Overview of statistical evidence presented to the Bristol Royal Infirmary Inquiry concerning the nature and outcomes of paediatric cardiac surgical services at Bristol relative to other specialist centres from 1984 to 1995.

David J Spiegelhalter

Stephen Evans

Paul Aylin

Gordon Murray

September 2000

The research on which this report is based was funded by the Bristol Royal Infirmary Inquiry. All views expressed in the report are the responsibility of the authors alone, and do not necessarily represent the views of the Bristol Royal Infirmary Inquiry.

Executive Summary

1. This overview provides a critical review of statistical evidence presented to the Inquiry regarding the nature and outcomes of paediatric cardiac surgery in Bristol between 1984 and 1995, focussing on the strengths and limitations of the available data sources, and the reliability of conclusions that have been drawn. Key published sources and commentaries have been taken into account. Such a comparative exercise raises a number of difficult issues concerning data quality, the need to aggregate over subgroups, risk-adjustment and so on (*Section 1*).

2. The focus of the analysis is on the performance of surgical services that existed in centres, rather than the performance of individual surgeons. It is therefore not appropriate to adjust for pre-operative risk-factors that may be influenced by preceding care (*Section 1.3.5*). Comparisons of performance were primarily restricted to analyses of 30-day mortality. The main findings were presented in terms of 'excess number of deaths'; namely the number of deaths observed in a given stratum at Bristol minus the number which would have been expected had Bristol been similar to other centres in the country. The excess deaths were summed over strata, and the assessed statistical significance of any excess took account of centre to centre variability (*Section 1.3.6*).

3. Case-mix adjustment was based on age at operation, operative grouping and epoch of operation. Coding of diagnoses and operative procedures in paediatric cardiac surgery is inherently complex and controversial: the operative grouping adopted was devised with substantial clinical input (*Section 2*).

4. All data sources were flawed, and no one source could be considered as representing the 'truth'. Sources used different definitions and variable degrees of quality control: data concerning follow-up of children after discharge from hospital, for example, were erratic. National data were administrative (Hospital Episode Statistics - HES) and professional (UK Cardiac Surgical Register - CSR). HES data have a poor reputation among clinicians, but a linkage exercise with national death registration showed a reasonably accurate correspondence with recorded 30-day in-hospital mortality (*Section 3.1.4*). There was evidence within CSR of highly variable

submissions from some units over the period in question. Although using different definitions and arising from relatively independent sources, HES and CSR data showed reasonable consistency at an aggregated level, although considerably poorer for individual procedure groups (*Section 3.3*). The crucial issue is not whether HES or CSR precisely measure activity and outcome, but the extent to which feasible data inadequacies could explain any observed divergent performance (*Section 3.4)*.

5. None of the five local data sources could be taken as a reliable basis for clinical audit (*Section 4*). In spite of all these problems, there was a surprising degree of agreement between the diverse sources regarding performance in Bristol, especially when restricted to looking at mortality rates following open surgery. This degree of consistency lends credibility to the conclusions drawn from the data (*Section 5*).

6. When compared with performance elsewhere, the main finding was a substantial and statistically significant number of excess deaths at Bristol (*Section 6.2*). Adjusting for operative case-mix did not influence this finding. Particular emphasis was placed on the analysis of data from 1991 to 1995, since data were available for that period from both of the national data sources. Depending on the precise approach to the analysis, the number of excess deaths for open surgery during this period was estimated to be of the order of 30 to 35. The excess mortality corresponded roughly to the mortality rate at Bristol being double that observed elsewhere in England for children aged under one year and even greater for children under 30 days. There was a trend observed outside Bristol for overall mortality rates to fall substantially over the Inquiry period, and this trend was not observed in the Bristol data. Further analysis showed that the excess was not restricted solely to switch and atrial-ventricular septal defect (AVSD) operations, and that missing data on outcomes in HES had minimal influence (*Section 6.4.1*). Evidence for excess mortality was robust to sensitivity analysis to a number of potential data inadequacies (*Section 6.4.3*).

7. Data sources were not of sufficient quality to make any firm conclusion concerning morbidity outcomes (*Section 6.3*).

8. Over the period 1991-1995, both HES and CSR data suggest performance in England (excluding Bristol) was roughly equivalent to published international sources (*Section 7*).

9. There is evidence of an association between lower volume of surgery and increased mortality in open operations on under 1s over the period 1991-1995, even when ignoring the data from Bristol (*Section 8.1*). However, this association only explains a small proportion of the excess observed in Bristol. Other factors regarding comorbidity and status at admission are not substantially related to the observed mortality pattern in Bristol (*Section 8.2*).

10. Between the years 1990 and 1994, there is a clear pattern of a concentration of operations just prior to the first birthday, particularly for AVSDs, and the operative mortality rate at this age is higher than that observed elsewhere. This pattern does not feature in any other centre, and the relevant operations in Bristol appear to be delayed rather than brought forward. This finding is associated with around 25% of the observed excess mortality in Bristol (*Section 8.3*).

11. The Clinical Case Note Review suggested that around 30% of children received less than adequate care, and that in just over 5% different management would reasonably be expected to have made a difference in outcome. Many aspects of the process of care were criticised, with no particular highlight on surgical performance. However, similar measures for other centres are not available, and so we cannot know whether similar criticisms could be made of procedures carried out elsewhere (*Section 8.4*).

12. In spite of the many flaws in the data, we do not believe that apparent divergent performance of this magnitude and consistency can be explained fully by statistical variability or systematic bias in data recording. Rather we conclude that there is strong evidence of poor performance at Bristol, especially for open surgery in children aged less than one year, over the period 1988 to 1995 (*Section 9.2*). Simple statistical analysis of available data might have suggested this pattern by around 1990 (*Section 9.4*), although the 1990 performance then matched the national average and so might have provided temporary reassurance. We must stress that this does not necessarily imply that there was poor performance by individual surgeons during this period. The whole system of care provided for these children, from diagnosis and referral through to post-operative care and discharge needs to be examined to look for an explanation for the observed poor performance (*Section 9.5*).

13. In terms of the future, it is clear that much greater attention needs to be paid to data quality in audit systems for paediatric cardiac services, in order to detect important but modest differences, rather than simply having the ability to flag grossly discrepant performance. We

recommend that: i) attention is paid to linkage between administrative, clinical and death registration systems; ii) that methods be developed so that all operations can be mapped onto a small number of risk categories, and iii) that simple but formal statistical procedures are introduced for institutional comparisons, monitoring individual performance, and providing for informed patient consent (*Section 10*).

List of contents.

Glossary.

Abbreviations.

ASD	Atrial Septal Defect.
AVSD	Atrial Ventricular Septal Defect.
BRI	Bristol Royal Infirmary.
CCNR	Clinical Case Note Review
CCR	Coded Clinical Records.
CSR or UKCSR	UK Cardiac Surgical Register.
CV	Coefficient of variation: the standard deviation divided by the mean.
HES	Hospital Episode Statistics.
OPCS	Office of Population Censuses and Surveys.
OPCS4	OPCS Classification of Operation and Procedures, Fourth Revision.
PAS	Patient Administration System.
PL	Perfusionists' Logs
RR	Relative risk.
SL	Surgeons' Logs.
SMR	Standardised Mortality Ratio
SWCHR	South West Congenital Heart Register.
TAPVD	Total Anomalous Pulmonary Venous Drainage.
TGA	Transposition of Great Arteries.
UBHT	United Bristol Healthcare NHS Trust.
UKCSR	UK Cardiac Surgical Register.
VSD	Ventricular Septal Defect.

1. Introduction.

1.1 Aims of the overview.

A key issue to be investigated by the Inquiry concerns the nature and outcomes of paediatric cardiac surgical services at Bristol relative to other specialist centres – referred to as Issue C in the Issues List published by the Inquiry in March 1999 (BRI Inquiry, 1999a). A large amount of oral and written evidence has been submitted to the Inquiry concerning Issue C, and major items are listed below. This overview is intended to provide a critical assessment of this evidence, with emphasis on the quality of the data sources and the reliability of any conclusions that may be drawn, and in particular to review the extent to which key data sources tell the same story or otherwise.

1.2 The main sources of published evidence.

These fall into four main categories. First, the Inquiry commissioned an initial set of expert reports on data sources (Evans, 1999; Aylin *et al*. 1999; Murray *et al*, 1999; Spiegelhalter, 1999), followed by a series of further analyses (Evans, 2000; Aylin *et al*, 2000; Spiegelhalter, 2000; Murray *et al*, 2000; Lawrence and Murray, 2000). Second, other work has been commissioned by the Inquiry, such as the Clinical Case Note Review (Hamilton and Silove, 1999; 2000) and the review on the published literature of outcomes in paediatric cardiac surgery Vardulaki *et al* (2000). Third, evidence has been submitted on general statistical issues (Curnow, 1999), use of HES data (Yates, 1997), and comments and responses on the initial statistical analysis for the Inquiry (Wisheart, 2000; Spiegelhalter *et al*, 2000; Stark, 2000a, Stark 2000b). Finally, some oral evidence is included where appropriate. In addition, the Inquiry has published papers on its approach to using relevant existing data (BRI Inquiry, 1999b), a preliminary overview of key data sources (BRI Inquiry, 1999c), and expert consultation process on analytic issues (BRI Inquiry, 1999d). Full titles are given in the Reference list, and papers should be available on the BRI Inquiry Web site.

Further responses to statistical reports (Wisheart, 2000b; Dhasmana, 2000) have been seen, but there has been insufficient time to allow any commentary to be included in this overview.

1.3 Issues in making institutional comparisons, and the approach taken in relation to the Inquiry evidence.

A number of issues have repeatedly arisen during the presentation and analysis of the statistical evidence to the BRI Inquiry. An attempt is made here to clarify the difficulties and to explain the approach taken in this overview.

