, this is not the case in paediatric congenital heart surgery. Some procedures are performed for a number of diagnoses, which undermines the extent to which the procedures



performed by a surgeon or within a unit accurately reflects case-mix. Similarly, the same heart defect may be managed using different surgical interventions, with the choice of intervention reflecting other aspects of the patient's condition (age, weight, comorbidities, severity of symptoms) but also with scope for there to be differences in surgical strategy between units based on local experience and, potentially, on a different balance being struck between long term objectives and short-term risks. An understanding of this complex issue is essential to determining whether and how diagnostic information may be usefully combined with information concerning the procedures performed in methods for risk-stratification.

### Appropriateness of the project team

Work led by the first applicant has resulted in the development of a predictive model of risk for paediatric cardiac surgery based on the RACHS-1 classification and age at operation [12]. This risk model has been developed using only data concerning patients that had surgery at Great Ormond Street Hospital (GOSH). One of the main motivations for this research proposal is to establish the extent to which this risk model is appropriate for use with data concerning patients of other hospitals; it is conceivable that, due to differences in the nature of the patients seen and the surgical strategies employed, the relationships between RACHS-1 classification, age and risk may differ at other centres.

The second applicant has trained in health services research, having undertaken the MSc in public health with health services research focus at the London School of Hygiene and Tropical Medicine in 2007. She has since taken on the role of clinical outcomes lead for cardiothoracic services at GOSH where she has worked with the CCAD to improve the data quality of cardiac surgery procedure records at GOSH and has functioned as an external assessor for CCAD at other institutions.

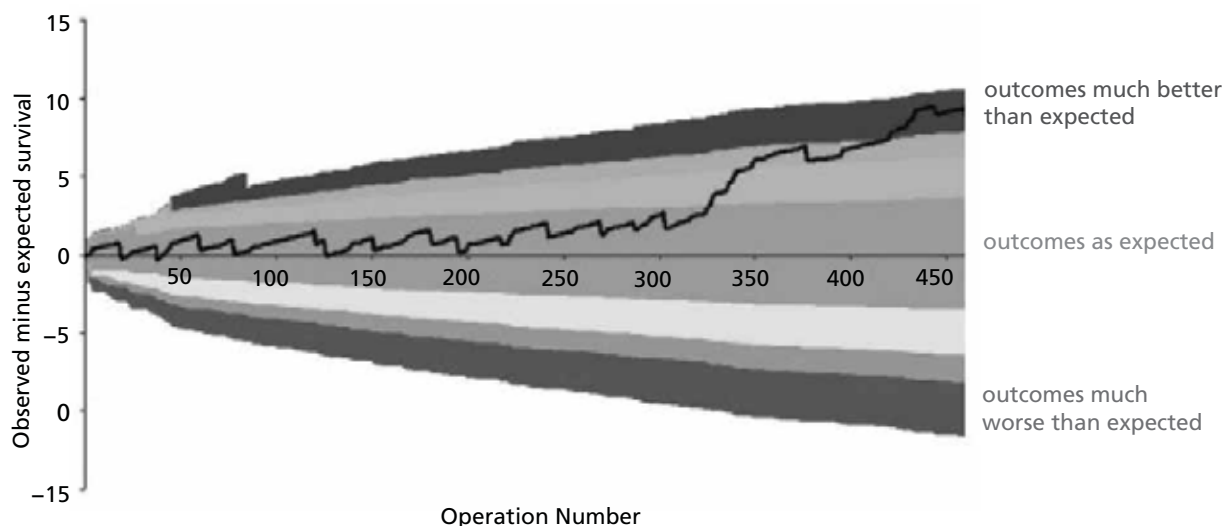
The third applicant is based with CCAD and has been the original project manager for the cardiac audits, since CCAD was founded in 1996. He has both technical training with a PHD and a clinical background in cardiac electrophysiology, a combination which facilitates the liaison between the clinical leads of the national audits and the technical staff in CCAD who build the applications. His responsibilities to date have included the analysis of the data and generation of the tables and funnel plots which are shown on the Congenital Public Portal ([www.ccad.org.uk/congenital](http://www.ccad.org.uk/congenital)).