1.3.1 Patients to be included in the comparison:

- *Issue*: It is important to distinguish between the evaluation of a health-care *system*, and one component of it, such as surgical performance. System evaluations should ideally be population-based so as to include, for example, those who died before surgery.
- *Approach*: The data available to the Inquiry for comparative purposes have not included information on a population basis, and so only those receiving surgery are considered. However, evaluation of the entire surgical service, rather than surgical performance alone, influences the approach to risk adjustment (see Section 1.3.5).

1.3.2 Mortality as a performance indicator:

- *Issue*: It is generally acknowledged that mortality is inadequate as a sole performance indicator.
- *Approach*: Apart from very limited components of HES and local data sources (Section 5.2), the Inquiry evidence only includes short-term mortality, and so we can make little comment on longer-term mortality, morbidity, or more subtle patient outcomes such as physical and cognitive functioning, dependency, or quality-of-life.

1.3.3 Imperfect quality of data:

- *Issue*: Institutional comparisons often need to make use of imperfect data sources, and none of the Inquiry's data sources can be considered to represent the 'truth' perfectly. Even if each data source were of perfect quality according to its own internal criteria, there would inevitably be disagreements on measures of activity and outcomes due to different sources adopting different definitions. The problem is made worse in this context by coding difficulties. All estimates of, say, the numbers of excess deaths are affected not only by statistical variability (which can be quantified) but error due to inadequate data (which is more difficult to quantify). There are particular dangers if centres differ in the quality of data, since more meticulous centres may be unreasonably penalised.

- *Approach*: Reasonable agreement between sources enhances the credibility of any conclusions, provided they are not explainable by the play of chance. Sensitivity analysis to plausible assumptions is important, and any observed differences should be large enough to overcome reasonable doubts about data quality. Ideally, all estimates would be clearly labelled as being based on imperfect data.

1.3.4 Drawing conclusions on subgroups of patients:

- *Issue:* There are two broad perspectives that need to be balanced. A 'clinical' view sees each patient as unique and hence may view aggregation into large groups as unreasonable, while a 'statistical' view seeks to make comparisons based on sufficient numbers of cases. This is made more difficult in this context by the undoubted complexity of individual cases.
- *Approach:* Coding problems and small numbers limit the reliability of conclusions on small subgroups, and it is inappropriate to seek 'causes' of individual adverse events in these sources of data. Rather, attention is focussed on the broad statistical picture. Two strategies have been taken and compared. First, simple aggregation of cases into larger groups such as all ` open' or ` closed' operations (which, of course,ignores differences in the mix of cases). Second, aggregation of individual subgroup comparisons using a stratified analysis – as discussed in Section 2.5, this should provide a more robust overall comparison.

1.3.5 Adjustment for case-mix and operative risk factors:

- *Issue:* Centres may differ in their *case-mix*, by which we mean the underlying cardiac anomalies of patients coming under their care. Surgeons' caseloads may differ in *operative-risk* factors, which include not only the type of anomalies that are included under case-mix, but also additional factors such as the age, previous medical history and current clinical condition of the patient at the time of operation. Fair comparisons of centres and/or surgeons should adjust for the appropriate factors.
- *Approach:* We have avoided use of the generic term 'risk-adjustment', since it always requires further definition depending on the purpose of the comparative exercise. When comparing whole surgical systems in centres, one should ideally concentrate on case-mix stratification: *i.e.* factors beyond all influence of the organisation. In contrast, if surgical performance alone were being compared, then a full 'operative-risk stratification' exercise may be appropriate, taking into account the precise clinical state and previous history of the patient just prior to their operation. However, this is *not* appropriate methodology when comparing the whole surgical

system, since many features at operation may be influenced by early care, timing of operation *etc*. - it is even arguable that one should not adjust for age at operation since the process of care could influence this factor (Section 8.3). Since the objective is a comparison of the systems in centres, results in the analysis have been broken into strata defined by broad procedure groups, epoch of operation and broad age-groups.

1.3.6 Summarising differences in performance:

- *Issue:* Aggregating differences in performance over many strata into a composite measure.
- *Approach:* The expected number of deaths (E) within each stratum may be estimated, assuming that Bristol were similar to the other centres in the country. This is compared to the observed number of deaths (O): the ratio O/E is equivalent to the Standardised Mortality Ratio (SMR). It is standardised for age group, procedure group and epoch. The evidence to the Inquiry also uses the difference O – E, which is termed the 'excess number of deaths'. This may be added up over many strata. Intervals and significance levels for this quantity may be calculated allowing for between-centre variability, assuming a specific statistical model: full technical details are provided in the Appendices of Aylin *et al* (1999, INQ 0013) and Spiegelhalter (1999, INQ 0015). A simpler analytic procedure has been adopted for rapid sensitivity analysis, which is described in the Technical Appendix.

1.3.7 The definition of 'divergent' performance:

- *Issue:* There are a number of reasons for centres to vary in performance (Spiegelhalter 1999, INQ 0015 0013). Chance variability is taken into account by standard statistical analysis, while measured case-mix factors can in principle be adjusted for. However, there will inevitably be further variability between centres due to unmeasured factors unrelated to quality of care. There must always be a centre which is bottom of a 'league table', and the vital issue is whether such a centre is substantially divergent from the spread among other centres.
- *Approach:* The statistical approach to this issue is discussed in detail in Aylin *et al* (1999, INQ 0013 0083) and Spiegelhalter (1999, INQ 0015 0014). Essentially, Bristol is removed from the analysis and the variability between other centres estimated. The number of excess deaths in Bristol is then estimated, and its 'significance' assessed taking into account between-centre variability.

1.3.8 Establishing causation:

- *Issue:* In the absence of randomisation or adequate controls, it is difficult to assign causation to any observed divergence in performance.
- *Approach:* Section 8 explores features at the institutional level (*e.g.* volume of surgery) and individual level (*e.g.* whether a patient has Down's syndrome) which may be associated with performance, but no definitive suggestions of causes of divergent performance can be made. When discussing factors that may explain divergent performance, a distinction should be made between 'exogenous' factors beyond the control of the system, and 'endogenous' factors such as the process of care. However, as discussed in Section 8, this is not always a straightforward distinction to make.

2. Criteria used for analysis.

The following definitions have been adopted: the extent to which the individual data sources can adhere to these definitions is covered in Section 3.

2.1 'Activity'.

An event has to be identified that measures activity and hence forms the basis for the denominator in any calculated mortality rate. The primary analysis focussed on the number of admissions/spells as the basis for comparison, although some of the data sources use operations as their measure of activity.

Critique: There is normally only one operation per admission and so there is limited difference according to which is chosen; Evans (1999, INQ 0012 0023/0026/0043) reports that there were only 5% more operations than admissions recorded in Bristol. See Section 3.1.2 for discussion on disagreements between the activity measured by HES and departmental records.

2.2 Coding of operative procedures.

Individual procedures within an admission have been coded according to the OPCS Classification of Operations and Procedures (OPCS4) (Aylin *et al* 1999, INQ 0013 0023). This coding scheme is not claimed to be a gold standard, but it is routinely used for PAS and HES data, experienced coders were available, and it allows a comparison to be made between all sources. Crucially, no suitable alternative was available (BRI Inquiry 1999d).

Critique: This system is generally unpopular with clinicians. Stark (2000a, WIT 0567, 0003) says this scheme is unfamiliar, that inappropriate codes are used, and reports anecdotal evidence of substantial disagreements with surgeons' own databases. Comparison with the surgeons' logs in Bristol (Section 4.3) reveals the difficulty of obtaining coding agreement in this complex area, and this is further discussed in Section 10.

2.3 Outcome.

A 'death' has been defined as death within 30 days of operation, whether or not it occurs in hospital, and whether or not it is related to the surgical procedure.
Critique: The highest post-operative risk is within 30 days. This is a standard choice and is the definition used by the UK Cardiac Surgical Register. The accuracy of the HES recording of deaths is discussed in Section 3.1.4.

2.4 Open/closed groups.

Two broad categories of operations have been defined: 'open' operations refer to those in which the heart is stopped and cardio-pulmonary bypass is required, while 'closed' operations do not require bypass. A scheme for mapping of OPCS4 codes to these two categories was derived through an iterative consultation process described in BRI Inquiry (1999d). Operations were excluded if they were adult, medical, either open or closed, or for unspecified reasons. A number of the excluded procedures are those that relate to diagnostic procedures such as catheterisation (Evans, 1999 INQ 0012 0022, section 4.15).
Critique: The scheme was not perfect. For example, Aylin *et al* (1999, INQ 0013 0030) report that some 2.5% of observed OPCS4 codes map onto the 13 consensus procedure groups (see Section 2.5) but not onto either of the open or closed group. Transplants were excluded, as they did not concern Bristol. A better mapping could be established, but might be expected to have little impact on the results of this exercise.

2.5 'Consensus' procedure groups.

Consultation with paediatric cardiologists and cardiac surgeons based on procedures carried out at Bristol gave rise to 13 'consensus' groups – see Aylin *et al* (1999, INQ 0013 0029) and BRI Inquiry (1999d). Consensus groups 1 to 11 were considered to be open, 12 and 13 closed. Table 2.1 shows the mapping of OPCS4 codes to the 13 groups, including whether mapping to a category of the UK Cardiac Surgical Register (CSR) was possible. Since one admission or operation may contain procedures falling in more than one group, the hierarchy shown in Table 2.2 was adopted in order that each admission could be classified into a single reasonably appropriate risk group. The code appearing highest in the table determined the group for that admission.

Critique: Stark (2000a, WIT0567 0004) identifies operations with apparently varying risks being mapped into a common procedure group, as well as significant exclusions from the exercise. Wisheart (2000a, SUB 0009) comments on disagreements between the numbers recorded in his own log and those in the procedure groups - see Section 4.3 for discussion of this. Obtaining professional agreement on an appropriate way to aggregate codes into a manageable number of groups is a difficult task due to the inherent complexity of the cases. Nevertheless, some form of grouping must be adopted or one is left with a list of individual operations and no comparison is possible. A smaller number of groups may have been preferable, for example that used by Hannan *et al* (1998), but the Inquiry had an interest in particular classes of operation and the scheme provided a means by which the available data sources could be mapped, to an incomplete extent, to a common grouping. No claim is made as to the ideal nature of this exercise, and possible improvements are discussed in Section 10.