The fourth, fifth and sixth applicants work at the Clinical Operational Research Unit (CORU), which is an academic group that applies mathematical and statistical modelling methods to a wide range of clinical problems. The unit has particular research interests related to cardiology and cardiac surgery [see, for example 2, 4, 5, 6, 11]. In collaboration with Professor Tom Treasure (now an honorary member of the unit), CORU helped to develop a simple graphical method (known as VLAD) that can be used to chart mortality outcomes in adult cardiac surgery for a unit or individual surgeon, taking due account of the risk stratification of caseload [2]. This method is used extensively and has become a standard analysis tool in the audit of adult cardiac surgery. CORU has also been active in methodological research associated with risk scores and outcomes assessment, being the first to use so called 'funnel plots' for assessing outcomes in cardiac surgery [14], having developed novel methods for assessing risk scoring systems [15] and having collaborated with the development of a risk scoring system for adult thoracic surgery [16]. There have been previous successful collaborations between CORU and Great Ormond Street Hospital including work on outcomes [5,11,14], safety [13] and hospital operation [17,18].

The project team will be assisted by a research fellow with skills in data analysis, mathematical modelling and inter-disciplinary collaborative work with clinical teams. The project goes well beyond formulaic data analysis. It is necessary for those involved in data analyses to be familiar with the operation of a paediatric cardiac surgery centre, to gain an appreciation of the surgical nomenclature, the context in which data are collected and also to have an appreciation of the ways in which outcomes monitoring can be used to promote good practice. Such a "hands-on" approach to analytical work within clinical research is a hallmark of CORU's methodology and will be facilitated by the make up of the project team.

### Monitoring outcomes in congenital heart surgery

As already discussed, charts of outcomes that take preoperative risk factors into account have been developed for use in adult cardiac surgery to display a surgeon's overall outcomes and to highlight any disturbing trends [2,6]. These charts are called Variable Life Adjusted Displays, or VLAD charts, which is their commonly used acronym. Members of the project team have been instrumental in establishing the technical feasibility of constructing VLAD charts relating to paediatric congenital heart surgery [12]. One such VLAD chart is shown in *Figure 1*. Here, successive operations are plotted on the horizontal axis and the jagged graph shown in bold gives a running tally of how much better (or worse) outcomes have been compared to what would be expected using a risk model. Systematic descent of this curve indicates possible cause for concern, the coloured regions indicating the chance that departure from the horizontal could be a chance coincidence. In this example, the VLAD curve rises, which indicates better than expected outcomes.



**FIGURE 1** An example of a Variable Life Adjusted Display (VLAD) Chart.

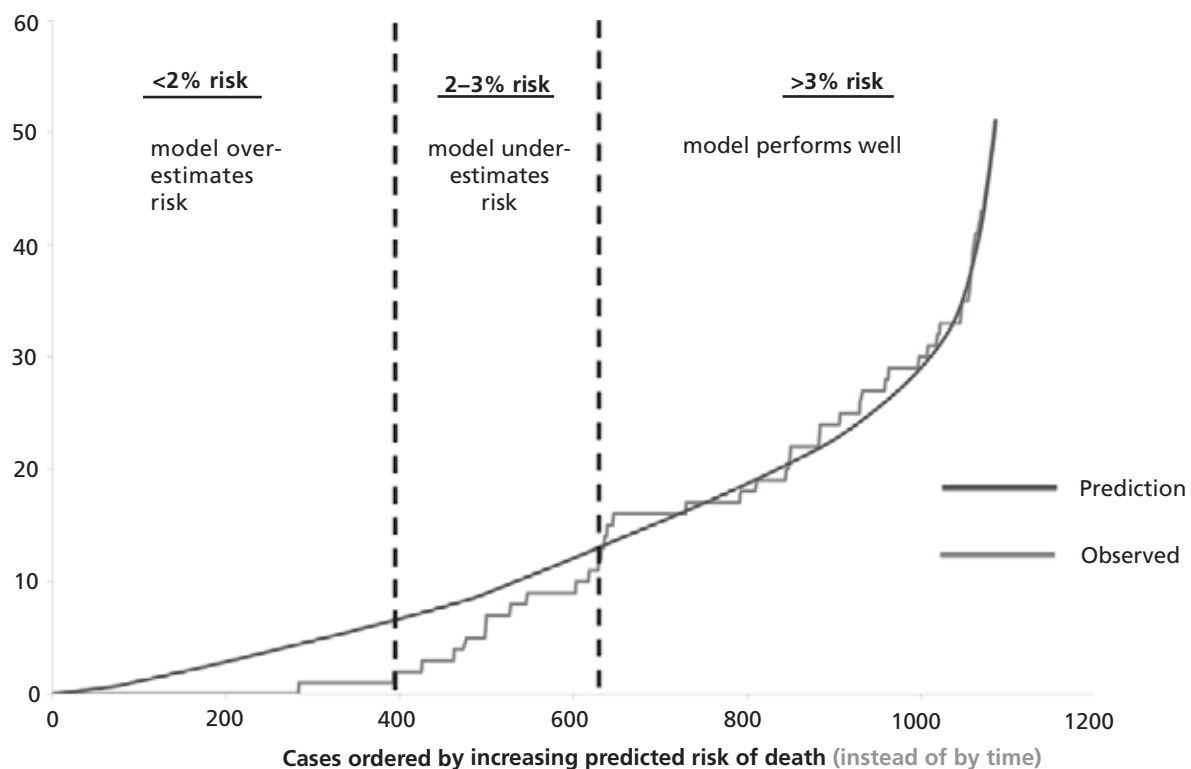
Before such methods are introduced into common use for paediatric cardiac surgery, there is a need to test both the underlying risk model and the monitoring methods to ensure that they are not misleading. The project team have begun this process by assessing the accuracy of the risk model across the range of risk, using data from a single unit (GOSH). As shown in *Figure 2*, we have established that, among the Great Ormond Street data used in the construction of the model, the risk model performs well for predicted risks over 3%, under-estimates risk for predicted risks in the range 2%–3% and over-estimates risk for patients with a predicted risk less than 2%. These features of the risk model do not significantly undermine its use in monitoring outcomes at GOSH [19] but the question arises as to whether such features could bias comparisons between centres with markedly different case mix.

Current NHS policy places increased weight on patient outcomes and quality improvement. Clearly one of the remits of a national audit database is to monitor outcomes and quality. The complexity in both diagnosis and procedure for congenital heart surgery has meant that this is currently not feasible. The main objective of this research project is to explore the feasibility of such routine monitoring.

### 3. Need:

In the wake of the Bristol Royal Infirmary Inquiry where one centre experienced a number of excess deaths in children after cardiac surgery, there has been an emerging culture of audit and quality improvement in the NHS. The CCAD public web portal for paediatric heart surgery data was launched in 2007 and up to the end of 2008, 68,000 procedures have been submitted from throughout the UK. During this period, outcomes of paediatric cardiac surgery have become subjects of great interest to the clinical community, patient families and the public. Despite this, relatively little is known about risk-stratified mortality rates in this context, because the case mix for paediatric cardiac surgery is very complex. Currently

## Cumulative deaths



**FIGURE 2** Comparing cumulative tallies of deaths compared to what would be expected given the risk model developed at Great Ormond Street Hospital.

no process for routine application of risk adjustment exists that addresses this complexity. Therefore, despite effective mandatory national audit data collection, clinical teams and others do not have tools for monitoring risk-adjusted outcomes in a fair and robust way.