A specific concern is random errors or systematic biases in coding. Pure random error in coding will tend to make patient groups more homogeneous and hence lead to high-risk groups having lower observed mortality, and low-risk groups having higher mortality. However, a systematic tendency for a centre to code cases into higher risk groups will bias their case-mix adjusted results. A sensitivity analysis (Section 6.4.3) has been carried out to address this issue.

2.6 Epochs.

Year-by-year comparisons provide insufficient cases to draw confident conclusions, and so periods greater than a year need to be compared. The following 4 epochs have been adopted: *1)* January 1984 to December 1987, *2)* January 1988 to December 1990, *3)* January 1991 to March 1995, *4)* April 1995 to December 1995. The boundaries of the epochs coincide with changes in the availability of the data from the different sources.

Critique: Not all data sources cover all these epochs. This overview primarily deals with Epochs 1 to 3, since Epoch 4 covers the period when the overwhelming majority of paediatric cardiac surgery was conducted by a new surgeon, and therefore is not the main focus of the comparative exercise. Epoch 4 is only considered when considering broad patterns of mortality rates in Section 6.2.

2.7 Age groups.

The following three groups for age-at-operation have been adopted: *1)* up to 90 days, *2)* 90 days to 1 year, *3)* 1 to 15 years. Analysis is mainly reported for under and over 1s since this is the finest grouping available from the CSR, although Section 8.3 considers a finer breakdown of age at operation.

2.8 Centres for comparison.

Comparison has been made only between 12 English centres, including Bristol: these are the 10 designated centres receiving supra-regional funding for paediatric cardiac surgery, plus two centres with considerable volume of activity. Bristol is numbered as Centre 1 in all comparisons – the others are identified in Table 2.3. Throughout this overview ‘elsewhere’ refers to the 11 other centres: other reports may include other smaller centres in their definition of ‘elsewhere’.

3. National Data sources.

There were two national and five local data sources available for analysis. Table 3.1 provides a summary comparison. See the individual reports for full discussion of these data sources, as well as the preliminary description provided by the Inquiry (BRI Inquiry, 1999c).

3.1 National Hospital Episode Statistics (HES).

3.1.1 Description:

This national administrative database has been in existence since 1987, and forms the basis for current performance indicators published by the Department of Health. Data are entered by non-medically qualified clinical coders as part of hospital administration, and no clinical data apart from diagnosis and interventional procedures are recorded. Yates (1997, WIT 0568) carried out a basic analysis which appeared to show the ability of HES to detect at least one high-mortality hospital in paediatric cardiac surgery, although his methodology has been subject to criticism (Gallivan 2000, INQ 0036)

3.1.2 Intrinsic quality of HES data:

Aylin *et al* (1999, INQ 0013 0015) review the evidence concerning the quality of the coding and the coverage. They conclude that HES could be reasonably reliable at a broad level of procedure groups, but judged that data before 1991 were unreliable. However, the quality of HES data has been questioned by Stark (2000a, WIT 0567 0002), particularly with regard to the lack of clinical input and the use of OPCS4 coding.

3.1.3 Accuracy of coding and 'activity':

Stark (2000a, WIT 0567 0004) reports substantially lower counts of activity (sum of operations identified as 'open' or 'closed') measured by HES and reported in Aylin *et al* (1999) and Spiegelhalter (1999), compared to the numbers of operations recorded in contemporary departmental records. Some undercount must be expected due to the Inquiry's use of admissions as a measure of activity, rather than operations as used in the departmental records. There will be additional contributions due to miscoding of records in HES, and in particular from admissions excluded from the open/closed groups (see Section 2.4)**.** It is difficult to interpret such discrepancies, as there is unknown variability between departmental record systems in, say, what constitutes an 'operation'. What is important for the Inquiry's analysis is that the same coding and exclusions (on the basis of OPCS4 codes) have been applied to all centres in a consistent manner. As noted at section 2.5 above, random errors in coding will tend to reduce differences between groups and hence between centres.

A possible marker of data quality is the ratio of episodes recorded by HES to those on KP70 (paper returns to DoH). Aylin *et al* (2000, INQ 0030 0017) found that there was excellent agreement both in Bristol and elsewhere for cardiothoracic surgery as a whole, but were unable to compare for paediatric cardiac surgery.

3.1.4 Accuracy of mortality data in HES - a linkage exercise:

A 'gold-standard' for mortality data is the ONS death registrations, and Murray *et al* (2000, INQ 0032) carried out an exhaustive and successful linkage exercise using the national HES data from April 1991 to March 1995. The results suggest HES is very accurate at capturing *in-hospital* mortality – of 714 deaths in hospital within 30 days of a procedure, only 6 (1%) were not present in HES. However, HES is only intended to record in-hospital deaths, and Table 3.2 summarises the deaths both in- and outside-hospital that are 'missed' by HES. Overall, HES did not capture 68 out of 806 30-day deaths (8.4%), with the rate for individual centres ranging from 3% to 19%. For open procedures under one year of age, 5.2% (21/407) of 30-day deaths were not included in HES – there is some variability between centres (0% to 12%), although Bristol is typical (2/47 = 5% not included). Of course, there is an unknown number of deaths not recorded in HES and for whom the linkage failed. However, in our judgement it is likely that 30-day deaths occurring after discharge will be the major source of missing 30-day deaths in HES.

3.1.5 Missing outcome data:

Aylin *et al* (1999, INQ 0013 0031) report that a number of admissions have missing outcomes, which may be due to failure to link episodes within a admission or simply that no outcome was recorded. Bristol had an excess of such missing data (19% and 3% for open and closed operations under 90 days, compared with 3% and 1% elsewhere). These incomplete records have generally been omitted from the analysis in this synthesis. Sensitivity to this is investigated in Section 6.4.

3.2 UK Cardiac Surgical Register (CSR).

3.2.1 Description:

This register was established by the Society of Cardiothoracic Surgeons of Great Britain and Ireland in 1977, and collects anonymised data from centres on activity and mortality rates. Ages are categorised into under or over one year, and the latter group includes congenital heart operations on over 15's. Collection followed calendar years until 1993, when it changed to financial years: hence data from January 1993 to March 1993 does not feature in the register.

3.2.2 Intrinsic quality of CSR data:

Lawrence and Murray (2000) report a survey of units that contributed data to the CSR during the period covered by the Inquiry. They conclude that there has been considerable variability in the way units collected data, in relation to staff, sources of data and the definitions applied. Nevertheless, surgeons tend to have more confidence in the data that they themselves have provided to the register, compared to that provided for HES, both for procedures and deaths.

3.2.3 Accuracy of coding and 'activity':

Lawrence and Murray (2000) report that the classification of complex diagnoses is very inconsistent between units, with the use of category 'Miscellaneous – other' varying from 8.1% to

0% across units, with Bristol being 3.1%. The CSR is primarily based on diagnoses and so mapping into operative procedure groups may be somewhat contrived. As noted by Stark (2000a, WIT0567 0006), of particular concern is the mapping to the consensus groups G2 and G3, where G3 is corrective repair of transposition of the great arteries (TGA), which in the OPCS4 coding scheme used for the other data sources corresponds to the later ‘switch’ operation. However, in the CSR there appears to have been substantial use of this category for earlier operations such as ‘Mustard’ and ‘Senning’, which leads to poor agreement between data sources for these groups in the earlier part of the period covered by the Inquiry. This problem was recognised in the initial analysis (Murray *et al*, 1999), and Stark (2000b, WIT 0567 0003) suggests these groups should be discounted within the CSR data. Table 6.2 acknowledges this issue.

3.2.4 Mortality outcomes in CSR:

Lawrence and Murray (2000) conclude that reporting of mortality is unreliable with some units suggesting under-reporting and, less commonly, over-reporting of deaths.

3.3 Comparison of HES and CSR data at a national level.

The HES and CSR data may be compared across all centres in this analysis for Epoch 3 (1991-March 1995). Murray *et al* (1999) have carried out a detailed analysis which is summarised here, combining age groups 1 and 2 and hence categorising by less than or greater than 1 year. Table 3.3 shows the number of cases and number of deaths from both sources, broken down by open/closed procedures, aged under and over 1, by centre, and by consensus procedure group. Although the ratios should be 1 if there were perfect agreement between HES and CSR, differences in definitions of activity, outcomes and coding schemes mean that one should not expect close concordance even were they both high-quality data sources. In particular, one might expect that the CSR would show more activity as there may be more than one operation per admission and over 15s are included up to 1993/4, and also more deaths since HES only records in-hospital mortality. However, it is plausible that there would be less systematic bias in using clinical coders for HES than in the highly variable submissions to the CSR.

The major conclusions are that there tends to be both more cases and more deaths reported in CSR than HES, leading to reasonably comparable death rates. Across the centres there is a broadly consistent pattern with Bristol (Centre 1) being typical. Centre 3 appears to have made very low returns to the CSR as to activity, although the number of deaths matches HES well. Agreement at the level of individual procedure groups is poorer: reasonable agreement is only seen for Groups 1, 3, 4, 8, 10. Groups 2 and 3 show better agreement if combined, as might be expected from the discussion in Section 3.2.3. Section 6.4.3 features a sensitivity analysis in which centres with high discrepancy are removed from the analysis.

3.4 Conclusions on national data.

The reasonably consistent patterns seen in Table 3.3 lend added weight to the HES evidence, as do the KP70 and linkage exercises carried out to assess the quality of the recorded activity and outcomes in HES. There is no evidence that Bristol was at variance with the national pattern in HES reporting. The CSR data must be treated with great caution at the level of individual procedure groups. The crucial issue is whether the undoubted inaccuracies are sufficient to cast doubt on any observed divergent performance.