The CCAD organisation is strongly supportive of this research: the most recent CCAD national contributors' meeting, which includes patient family representation, and the CCAD Newsletter noted the deficiencies in access to data analysis as significant impediments to the audit program. The current approach taken by CCAD in terms of assessing and reporting patient outcomes is to evaluate 36 specific paediatric cardiac surgery procedure categories, summarising the number of deaths in each one for individual centres and comparing these between centres. There are two major problems with this system of evaluation: the number of procedures is large and the number of individual patients in each category, even across the whole country is relatively small and therefore the system is very unlikely to detect true differences in outcome in a timely fashion. The second is that the 36 specific procedure categories do not capture all of the operations performed (RACHS-1 identifies 79) and it excludes many operations that result in mortality, rendering those procedures relatively inaccessible to the monitoring system. For these reasons we currently lack the necessary insights into how to monitor and assess outcomes to ensure that optimal standards are maintained.

Outcome monitoring with appropriate risk adjustment is well advanced in adult cardiac surgery, which has been of clear benefit to the surgical profession in relation to its quest for improved quality assurance. There is evidence that since outcome monitoring in adult cardiac surgery became mandatory and routine, outcomes have improved and there has been no consequent negative effect in terms of centres turning away high risk cases, as was originally feared [20]. There were consequent benefits to patients and their families in terms of the quality of care and the improved information they received. It is hoped that new knowledge generated by the results of this proposed research will help to promote equivalent quality assurance and transfer of information in the paediatric cardiac surgical community in a sustainable way.

This proposed project, which seeks to explore the feasibility of routine risk adjustment to paediatric cardiac surgery in the UK, is possible since a high quality national database now exists. The only way to effectively perform a project of this kind is to use a large national dataset such as the CCAD database in order to remove the biases associated with optional reporting of procedures such as occurs in other large databases [11] and avoid the case mix bias that may arise when using data from a single centre. Since patient mortality outcomes are independently checked by CCAD with the Office of National Statistics, the chance of underreporting of poor outcomes, as has also been reported in a recent study from North American and European Databases [11], will be avoided. As stated, the project team includes a senior figure within the CCAD organisation and has put in ground work in this area over the previous few years as laid out in BACKGROUND. The project is in line with the mission of the HSR program since it seeks to apply a risk adjustment model to an existing national database, which is one of the stated areas of its remit. Furthermore, the proposed research complements current wider NHS policy in terms of quality improvement and the generation of quality accounts, since adjustment for case mix in paediatric cardiac surgery would increase the likelihood of such accounts being useful and interpretable.

#### 4. Methods:

##### a. Setting

###### *Preparatory work*

The national dataset of paediatric cardiac surgery procedures (68,000 records) is currently being prepared for use in this programme of research by the CCAD and will be released to the analysts shortly. Over the period January to May 2010 inclusive, we will develop an algorithm to automate the classification of surgical procedures according to the RACHS-1 scheme. Additionally we will liaise with the CCAD to ensure that the analysts, who have already spent time familiarising themselves with the context of paediatric heart surgery, are fully conversant with the methods employed by CCAD in collating, "cleaning" and coding the data it receives from surgical units. We are not seeking funding for this preparatory work.

##### b. Data analysis

###### *Data analysis and model development*

Given the features of the performance of the current risk model that are evident using data from Great Ormond Street Hospital, it is inevitable that the development of a more refined model of peri-operative mortality risk will be warranted. A carefully designed analysis strategy will be used, subdividing data gathered into two sets – the development set and the test set. These sets will be compiled using a randomisation process to select of 70% of all data available at the outset of the project, stratified by centre and calendar year, to form the development set. The test data set (the remaining 30% of the available data) will not be used until the latter phases of the study and will play no part in risk model development. It is necessary to observe such a quarantine process to avoid the bias that inevitably results if the same data that are used to develop a risk model are used to test it, which almost always leads to overoptimistic test findings. We anticipate that some of the work on this section of the project, developing the procedure based model may be complete by mid way through 2010.