4. Local Data sources on Bristol.

4.1 Patient Administration System (PAS).

Evans (1999, INQ 0013) reports that the Bristol PAS provides both returns on activity to the Department of Health and supports administration of UBHT. Its characteristics are summarised in Table 3.1. The Inquiry has heard that the coding team at UBHT was considered of good quality (see the Transcript of oral hearings for 14 July 1999, paragraphs 12 to 15). Deaths out of hospital may not be recorded, although such deaths are sometimes added in later and these may not feature in the return made to the Hospital Episode Statistics (HES).

4.2 Coded Clinical Records (CCR).

Evans (1999) describes how UBHT provided to the Inquiry the medical records of all children who underwent cardiac surgery over the period 1984 to 1995, identified through the PAS and Surgeons' Logs (SL). Relevant cases may not have been identified, and incompleteness of clinical notes is a problem that may limit the conclusions that can be drawn from this source, though it is clear that completeness for open cases is very high when comparing activity across all the sources of data for Bristol.

4.3 Bristol Surgeons' Logs (SL).

Hand-written and typed logs of the operations of two surgeons (Mr Dhasmana and Mr Wisheart) have been provided to the Inquiry. These cover the whole period of interest, contain details of the patient and the operation and its outcome, and had been used as a basis for internal audit and submissions to the UK Cardiac Surgical Register (CSR) although with no formal validation. Only operations at the Bristol Royal Infirmary, which would be expected to be only 'open' operations, are covered.

Mr Wisheart later provided a computerised version of his log, and Evans (2000, INQ 0029 0018) has matched it against the Inquiry's coded version of the Surgeons' logs (SL). There is good concordance in overall numbers of activity and mortality, but comparison within operation groups is difficult as there are nearly 200 different operation descriptions used by Mr Wisheart. No attempt could be made to translate these into the 13 consensus groups, and there was no similar source for other surgeons at Bristol.

4.4 South West Congenital Heart Register (SWCHR).

The background and potential quality limitations of this cardiologists' database are discussed in detail in BRI Inquiry (1999c) and Murray *et al* (1999). They conclude that although there have been no systematic data-collection procedures, definitions or follow-up, the maintenance of common staff should help consistency. It could form the basis for a comprehensive audit, but substantial work would be required to validate the data.

4.5 Perfusionists' Logs (PL).

Evans (2000, INQ 0029 0027) describes the coding of this data source, which should be both an accurate source of information on deaths at operation (although not afterwards) and, as all children in the log had open heart surgery by definition, the 'gold-standard' for defining an 'open' procedure.

5. Activity and outcomes in Bristol

5.1 Mortality outcomes.

5.1.1 Comparison of sources on mortality:

Spiegelhalter (1999, INQ 0015 0021) compares six sources with regard to apparent activity and number of deaths, using both the 13 consensus procedure groups and (except for SWCHR) the open/closed classification. Certain sources do not provide data on all consensus groups, and occasional operations inappropriately classified (using OPCS4 codes) as closed in Surgeons' Logs or Perfusionists' logs are not considered.

In Table 5.1 we consider just Epoch 3, with revised data for Fontans (G9) for the CCR and SL, and using the perfusionists' logs (PL) only for activity. The relative variability between the sources is summarised by the Coefficient of Variation (CV) - values of CV around 20% could be considered as having reasonable agreement, and less than 10% as having good agreement.

For many of the individual procedure groups the agreement is reasonable: for example, Fallot (G1: CV = 15), TAPVD (G4: CV = 22), AVSD (G5: CV = 11) and the sum over procedure groups for which all sources are available (CV = 20). For open operations in general the agreement is remarkably good (CV = 4). There is poor agreement of CSR with other sources for Groups 2 and 3 for reasons discussed in Section 3.2.3 but, if CSR is ignored, agreement is fairly good for G3 (switches). Better agreement may be attributable to procedures that can be fairly unambiguously coded. PAS appears to record more admissions, and the set of clinical records in CCR was partly derived from surgeons' logs, so CCR should include all cases in SL. Disagreement on operation dates between different clinical sources can lead to minor differences between SL, CCR and PAS.

5.1.2 Annual mortality in open operations for under 1s.

Table 5.2 focuses on open operations on under 1s, and the different sources agree reasonably well with regard to both activity and the number of deaths. There is an apparent drop in the mortality rate in 1990, although according to the Surgeons' Logs (SL), only one baby less than 90 days was operated on in 1990 compared to around 7 in other years. However, the numbers each year are too small to draw any firm conclusions on individual years. Linkage of HES with ONS data does not tend to find deaths among those in whom outcome is unknown, suggesting that missing outcomes might reasonably be expected to be survivors.

5.2 Morbidity outcomes.

In response to the findings of Aylin (1999, INQ 0013 0028) of an apparently higher rate of neurological complications in Bristol, Evans (2000, INQ 0029) examined evidence on complication rates in local data sources. The Surgeons' Logs (SL) did not, predictably, contain good information on longer-term outcomes, while both in the coded clinical records (CCR) and PAS the recorded neurological complication rates among survivors of open surgery was very low (1.9% and less than 1% respectively). There was poor agreement between sources and Evans (2000, INQ 0029 0016) concluded that there was under-reporting in all centres, with Bristol possibly being slightly more accurate in its reporting. The Clinical Case Note Review (CCNR) did look in detail at the possibility of disability in those who had not died at 30 days, but with only 40 cases, even though they were preferentially sampled from high risk groups, the number with any disability was very small (4, all "moderate" disability). It is therefore not possible to draw confident conclusions on the true morbidity rate or make comparisons with other centres.

5.3 Conclusions on Bristol activity and outcomes.

There are clear limitations to all sources, and none is subject to defined procedures for data collection, follow-up and validation. It would be fair to say that none is held in high regard as a source of reliable evidence for clinical audit. However, Evans (1999) concludes that where direct comparison is sensible, the pattern is similar and there are no startling discrepancies. Although there is no gold standard for comparison, the Bristol PAS system appears of reasonable quality, and hence this lends confidence to Bristol returns to the national HES database. Our overall comparison suggests that the different sources agree well on the open operations in general and for many specific procedures.

The main findings of interest concern mortality rate for open surgery in under 1s. Overall, sources agree that the mortality rate was around 25 – 30% during the period under scrutiny, although with considerable variability between different procedures. The routine data sources available form an inappropriate basis for any firm conclusions concerning morbidity rates in Bristol.

6 Comparison of Bristol with national performance.

6.1 Analyses carried out.

Murray *et al* (1999, INQ 0014) and Aylin *et al* (1999, INQ 0013) report comparisons between Bristol and elsewhere using CSR and HES data respectively, and including detailed analysis of the relative rank of Bristol in the 12 centres carrying out surgery in England. Spiegelhalter (1999, INQ 0015) synthesised this evidence and discussed each of the procedure groups in detail. Aylin *et al* (2000, INQ 0030) repeated the HES analysis using age categories of under and over 1. Each of these analyses was based on a common method: examine variability among centres other than Bristol, predict the number of deaths expected in a centre with Bristol's activity were it 'typical' of centres elsewhere, and subtract this expected number of deaths from the observed number to estimate 'excess' mortality.

6.2 Mortality outcomes.

6.2.1 Overall summary:

Table 6.1 summarises the results for all open surgery, case-mix stratified open surgery and closed operations, using CSR and HES data for all epochs, under and over 1s, using results contained in Spiegelhalter (1999, INQ 0015) and Aylin *et al* (2000, INQ 0030 0073). Although the CSR data report statistically significant excess mortality for Bristol in over 1s during 1988-1990, the primary finding from both CSR and HES is of excess mortality from 1991-1995 in open operations in under 1s, in which the mortality rate in Bristol was around double that in other centres. This difference is retained after stratifying for operative group, which is the available determinant for case-mix. There is no evidence for excess mortality in closed operations, or for open operations in over 1s from 1991-1995. Reported mortality for open operations in under 1s fell in other centres from 21% in 1984-1987 to 12% in 1991-1995. Bristol appears not to have followed that pattern of improvement. There is no evidence of excess mortality in Bristol during Epoch 4, although activity in Bristol was too small to draw any firm conclusion.

We emphasise that the estimated total excess deaths for HES depends on the age-stratification used: the excess risk is greater in younger children: for all open operations in epoch 3 the total is 30.1 when dividing only into under and over 1s (Table 6.1) and 34.3 when including a < 90 day category (Aylin *et al*, 1999).

6.2.2 Mortality in procedure groups during 1991-1995: under 1s.

Table 6.2 summarises the analyses for Epoch 3 (1991-1995) for under 1s, using results from Spiegelhalter (1999, INQ 0015 0060) and Aylin et al (2000, INQ 0030 0032). For CSR, Groups 2 and 3 have been highlighted as being unreliable for reasons discussed in Section 3.2.3. There are predictable disagreements between the two sources of data. HES identifies excess mortality with 95% confidence for switches (G3), AVSD (G5), ASD (G6), open operations stratified for case-mix, (G1 to G11), and all open operations taken together.

6.2.3. Open surgery – comparison with other centres:

Figure 6.1 shows the mortality rates and 95% confidence intervals for each of the 12 centres carrying out open surgery, based on CSR and HES data for relevant epochs and divided into under and over 1s. The variability between the centres is immediately apparent. In under 1s, Bristol (Centre 1) had the third highest mortality rate reported to the CSR in 1988-1990 and the highest rate in both CSR and HES 1991-1995. The estimated probability that Bristol had the highest true mortality rate in under 1s during 1991-1995 is 88% using CSR data (Spiegelhalter 1999, INQ 0015 0060) and 97% using HES data (Aylin *et al*, 2000 INQ 0030 0073).

Table 6.3 presents the annual mortality rates for open surgery in under 1s for Bristol and elsewhere. The CSR results show that each year between 1988 and 1994 (with the exception of 1990), Bristol had either the highest or near the highest mortality rate for open surgery in under 1s. This is reinforced by the HES data between 1991 and 1994. It is clear that Bristol's activity was consistently below the median in the country, and the possible association of mortality with volume is discussed in Section 8.4.