**Initial exploratory analyses** The development set will first be used as the basis for a number of exploratory analyses in order to learn more about pre-operative factors reported to CCAD that influence outcome and other issues associated with the development and use of risk scores. Since the development set is quarantined from the test set, however many analyses of this sort are performed, no accidental bias can be introduced to the final evaluation of the resultant risk model. Such preliminary analyses will include:

1. Identification of inter- and intra-centre differences in practice;
2. Identification of any time varying patterns of operative complexity (as assessed by the RACHS-1 score), predicted risk (assessed using the existing Great Ormond Street risk model) or outcome;

3. Examination of the sensitivity of risk scoring to the classification used within the RACHS-1 method;
4. Examination of outcomes by diagnosis, and any subsequent implications for use of risk models based only on procedure.

A particularly important topic for investigation concerns the treatment of data in cases where a patient has several procedures during the course of a single operation, several operations during a single episode of hospital care or several operations at different points in time in the management of single heart defect. This is not uncommon with patients who have congenital heart abnormalities. The usual manner in which such cases are used in risk scoring exercises is to assign the outcome to the procedure that has the highest complexity score however it is far from clear whether this is sound. We intend to examine this by examining outcomes for patients assigned to a given complexity score to see whether those having multiple procedures have a higher mortality than those who don't. This issue may partly be addressed by the exploration of risk adjustment based on diagnosis.

These issues will be addressed both by data analysis and by mathematical modelling, which has proved useful in this context [21].

**Development of a more refined risk model** The next phase of analysis will be concerned with using the development data set in order to derive a risk model more refined than that presently available, better reflecting current practice across the UK. Development of a new risk model will have two phases, the first concerning the choice of risk factors to include from those routinely reported to CCAD, the second being one of calibration.

**Assessing the performance of the new risk model** The final stage of the analysis will be to use the test data set (previously quarantined from the analysts) to assess the performance of the risk model. Formal statistical testing will be carried out using the Hosmer Lemeshow test [22] and related statistical methods. In addition, analysis using the MADCAP method [14] will be performed to provide a visual guide to the performance of the risk model over the entire range of risk, [c.f. *Figure 2*].

**Assessing the robustness of the routine monitoring of risk-adjusted outcomes** Having assessed the performance of the risk model, we will go on to establish whether its use in the routine monitoring of outcomes is robust. To this end a further set of data will be obtained from the CCAD covering the period since the start of the research project. These data will be used as test data for examining the performance of the monitoring tools developed to mimic the prospective monitoring of risk adjusted data at the level of individual units, enabling the analysts to establish the extent to which any features of the risk model may undermine the validity of its routine use. Importantly, this activity will also give some indication as to how often the risk model should be updated to reflect changes in the underlying risk of peri-operative mortality faced by paediatric congenital heart surgery patients in the UK.

Based on this final analysis, we will establish caveats to the interpretation of risk adjusted outcome data that will be disseminated alongside tools for prospective monitoring of outcomes in paediatric congenital heart surgery.

**Power Calculations** Access has been granted by the CCAD to available data concerning all congenital heart operations performed in UK sites during the period 2002-2008 inclusive and on an ongoing basis, the size of sample available to the research team (approximately 68,000 cases in the first instance) is ample for the purposes outlined.

### **5. Contribution to collective research effort and research utilisation:**

It is hoped that the main product of this research will be a risk model that is suitable for routine monitoring of outcomes, although it is possible that analysis will show that the complexity of congenital heart surgery is not adaptable to such routine monitoring. Although certain procedure based mortality rates are

currently available, and are of interest for stake holders who may require information on a particular procedure, the overall performance of a cardiac programme cannot be monitored on the basis of a single procedure. We plan to offer graphical displays of results over time, which could be used to identify increases or decreases in the performance of the programme, thus rendering these more accessible to quality control. The output of our research would be better powered to detect problems arising in a programme at an earlier stage. We plan to work with key stake holders to ensure that our work is presented in a fashion which may be as easily interpreted as possible, including displaying the information on the CCAD public portal for all to see.

### 6. Plan of Investigation:

A timeline setting out the plan for achieving the aim of this project is shown below. Important milestones are shown by bold vertical lines indicating: the end of the model development period; the end of work establishing the feasibility of routine risk adjustment and the end of funded dissemination activity.

Activity	Jun 2010	Jul 2010	Aug 2010	Sep 2010	Oct 2010	Nov 2010	Dec 2010	Jan 2011	Feb 2011	Mar 2011	Apr 2011	May 2011
Test of previously developed RACHS and age risk model on national data												
Explore the interaction between diagnosis, procedure and outcome in evaluating risk												
Development of a more refined risk model possibly incorporating diagnosis & comorbidities												
Assessing performance of new risk model												
Explore the feasibility of using the risk model in routine monitoring												
Prepare papers detailing research findings for submission to peer-reviewed journals												

### APPROVAL BY ETHICS COMMITTEES

The CCAD database is exempt under section 60 of the Health and Social Care Act 2001, and as such is permitted to hold information relating to patients undergoing cardiac interventions in the UK. A data sharing agreement has been formed between the investigators based at the Clinical Operational Research Unit (CORU), University College London and the Information Centre for Health and Social Care, which sets out the conditions under which data may be transferred to CORU for the purposes of this research.

### 7. Project Management:

The project will be jointly managed by Dr Brown (at Great Ormond Street Hospital) and Professor Utley (at CORU) under the leadership of Mr Tsang. Formal progress meetings will be held monthly with the physical proximity of UCL and Great Ormond Street Hospital permitting frequent informal discussions concerning clinical and analytical issues as they arise during the course of the project. The analytical team hold honorary contracts at Great Ormond Street, which will enable them to gain essential insights concerning the context in which any risk-adjusted monitoring of outcomes will take place. The bulk of the analytical work will be conducted at CORU by a research fellow under the line-management of Dr Pagel.

### 8. Service users/public involvement:

The output from this study will be subjected to similar procedures to the data that are currently presented on the CCAD portal, which includes review by patient group representatives and other key stakeholders that are represented. Views will be sought on the interpretability and usefulness of the data and final presentation on the public portal will reflect this.

### EXPERTISE AND JUSTIFICATION OF SUPPORT REQUIRED

The project team is a collaboration between clinicians, mathematicians with considerable experience of working with clinical data and the relevant national audit body (the Central Cardiac Audit Database), which combines:

1. Relevant clinical expertise developed working in the field of paediatric cardiac surgery (applicants 1 and 2) that is essential for successful completion of a complex project with direct clinical relevance such as this;
2. Data analysis and data management expertise (applicant 3) accrued by project management of the cardiac audit database for the previous 13 years from its inception;
3. Mathematical modelling and operational research expertise (applicants 4, 5 and 6) which has contributed to a range of projects in this area including the development of graphical monitoring tools used worldwide, the development of methods to evaluate the performance of risk models over the full spectrum of risk and the development of a risk model for use in monitoring outcomes of thoracic surgery.

The project team already have a successful track record of working together with a number of peer reviewed publications and presentations, most recently (Tsang V T, Brown K L, Synnergren M J, deLeval M R, Kang N, Gallivan S, Utley M, 'Monitoring risk-adjusted outcomes in congenital heart surgery: Does the appropriateness of a risk model change with time?', *The Annals of Thoracic Surgery* 87(2):584-587, 2009).