6.2.4. Divergence of other centres:

Spiegelhalter (1999) and Aylin *et al* (2000) provide estimates of excess mortality for each of the 12 centres, treating each centre in the same manner as Bristol in the main analysis. From this analysis and Figure 6.1 it can be seen that only one other centre, Centre 10, had consistent evidence of divergent performance and this was for open operations in over 1s. It was revealed by the Inquiry in November 1999 that this was Harefield hospital. This finding must be treated with caution. Harefield has been an innovative centre for transplant surgery and these operations are included in the CSR (although not in the HES open category), and it also has a reputation for taking difficult cases from abroad.

Stark (2000b, WIT 0567 0002) observes that the excess mortalities for each centre calculated using HES and CSR data do not always closely coincide. In particular, as observed in Section 3.3 and shown clearly in Figure 6.1, Centre 3 reports far higher mortality to the CSR than may be calculated from HES. Section 6.4.3 considers the sensitivity of the conclusions on Bristol to the removal of this and other centres.

6.3 Non-mortality outcomes.

Aylin *et al* (1999) examined outcomes other than mortality using the HES data for 1991 to 1995, although they emphasise the limitations of this approach. They found that for open operations, Bristol recorded a higher proportion of admissions where central nervous system, cardiac, respiratory and urinary complications occurred, when compared with other centres. However, as reported in Section 5.2, Bristol's reporting of complications may be more complete than in other centres, and in any case the data sources are very unreliable. Aylin *et al* (1999) also report that for both open and closed operations, substantially fewer patients were discharged from Bristol within 7 days compared to elsewhere. This finding must be interpreted with caution, since many factors could influence length of stay.

6.4 Further questions.

6.4.1 Is the excess mortality restricted to switches and AVSDs?

Table 6.2 shows that switches and AVSDs are prominent contributors to the observed overall excess mortality, and Wisheart (2000a, SUB 0009) questions whether there are other contributors. The information in Table 6.4 can be extracted from Tables 7.3.1, 7.3.2 and 7.4.5 of Spiegelhalter (1999).

Excluding switches and AVSDs, the CSR show a significant 83% increase in mortality over other centres. The HES data show a 44% increase in mortality over centres elsewhere, although this is not statistically significant at conventional levels. Because of the known lack of distinction in the CSR between switch (group 3) and inter-atrial repair (group 2), group 2 might also be excluded: Table 6.4 shows that this slightly increases the contrast between Bristol and elsewhere.

6.4.2 Is the excess mortality influenced by the missing outcomes in HES?

The majority of the HES analyses have ignored admissions with missing outcome data, Wisheart (2000a, SUB 0009) questions whether this seriously biases the results. We have carried out a simple analysis to examine what the impact of these missing outcomes might be, taking the most optimistic view that in Bristol they all were survivors. The data shown in Table 6.5 were taken from Aylin *et al* (1999, INQ 0013 0055-0057), and only consider pooled open operations. There were 48 cases in Bristol with missing outcomes. If they had been included in the analysis, and had they all survived, then they would have added 0 to the observed number of deaths, and added around 3.6 to the expected number of deaths. Thus the total excess deaths would have been reduced by around 3.6, from 34.3 (the estimated total when using the age-stratification in Table 6.5) to 30.7. Note that this analysis does *not* assume that missing outcomes elsewhere were survivors (although the linkage exercise does not suggest missing outcomes elsewhere are at increased risk of being deaths). Thus, even if we assume that all missing outcomes at Bristol were survivors, there is little effect on the findings. It therefore does not appear that missing outcomes makes the HES analysis unreliable.

6.4.3 Can we base conclusions on imperfect data?

Given that the data sources for these comparisons have such clear limitations, it is reasonable to ask whether in this light any reliable conclusions can be drawn. Statistical significance alone is not a sufficient guide, as it only indicates quantifiable random error and not systematic reporting or coding biases. **The crucial issue is not whether the data are 'true', since they manifestly contain errors, but whether such errors are likely to be great enough to overcome the observed pattern in the data.** To address this issue a number of sensitivity analyses to possible shortcomings in the data have been carried out and the detailed results are shown in the Technical Appendix for the data of primary interest: open operations on under 1s from 1988. The summary conclusions are:

1. *Centres for which doubts exist concerning national data.* Centres 3, 4, 5 were removed from all analyses, since Table 3.3 reveals these as having the highest discrepancy between HES and CSR data, with more than 20% difference in death rates. This has the effect of increasing Bristol's divergence, presumably because the removed centres are smaller and tend to have higher mortality on such patients (see Section 8.1).
2. *Procedure groups.* Procedure Groups 2 and 3 (Interatrial and other repairs of TGA) and 5 (AVSDs) were removed from the analysis, as described in Section 6.4.1. Groups 2 and 3 suffer from known coding overlap in CSR. This increases the divergence in 1988-1990, but reduces it

considerably in 1991-1995. However, as suggested in the simple analysis of Table 6.4, there is still considerable evidence of divergent performance even without these higher-risk groups.

3. *Undercount of mortality in HES.* The mortality rate in the HES data for each centre was increased by the 'undercount' for open operations in under 1s that was detected in the linkage study (Section 3.1.4) and shown in Table 3.2. The undercount ranges from 0 to 12% of deaths, with an average of 5.2%. This has little effect on the conclusions, presumably because Bristol has an average undercount (5%).
4. *An analysis favouring Bristol.* An 'extreme' scenario is one in which we choose, for each centre and procedure group, the results from HES or CSR according to the following rule: for Bristol, we select the results with the *lower* mortality, for each other centre we select the results with the *higher* mortality. This stringent comparison still shows strong evidence of divergent performance for all open operations, but stratifying for case-mix leads to borderline evidence for excess mortality in Bristol, with an estimate of around 50% increase in odds of death.

These sensitivity analyses certainly have an influence on the accuracy with which excess mortality in Bristol can be estimated. However, in our view, the magnitude of the observed divergent performance is such that reasonable variations in assumptions are not sufficient to cast the conclusions into doubt. This is discussed further in Section 9.

7. Comparison with published sources on operative mortality.

7.1 Sources of published data.

Vardulaki *et al* (2000, INQ 0039) have carried out a systematic review of published research on mortality data for five main procedures, corresponding to our consensus groups G3 (switch operations for transposition), G4 (TAPVD), G5 (AVSD), G7 (Truncus) and G9 (Fontan). They acknowledge the difficulty of generalising from published sources, as there is likely to be selective reporting from centres of excellence. There is substantial heterogeneity between sources, and there is a general pattern of improvements over time.

In addition, Hannan *et al* (1998) report a study on 7169 cases in New York between 1992 and 1995 and provide mortality rates for many procedures, including those studied by Vardulaki *et al.* This covers almost exactly the period of Epoch 3.

7.2 Comparison with published sources.

Table 7.1 compares the results derived from Spiegelhalter (1999) with those reported by Vardulaki *et al* (2000) and Hannan *et al* (1998). Direct comparison is difficult, as Vardulaki *et al* (2000) do not stratify for age but do report results at a finer level of detail than our consensus groups. In addition, results for the period 1991-1995 are not directly reported, and so the rates given in Table 7.1 are taken by eye from their Figures and so can only be considered rough estimates. We note that the Hannan results fit closely with those reported by Vardulaki *et al* (although they do not contribute to their analysis since the publication date (1998) lies outside the range adopted by Vardulaki *et al*). The HES and CSR results agree well with the international data.

7.3 Conclusions.

Stark (2000b, WIT 0567) suggests that the mortality rates given for England in Inquiry reports appear low, and reports anecdotal mortality rates from Toronto Children's Hospital of 26% in open surgery between 1991 and 1995 (age-group unknown). However, Table 7.1 suggests that the results from non-Bristol centres derived from both HES and CSR are compatible with published data from elsewhere, and in particular New York State. Bristol appears to have divergent performance from international published sources.

8. Investigation of possible factors associated with divergent performance in Bristol.

Having observed evidence of divergent performance in Bristol, a number of possible explanatory factors have been investigated. For each factor it is preferable to identify two characteristics:

1. The **level** at which it is measured (*i.e.* institutional / patient). Purely institutional factors, such as staffing level, organisation of care, experience of staff, and volume of surgery, can only provide indirect explanation for variability between centres since it is not clear how they directly influence the risk experienced by individuals. The only institutional level factor available for investigation was volume of surgery.

2. The extent to which the factor is '**exogenous/endogenous'** to the system being evaluated, *i.e.* the extent to which the factor is susceptible to influence or change by the system. This is more of a grey-scale than an absolute classification of factors. For example, geographic clustering in births of difficult cases should be an exogenous factor since it is not under any control of the system – the adjustment for broad procedure group attempts to deal with this, although even the choice of procedure is to some extent subject to clinical influence. Similarly, comorbidity factors such as Down's syndrome should be exogenous but the incidence may be influenced by referral practices. Status at admission, comorbidity and timing of surgery have been examined at the individual level – however, for each of these it is not immediately clear to what extent they are directly influenced by the cardiac surgical system under evaluation, and hence they cannot clearly be labelled as either exogenous or endogenous..

8.1 Institutional factors: volume of surgery.

8.1.1 Results of analysis.

Spiegelhalter (2000, INQ 0031) reports an analysis of the association between volume of surgery and mortality outcomes, using data from the CSR and HES. For open operations in under 1s, and for arterial switches and AVSD in particular, there was strong and consistent evidence for an association between mortality rates and volume (not taking into account any data from Bristol), in which higher-volume centres have lower mortality. Stratifying for operation-mix, or including the results from Bristol, strengthened this association. Figure 8.1 summarises the results for open operations in under 1s for 1991-1995, estimating the relative reduction in risk per additional 10 cases per year to be around 3% and 4% in CSR and HES respectively. We note that, according to the HES data, centres carrying out less than 200 cases in four years (one a week) had a mortality rate of 15% (not including Bristol) or 17% (including Bristol), while those carrying out more than one a week had a mortality rate of 10%. The relationship also appears to hold in earlier epochs: the CSR data estimate the relative reduction in risk per additional 10 cases per year to be 9% (95% interval -6% to 22%) in 1984-1987 and 6% (95% interval 2% to 10%) in 1988-1990.