Work led by the first applicant Mr Tsang, who is an experienced cardiac surgeon, has resulted in the development of a predictive model of risk for paediatric cardiac surgery based on the RACHS-1 classification and age at operation. This applicant will supervise a more junior fellow in cardiac surgery, both on a part time basis, to provide clinical input for any procedure coding issues that may arise. The second applicant has trained in health services research, and has several years' experience of data quality in relation to CCAD and will assist with clinical and data quality aspects of the work. Funding is sought for the first and second applicants and a cardiac surgical fellow, on a part time basis for one year to work on the clinical aspects of this project. The availability of diagnostic information and our plans to explore the potential use of this information make the success of this endeavor dependent on the clinical members of the team having sufficient, protected, time to work alongside the analytical team.

No funds are sought for the third applicant who is based with CCAD as the project manager and will function as the main link with this organisation.

The fourth, fifth and sixth applicants work at the Clinical Operational Research Unit (CORU), which is an academic group that applies mathematical and statistical modelling methods to a wide range of clinical problems and have been instrumental in developing monitoring tools for use in adult cardiac surgery. CORU's previous work on adult cardiac surgery audit contributed constructively to this area as may be seen from the publications list of related papers at the end of the full proposal attached in *Appendix B*. Funding is not sought for the team's preparatory work or for the contribution of Professor Gallivan and Dr Pagel; this will be funded from CORU's grant with the UK Department of Health Policy Research Programme. The project team will be assisted by a research fellow with skills in data analysis, mathematical modelling and inter-disciplinary collaborative work with clinical teams. Funding is sought for this research fellow for 7 months and for Professor Utley's contribution.

#### PLANNED OR ACTIVE RESEARCH GRANTS

**2008 – 2011** Cancer survivors and cancer survivorship: quantifying cancer prevalence and modelling its dynamics in England and the UK, Macmillan Cancer Support (£172,000 over 3 years). Co-I Martin Utley.

**2006 – 2010** *The Clinical Operational Research Unit*, funded by the UK Department of Health Policy Research Programme (£2,400,000 over 5 years). PI since March 2007 Martin Utley, PI Jan 2006 – March 2007 Steve Gallivan.

**2006 – 2009** *Developing Evidence Based and Acceptable Stepped Care Systems in Mental Health Care: An Operational Research Project*, funded by the NHS Service Delivery and Organisation R&D Programme (£300,000 over 3 years). Co-I Steve Gallivan, Martin Utley.

#### HISTORY OF PAST OR EXISTING NIHR PROGRAM RESEARCH

None.