8.1.2 Interpretation.

Spiegelhalter (2000, INQ 0031) estimated that a hospital carrying out 120 open operations a year on patients aged under 1 in 1991-1995 would be expected to have an underlying mortality rate 25% lower than one carrying out only 40 such operations. If the hospitals had exactly the same age- and operations mix, this reduction is increased to 35%. These are percentage changes relative to the underlying risk, and so implications in terms of the difference in numbers of deaths depend on the context. However, considerable caution is needed in interpreting these results, and it does not necessarily follow that concentrating resources in fewer centres would reduce mortality rates, since volume may be associated with lower mortality without being a direct cause. Using the association found in other centres, it was estimated that only around 12% (HES) or 17% (CSR) of the excess mortality observed in Bristol in open operations in under 1s might be explainable by the lower volume of surgery being carried out in Bristol.

8.2 Patient factors: status at admission and comorbidity.

8.2.1 Results of analysis.

Aylin *et al* (1999;2000) explored a number of factors which might account for the high reported mortality following operations at UBHT. Table 8.1 summarises some of these findings for open operations. It suggests that age-mix cannot account for the high mortality at UBHT, not only because age-specific mortality was higher in all age groups compared with elsewhere, but also that UBHT operated on a much smaller proportion of the youngest (higher-risk) babies aged under 90 days (7%) than elsewhere (16%). Mortality in children with Down's syndrome from other centres (excluding UBHT) is not significantly greater (8%) than children operated on without this disorder (7%), so differences in the proportion of children with Down's syndrome treated at UBHT are not likely to account for differences in mortality. Patients transferred from other units to centres (excluding UBHT), have a higher mortality (14%) than patients admitted by other means (5%), but UBHT had a much lower level of transferred patients (6%) than elsewhere (22%), so this again cannot account for higher mortality in UBHT. Emergency admissions have a higher mortality (12%) than non-emergency admissions (7%) in other centres (excluding UBHT), however UBHT admitted a smaller proportion of emergencies (7%) than other units (10%), making this an unlikely explanation for their high mortality. Aylin *et al* (1999) also found that mortality did not vary by levels of socio-economic deprivation of patients and that the distribution of primary diagnoses in UBHT patients was similar to other centres.

8.2.2 Interpretation

The analyses are summarised below.

Factor	What is associated with higher mortality	How Bristol compares to average	Comments
Volume	Low	Lower	Explains small proportion of excess
Age at operation	Low	Higher average age	Marked divergence of practice at Bristol
Proportion of Down's	High	Lower proportion	Does not explain excess
Transfers	High	Lower proportion	Does not explain excess
Emergency Admission	High	Lower proportion	Does not explain excess
Socio-Economic Deprivation	High	No difference	Does not explain excess

HES data is limited in the information it provides about status at admission and comorbidities, and the analyses suggest that these factors cannot explain the high mortality reported at UBHT. The role of age at surgery is now examined in more detail.

8.3 Patient factors: timing of surgery.

8.3.1 Results of analysis.

The age at which surgery takes place may be influenced by the system of care, and hence may be an explanatory factor in divergent performance. Only HES data provide comparative data with precise dates of operation, and this is only available for Epochs 3 and 4 (April 1991 to December 1995). Aylin *et al* (2000) derived Figure 8.2, which shows the number of open operations taking place at each month of age up to 18 months, in Bristol and in the other centres combined in Epoch 3. It is clear that Bristol has a peak of activity at 11 months, in contrast with a steady decline in activity with increasing age seen in other centres. Aylin *et al* (2000) show that this peak in activity apparently only occurs before March 1994, and an 11th -month peak in Bristol between 1990 and 1993 is confirmed by examination of local data sources, specifically PAS, CCR, Surgeons' Logs and perfusionists' logs: Figure 8.3 shows the age distribution recorded in the PAS system. Figure 8.4 shows the age-specific activity in all 12 centres between April 1991 and March 1994, which shows the 11th -month peak was unique to Bristol. The peak is apparent in the larger groups (G2 Inter-atrial repair, G5 AVSD, G7 VSD) but other groups are too small to judge.

Table 8.2 considers age at operation in 3 month intervals up to two years old, and annually and five-yearly thereafter. Bristol had only carried out 21% of its surgery before the age of 9 months, compared to 39% elsewhere. However, in the following three months up to their first birthday, 14% of all surgery (60 operations) was carried out compared to 4% elsewhere. The mortality rate elsewhere is reasonably constant after the first three months, and can be used to estimate the expected number of deaths expected in Bristol. The estimated total of excess deaths is 34.1 (slightly different from previous estimates due to the finer age-stratification): those operated on in the first three months of life contribute 16.8 and those in the final three months before their first birthday contribute 7.8, approximately 25% of the total. This excess of 7.8 is due both to the number of operations taking place, mainly in the 11th month, and the fact that the mortality rate of 18% (19/60) is significantly higher than the mortality elsewhere (5%, 49/381, P< 0.001).

8.3.2 Interpretation.

According to the HES data, around a quarter of the age-stratified excess mortality (7.8 out of 34.1) in open surgery in 1991-1995 is associated with operations performed within three months of the first birthday. Aylin *et al* (2000) identify AVSD operations as a primary contributor to this: 41% (14/34) of AVSD surgery was in this period with a 50% mortality rate (7/14).

The pattern for timing of surgery shown in Table 8.2 suggests that the operations carried out just prior to the first birthday may have been delayed from earlier rather than brought forward, since Bristol had carried out only 21% of its open operations before the age of 9 months, compared to 39% elsewhere. In particular, 40% of AVSD operations performed elsewhere during 1991-1995 were carried out in the first six months of life, compared to 9% in Bristol.

8.4 Patient factors: the process of care.

A Clinical Case Note Review (CCNR) (Hamilton and Silove1999, INQ 0016) was carried out on a stratified sample of UBHT medical records in order to provide peer judgement as to the adequacy of care received. Full interpretation of the results of the CCNR is given elsewhere (Hamilton and Silove, 2000). Their Executive Summary concludes that the care received by 70% of the children was adequate, leaving 30% whose care was less than adequate to different degrees. For just over 5% of children, it was considered that different management would reasonably be expected to have made a difference to outcome. The reviewers provided criticisms of a range of aspects of process of care throughout the surgical system, including delays in treatment, shortcomings in cardiological contribution and organisation of intensive care, the split site and general organisational failings. The conduct of surgery was one of the criticised factors but was not particularly highlighted. In the stratified sample, over half the deaths (21/40) were considered to have received less than adequate care in which different management might have made, or would reasonably be expected to have made, a difference in outcome. We do not, however, have a comparative group in order to see whether other centres have similar systemic difficulties, and so cannot know whether similar criticisms could be levelled at other centres over this period.

9. Overall Conclusions.

9.1 The available data sources.

The two national sources, HES and the CSR, are admittedly imperfect. Both suffer considerably from lack of agreed operating procedures for ensuring completeness and accuracy of activity, coding and outcome results. Both the OPCS4 coding scheme and the use of non-clinical coders lead HES to be viewed with suspicion by clinicians. There are also strong concerns about variability between centres in the CSR's coding procedures and recording of mortality. Even if they were meticulously completed, agreement between the two sources could not be expected due to their different criteria. However, HES was found to be surprisingly accurate in its recording of in-hospital mortality and, with certain clear exceptions, the sources described the same broad picture.

The local sources were found to provide good agreement on activity and overall mortality, although comparison at a finer level was sensitive to the coding conventions used. Nevertheless, the six sources on Bristol's activity and outcome agree well for open operations in general and, to a lesser but still reasonable extent, for finer consensus procedure groups of interest. Where there is disagreement, then there are clear reasons, usually resulting in transfer of operations between two groups.

9.2 Evidence for divergent performance of Bristol.

There is no evidence of excess mortality in closed operations carried out in Bristol, and limited evidence in open operations on children aged over 1 year. However, there is strong and consistent evidence of excess mortality in open operations in children less than 1 year old at operation. It is estimated from HES data that in the period 1991-1995, 24.1 (95% confidence interval 12 to 34) of 41 recorded deaths are in excess of that expected were Bristol a 'typical' centre: finer age-stratification increases the estimated excess mortality. CSR data suggest the excess mortality dates back at least to 1988. Open procedures on children aged less than 1 that can be identified with reasonable consistency as having excess mortality include 'switches', operations for TAPVD, AVSD and, although rare in this age group, ASD. It is to be expected that excess mortality is easier to detect in higher risk groups.

The excess mortality was not just restricted to AVSDs and switch operations, and the conclusions are robust to admissions with missing outcomes. National mortality rates were comparable to those in the international literature. One other centre had a consistent pattern of excess mortality in open operations in children over 1 year, but there were no other centres with consistently divergent raised mortality in the younger age group.

9.3 Explanation for divergent outcomes.

At an institutional level, Bristol is a low-volume centre and other low-volume centres have been associated with higher mortality rates. Regardless of the policy implications of this finding, it is apparent that only a limited proportion of Bristol's excess mortality can be 'explained' by this indirect risk factor.

Bristol differed from the national pattern in some aspects of status and comorbidity (Section 8.2), but these characteristics do not apparently explain divergent performance. There is also no evidence that Bristol had systematically higher-risk case-mix. The most striking factor is the high incidence of surgery in the period immediately preceding the first birthday between 1990 and March 1994. Around 25% of the excess mortality in open surgery is associated with a peak of operations in the three months before their first birthday, mainly in the 11th month. The evidence in Table 8.2 suggests that these cases may have been delayed from earlier surgery, rather than being operations that might normally have been carried out after their first birthday.

9.4 What might have been known?

It is possible to consider what simple analyses might have been performed using the data and the statistical tools that would have been readily available to the surgeons at the time. The participating centres in the CSR were supplied with detailed annual reports giving mortality rates split by age and procedure, aggregated over all participating centres. This would have allowed a centre to compare its mortality rates with corresponding national figures.

Open surgery performed in children aged under one year is an appropriate subgroup to monitor, since these children are at high risk and include the majority of deaths. In this group the ratio of the overall mortality rates at Bristol to the rates for other centres in England for 1985 through to 1995 were 1.18, 1.21, 1.24, 2.04, 1.93, 0.79, 2.05, 1.19, 3.18, 1.67 and 0.50 respectively. A chi-squared test performed each year would have given a crude indication of whether the local mortality rate differed from the national rate by more than could be explained by chance. Using such a test, the data for 1988, 1989, 1991 and 1993 are statistically significant at the 5% level. If years had been pooled in pairs or triplets to give larger numbers, then the results for 85/86 and 86/87 are non-significant, as are the results for 85/86/87, but the results for 87/88 and 86/87/88 are statistically significant. Thus with any of these approaches, it is not until the data for 1988 were included that the divergence from the national rates became statistically significant, and this was reinforced by the data for 1989. Given that there was a delay of the order of 18 months before the CSR data were fed back to centres, it would have been 1990 before the data from the CSR might have given any reason for concern, and the independent reinforcement for the 1989 data, which would become available during 1991, would have heightened this concern. However, the data for 1990 then came back into line with national figures (see Table 6.3), which might have been taken as reassurance that any problems which might have existed previously had been resolved.

This final point illustrates the difficulty of interpreting crude data based on small numbers of patients each year. Taking running totals from three year periods the data are statistically significant for 86/87/88, 87/88/89, 88/89/90, 89/90/91, (borderline non-significant for 90/91/92), 91/92/93, 92/93/94 and 93/94/95. Clearly there is a consistent and on-going pattern of poor outcomes, but it is difficult to know what weight should have been put on these data at the time, with there being questions over the data quality and with inadequate statistical tools to adjust for case mix and to analyse accumulating data from many different centres. A related difficult question is the extent to which the responsibility lay with individual centres to interpret their own data, versus the role of the Society of Cardiothoracic Surgeons which with access to the full data for each centre was in a better position to analyse and interpret the data. Of course, statistical analysis is only one aspect of monitoring clinical performance.

9.5 Conclusions.

We again emphasise that statistically significant findings, taken on their own, are insufficient grounds for confidently identifying divergent performance when there are grave and well-founded doubts about the quality of the data sources. It is also important to emphasise that there are many areas in which there was no evidence of poor performance in Bristol. Nevertheless, although no data source can be considered as exactly representing the true state of affairs, their consistency, and the fact that they are derived in very different manners, suggests that their findings reinforce each other.

The single most compelling aspect of the data is the magnitude of the discrepancy between the outcomes observed at Bristol and those observed elsewhere. For children aged under one year undergoing open surgery between 1988 and 1994, the observed mortality rate at Bristol was roughly double that observed elsewhere in 5 out of 7 years. While the national trend over this period was for mortality rates to fall substantially, no such trend was seen in the Bristol results. In spite of the many flaws in the data sources, we do not believe that statistical variation or any systematic bias in data collection can explain a divergence of this magnitude. We therefore conclude that there is strong evidence of divergent performance at Bristol in the areas identified above, and we believe that the imperfections of the data do not cast serious doubt on these conclusions.

10. Proposals for the future.

In the light of our combined experience in working on the Inquiry data sources, we would now like to make a range of proposals regarding future monitoring systems in paediatric cardiac surgery, which may also have more general relevance to other settings.

10.1 What kind of comparative studies?

Given the limitations of key data sources, the Inquiry's statistical evidence is necessarily focussed on short-term mortality outcomes of those who received surgery. However, our investigation suggests that a much broader perspective is appropriate if comparing systems of care, perhaps more in line with public-health investigations. Such a population-based approach examines all cases of interest, whether or not they come to surgery, and is ill served by current systems and initiatives. Such a perspective becomes particularly important if, for example, surgeons started to avoid operating on high-risk patients in order to improve apparent mortality rates – this may be a consequence of a "blame culture" that emphasises penalties for apparent poor performance.

Overall, there needs to be clarity as to the precise objectives of any comparative exercise. A crucial distinction is whether the objective is to identify grossly discrepant performance, or whether the aim is more educational, with individual surgeons or units following their performance year by year, looking for minor problems, or seeking to identify the benefits of minor changes in practice ('closing the audit loop'). Many articles on clinical audit see the latter as being the aim of audit. However, the statistical work commissioned for the Inquiry shows that, even given perfect data sources and even if there were no difference in case-mix, statistical variability would mean that data would need to be accumulated over many years to detect modest but important differences in mortality rates. Given the many flaws that have been identified in existing data sources, it is clear that only gross divergence could have been identified with any degree of confidence. If, for example, the mortality rate for open operations in under 1s observed at Bristol had been 50% higher than elsewhere rather than 100% higher, it would have been very difficult to exclude the possibility that the difference had arisen through a combination of differences in case-mix, in the coding of operative procedures, and in the thoroughness of achieving follow-up data.

Existing data sources can and should be improved, for example by introducing routine linkage of HES records to national mortality records in order to confirm mortality data. Equally, data collection procedures require much greater standardisation, with adequate training of the staff involved, and regular feedback of data so that quality can be maintained. The objectives of any audit exercise need to be reviewed carefully in the light of the sample size that is likely to be available for any comparisons, and the magnitude of the likely biases.

10.2 What source of comparative data?

There are currently two broad approaches to sources of data for comparative exercise: an administrative model and a clinical model. A simplistic comparison of their advantages and disadvantages might comprise:

Administrative model	**Clinical model**
Example: PAS/HES	Example: Clinical databases
Pro:	**Pro:**
Established system for pooling data	Contribute data to refined CSR
Trained coders	Data controlled by clinicians
Facility for linkage for population studies	Clinical data
Accurate mortality records	Individualised risk-assessment
All centres contribute	
Anti:	**Anti:**
Non-medical coders	Lack of standardisation between centres
OPCS4 not ideal	No agreed coding scheme
No adjustment for clinical risk factors	Lack of linkage for mortality *etc*
Only mortality outcomes	Voluntary involvement

This is clearly a simplification and many compromises are possible between these archetypes. We believe that each approach has a role, but that development in isolation to each other is wasteful and inefficient. Our experience in this exercise has been that neither approach has been satisfactory.

We therefore strongly recommend the development of linkage schemes between ONS national statistics and administrative systems, and between administrative and clinical systems.

A separate but related issue is the question of how to raise the credibility of routine data, especially for clinicians. It is clear that, for whatever reason, many clinicians have no confidence in the HES data. Any future developments of routine data systems needs to address the issue of how best to ensure data are clinically valid and meaningful, possibly based on the promotion of a sense of 'ownership' of the data by clinicians.

10.3 What kind of coding scheme and groupings?

This investigation has revealed the difficulty in developing an agreed coding scheme for complex cases in paediatric cardiology, that allows both accurate description of individuals and a facility for pooling cases in a clinically acceptable way. The fact that such a scheme was not in general use in this country forced the Inquiry to use data sources and coding schemes that were criticised by clinicians. Furthermore it is unclear how coding is this context will develop, in the light of the recent publication of two independent schemes under the auspices of the Society of Cardiothoracic Surgeons (Mavroudis and Jacobs, 2000) and the Association of European Paediatric Cardiology (Franklin *et al*, 1999) respectively.

While it is desirable to make comparisons between precisely-defined homogenous groups of patients, we feel this has been over-emphasised and that for monitoring purposes it is better to develop broad groups into which activity can be allocated with reasonable accuracy. Finer distinctions can always be made for more focused clinical purposes.

We recommend the adoption of a scheme in which each procedure is placed in one of a small number of risk categories. Whatever detailed clinical coding scheme is adopted, it is important that it can be mapped both onto such a simplified system for monitoring, and into the codes used by administrative systems.

10.4 How can statistical methods help in analysing performance?

Comparative data may be useful in many ways, and a variety of statistical tools are available to help exercise due caution.

Institutional comparisons: Curnow (1999, WIT 0361 0002) emphasised that statistical techniques may be used to indicate when an institution may have passed either a ‘warning’ threshold, which might trigger further investigation, or an ‘alarm’ threshold which might indicate immediate action. The setting of such thresholds requires a combination of statistical and clinical judgement, and allowance for random error and inevitable between-institution variability. Outcomes should be risk-adjusted where feasible, although this might be only into broad groups (see Section 10.3) as to much disaggregation reduces precision. Statistical methods can also prevent undue attention to spurious ranking into ‘league tables’ (Marshall and Spiegelhalter, 1998).

Clinical comparisons within institutions: Availability of good data sources would allow, for example, the cumulative monitoring of risk-adjusted excess mortality (or another performance indicator) for individual clinicians, as is being increasingly adopted in adult cardiac surgery. Care is required if formal thresholds are used for monitoring.

Patient information: There is likely to be increased demand for patients to be given numerical risk assessments when asked for consent for surgery. This is not a straightforward matter: does one give the data for the individual surgeon, institution, or nationally, and for what period? How much should data-based statistics be adjusted for subjective opinions concerning the individual patient? There are statistical methods that can help with individualised risk-assessment, discounting historical data, pooling local with national data, and critiquing past numerical risk assessments.

We recommend the informed introduction of formal statistical procedures for institutional comparisons, monitoring individual clinical performance, and providing for informed consent of patients.

Acknowledgements

The help of the Secretariat to the Inquiry, particularly Ruth Chadwick, is gratefully acknowledged. Mr Leslie Hamilton and Dr Kate Bull provided invaluable help on coding, and members of the statistical Expert Panel to the Inquiry provided valuable suggestions. The authors, however, take full responsibility for the opinions expressed in this report.

References

Aylin P, Alves B, Cook A, Bennett J, Bottle A, Best N, Catena B and Elliott P. (1999). Analysis of Hospital Episode Statistics for the Bristol Royal Infirmary Inquiry. *Bristol Royal Infirmary Inquiry*, INQ 0013, November 1999.