## 9. References

1. Parsonnet V, Dean D and Bernstein AD (1989) A Method of Uniform Stratification of Risk for Evaluating the Results of Surgery in Acquired Adult Heart Disease. *Circulation* **79**:13–12.
2. Lovegrove J, Valencia O, Treasure T, Sherlaw-Johnson C, Gallivan S (1997) Monitoring the result of cardiac surgery by variable life adjusted display (VLAD). *Lancet* **350**:1128–1130.
3. Nashef SA, Roques F, Michel P (1999) European System for Cardiac Operative Risk Evaluation (Euroscore). *Eur J Cardiothorac Surg* **16**:9–13.
4. Sherlaw-Johnson C, Lovegrove J, Treasure T, Gallivan S (2000) Likely variations in perioperative mortality associated with cardiac surgery: when does high mortality reflect bad practice? *Heart* **84**:79–82.
5. Gallivan S, Davis K, Stark J (2001) Early identification of divergent performance in congenital cardiac surgery. *Eur. J. Cardiothoracic Surgery* **20**:1214–9.
6. Sherlaw-Johnson C, Gallivan S, Treasure T, Nashef SAM (2004) Computer tools to assist the monitoring of outcomes in surgery. *European Journal of Cardio-Thoracic* **26**:1032–6.
7. Kang N, Cole TJ, Tsang VT, Elliott MJ, de Leval MR (2004) Risk stratification in paediatric open-heart surgery. *Eur. J. Cardiothorac. Surg* **26**:3–11.
8. Jenkins KJ, Gauvreau K, Newburger JW, Spray TL, Moller JH, Iezzoni LI (2002) Consensus-based method for risk adjustment for surgery for congenital heart disease. *J Thorac Cardiovasc Surg* **123**(1):110–8.
9. Lacour-Gayet F, Clarke D, Jacobs J *et al* (2004) The Aristotle score: a complexity-adjusted method to evaluate surgical results. *Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu* **7**:185–91.
10. Kang N, Tsang V, Elliott M, de Leval M, Cole T (2006) Does the Aristotle Score predict outcome in congenital heart surgery? *Eur J Cardiothorac Surg* **29**(5):693–7.
11. O'Brien S, Clarke D, Jacobs J, Jacobs M, Lacour-Gayet F, Pizarro C, Welke K, Maruszuski B, Tobota Z, Miller W, Hamilton L, Peterson E, Mavroudis C, Edwards F (2009) An empirically based tool for monitoring outcome in congenital heart surgery. *J Thorac Cardiovasc Surg Nov* **138**(5):1139–53.
12. Kang N, Tsang V, Gallivan S, Sherlaw-Johnson C, Cole T, Elliott M, de Leval M (2006) 'Quality assurance in congenital heart surgery' *Eur J Cardiothorac Surg* **29**(5):693–7.
13. Catchpole K, Giddings AE, De Leval MR, Peek GJ, Godden PJ, Utley M, Gallivan S, Hirst G, Dale T (2006) Identification of systems failures in unsuccessful paediatric cardiac surgery, *Ergonomics* **49**(5–6):567–88.
14. Stark J, Gallivan S, Lovegrove J, Monro JL, Pollock JCS, Watterson KG (2000) Mortality rates after surgery for congenital heart defects in children and surgeons performance, *The Lancet* **355**:1004–1007.
15. Gallivan S, Utley M, Treasure T, Pagano D (2006) MADCAP: A graphical method for assessing risk scoring systems. *European Journal of Cardio-thoracic Surgery* **29**:431–433.
16. Berrisford R, Brunelli A, Rocco G, Treasure T, Utley M, (2005) The European Thoracic Surgery Database project: modelling the risk of in-hospital death following lung resection', *European Journal of Cardio-thoracic Surgery* **28**:306–311.
17. Pagel C, Peters M, Utley M (2007) Modelling pediatric intensive care capacity and activity by identifying patients with similar length of stay characteristics, *Paediatric Critical Care Medicine* **8**(3 suppl.):A206.



18. Peters M, Pagel C, Gallivan S (2007) A computerised model of bed-demand in a paediatric intensive care unit, *Paediatric Critical Care Medicine* **8**(3 suppl.):A207.
19. Monitoring risk-adjusted outcomes in congenital heart surgery: does the appropriateness of a risk model change with time? (2009) Tsang VT, Brown KL, Synnergren MJ, Kang N, de Leval MR, Gallivan S, Utley M. *Ann Thorac Surg* **87**(2):584–7.
20. Bridgewater B, Grayson A, Brookes N, Grotte G, Fabry B, Au J, Hooper G, Jones M, Keogh B (2007) Has the publication of cardiac surgery outcome data been associated with changes in practice in northwest England: an analysis of 25,730 patients undergoing CABG surgery under 30 surgeons over eight years. *Heart* **93**(6):744–8.
21. Gallivan S (2006) Data-free modelling of errors in databases, In: Lagergren, M. ed. *Proceedings of 30th Meeting of the EURO Working Group on Operational Research Applied to Health Services* Stockholm, pp 217–232, 2006.
22. Hosmer DW, Lemeshow S (1989) *Applied Logistic Regression*. New York: Wiley.

*This protocol refers to independent research commissioned by the National Institute for Health Research (NIHR). Any views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HSR programme or the Department of Health.*





A decorative graphic consisting of numerous thin, parallel green lines that curve from the left side of the page towards the right, creating a sense of movement and depth.

**EME  
HS&DR  
HTA  
PGfAR  
PHR**

Part of the NIHR Journals Library  
[www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

*This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health*

***Published by the NIHR Journals Library***