Aylin P, Alves B, Cook A, Best N (2000). Supplementary analysis of Hospital Episode Statistics for the Bristol Royal Infirmary Inquiry. *Bristol Royal Infirmary Inquiry*, INQ 0030.

BRI Inquiry (1999a) Issues list for Part I of the Inquiry. *Bristol Royal Infirmary Inquiry*, February 1999.

BRI Inquiry (1999b) The Inquiry's approach to making use of relevant data sources. *Bristol Royal Infirmary Inquiry*, March 1999.

BRI Inquiry (1999c) Preliminary overview of existing data sources relevant to the Inquiry's remit. *Bristol Royal Infirmary Inquiry*, July 1999.

BRI Inquiry (1999d) Synthesis of statistical sources: a note on expert consultation on key analytical issues. *Bristol Royal Infirmary Inquiry*, November 1999.

Curnow R N (1999). Potential contribution of statistical evidence to the Inquiry. *Bristol Royal Infirmary Inquiry*, WIT 0361, November 1999.

Dhasmana J D (2000) Response to supplementary statistical evidence to the Inquiry. . *Bristol Royal Infirmary Inquiry*.

Evans SJW (1999). A report on local data relating to children who received cardiac surgery under the terms of reference of the Bristol Royal Infirmary Inquiry. *Bristol Royal Infirmary Inquiry*, INQ 0012, November 1999.

Evans S J W (2000). Further Reports based on Local Sources of Data for the Bristol Royal Infirmary Inquiry. *Bristol Royal Infirmary Inquiry*, INQ 0029.

Gallivan S (2000). Issues related to papers by Professor John Yates presented to the BRI Inquiry. *Bristol Royal Infirmary Inquiry*, INQ 0036.

Franklin R C G, Anderson, RH, Daniels, O, Elliott, M, Gewillig, MHML, Ghisla, R, Krogmann, ON, Ulmer, HE, Stocker, FP (1999) Report of the Coding Committee of the Association for European Paediatric Cardiology. *Cardiology in the Young*, **9**, 633-658.

Hamilton, L and Silove, E (1999): Preliminary Report on the Clinical Case Note Review. *Bristol Royal Infirmary Inquiry*, November 1999.

Hamilton, L and Silove, E (2000) Report on the Clinical Case Note Review. *Bristol Royal Infirmary Inquiry*.

Hannan E L, Racz M, Kavey R E, Quaegebeur J M, Williams, R (1998) Pediatric cardiac surgery: The effect of hospital and surgeon volume on in-hospital mortality. *Pediatrics*, **101**, 963-969

Lawrence A E and Murray G D (2000). The UK Cardiac Surgical Register: assessment of data quality issues for the Bristol Royal Infirmary Inquiry. *Bristol Royal Infirmary Inquiry*, INQ 0033.

Marshall EC and Spiegelhalter DJ (1998). League tables of *in vitro* fertilisation clinics: how confident can we be about the rankings? *British Medical Journal*, **316**, 1701-4.

Mavroudis C and Jacobs J P (2000) Congenital heart surgery nomenclature and database project: overview and minimum dataset. *Ann Thoracic Surgery*, **69**, 2-17.

Murray G D, Lawrence AE and Pollock J (1999). UK Cardiac Surgical Register and South West Congenital Heart Register: a statistical analysis and review of key data sources relevant to the inquiry’s remit. *Bristol Royal Infirmary Inquiry*, INQ 0014, November 1999.

Murray G D, Lawrence A E and Boyd J (2000) Linkage of Hospital Episode Statistics (HES) data to Office for National Statistics (ONS) mortality records, *Bristol Royal Infirmary Inquiry,* INQ 0032.

Spiegelhalter D J (1999) An initial synthesis of statistical sources concerning the nature and outcomes of paediatric cardiac surgical services at Bristol relative to other specialist centres from 1984 to 1995. *Bristol Royal Infirmary Inquiry,* INQ 0015, November 1999.

Spiegelhalter D J (2000) An investigation into the relationship between mortality and volume of cases in surgery for congenital heart disease from 1984 to 1995. *Bristol Royal Infirmary Inquiry*, INQ 0031.

Spiegelhalter D J, Evans S, Aylin P and Murray G D (2000) A response to 'Submissions on behalf of Mr J D Wisheart: Appendix 2. The Inquiry's Statistical Analysis'. *Bristol Royal Infirmary Inquiry*, INQ 0034.

Stark J (2000a) Comments on statistical analysis and review of outcomes of paediatric cardiac services at Bristol and other specialist centres. Part 1 *Bristol Royal Infirmary Inquiry*, WIT 0567.

Stark J (2000b) Comments on statistical analysis and review of outcomes of paediatric cardiac services at Bristol and other specialist centres, Part 2. *Bristol Royal Infirmary Inquiry,* WIT 0567.

Vardulaki K A, Bennett-Lloyd B D, O'Riordan P A, Reeves B C, Tsang V T C and Black N A (2000) A systematic review of the outcomes of paediatric open-heart surgery, *Bristol Royal Infirmary Inquiry*, INQ 0039.

Wisheart J (2000a) Submissions on behalf of Mr J D Wisheart: Appendix 2. The Inquiry's Statistical Analysis. *Bristol Royal Infirmary Inquiry*, SUB0009.

Wisheart J (2000b) Response of Mr J D Wisheart to a group of statistical papers. *Bristol Royal Infirmary Inquiry*, WIT 0120 0468-0499.

Yates J (1997) A case study exploring the early identification of performance failure in an acute hospital, *Bristol Royal Infirmary Inquiry*, WIT 0568.

Technical Appendix

The full analysis is described in the Appendices of Aylin *et al* (1999, INQ 0013) and Spiegelhalter (1999, INQ 0015). This involved a logistic regression model for each stratum defined by epoch, age-group and procedure group, in which a random effect was associated with each non-Bristol centre. The variance component within each stratum were assumed to be drawn from a hierarchical prior distribution, and this provided a predictive distribution over the effect in a new centre, and hence a predictive distribution over the number of deaths in a centre with Bristol's activity. The difference to the observed number of deaths gave both an estimate and interval for the excess mortality. This analysis is intended to allow for important sources of variation and so will be fairly conservative.

The full analysis is time consuming and unsuited for repeated sensitivity analyses. For this reason an intermediate analysis has been carried out, again based on a logistic regression but assuming independent fixed effects for each centre in each stratum defined by epoch and age-group, with main effects fitted for procedure group. The contrast between Bristol's effect and the average of the effects in the other centres was obtained using 'Helmert contrasts'. The natural way to report the results is by odds ratios and their 95% confidence intervals: the more restrictive assumptions in this simpler model tend to make the results somewhat less conservative than the full analysis.

The baseline analysis in the table below shows that for 1991-1995, both HES and CSR data, taken at face value, provide strong evidence for excess mortality in Bristol (odds ratio greater than 1). Stratifying for case-mix does not decrease this estimate. The CSR data from 1988-1990 provide some evidence for excess mortality. The results for the sensitivity analyses are given below and summarised in Section 6.4.3.

Appendix Table. Results of sensitivity analysis on open operations under one year of age.

			Unstratified			***Stratified for case-mix***		
Data	Epoch	Analysis	Odds ratio	95% interval	p-value	Odds ratio	95% interval	p-value
CSR	**1988-1990**	**Baseline**	**1.54**	**1.03 to 2.30**	**.02**	**1.62**	**0.96 to 2.72**	**.04**
		a) Without centres 3,4,5	1.73	1.01 to 2.98	.03	1.86	0.93 to 3.73	.04
		b) Without groups 2,3,5	2.38	1.44 to 3.94	.001	2.88	1.5 to 5.53	.001
CSR	**1991-1995**	**Baseline**	**1.95**	**1.40 to 2.72**	**.0001**	**1.90**	**1.22 to 2.95**	**.002**
		a) Without centres 3,4,5	1.98	1.27 to 3.10	.001	1.97	1.09 to 3.56	.01
		b) Without groups 2,3,5	1.96	1.28 to 3.01	.001	1.90	0.95 to 3.82	.04
HES	**1991-1995**	**Baseline**	**2.72**	**1.92 to 3.86**	**.0001**	**3.43**	**2.21 to 5.32**	**.0001**
		a) Without centres 3,4,5	2.74	1.71 to 4.38	.0001	3.46	1.85 to 6.48	.0001
		b) Without groups 2,3,5	1.80	1.10 to 2.95	.01	2.58	1.46 to 4.58	.001
		c) Compensating for deaths missed by HES	2.67	1.71 to 4.17	.0001	3.2	1.47 to 6.95	.002
HES + CSR	**1991-1995**	**d)** "Extreme" choice of HES and CSR	1.83	1.31 to 2.54	.0001	1.46	0.97 to 2.21	.04

a) Removes centres with higher discrepancies between HES and CSR data
b) Removes procedures for TGA (Groups 2 and 3) and AVSD (Group 5).
c) Increases mortality rates in each centre by apparent undercount indicated in Table 3.2
d) For each procedure group, selects HES or CSR data depending on which has the ***lower*** mortality (if Bristol), or whichever has the ***higher*** mortality (if elsewhere).

'Odds ratio' is the estimated ratio of the odds of dying in Bristol compared to the average odds of dying in other centres. P-values are 1-sided.

Table 2.1. Paediatric Cardiac Surgical Procedures by Group: OPCS4 Codes mapped by UKCSR Categories

Group	OPCS4 Procedure Code	Description	Map to UKCSR
G1	K04	Tetralogy of Fallot	Yes
G2	K05	Interatrial TGA	Yes
G3	K06	Other TGAs (- switch)	Yes
G4	K07	Repair of TAPVD	Yes
G5	K09 excluding K09.4	Repair of CAVSD (complete not partial)	Yes
G6	K10, K20 and K09.4	Closure of secundum and sinus venosus ASD	Yes
G7	K11 (only on its own or with K10 or +/- L02;	Closure of VSD	Yes
	K11 is superior code to K10)		
G8	L01.1	Truncus arteriosus	Yes
G9	K19.1, K19.2, K19.4 + L09	Fontan type operations	Yes
G10	K26, K28, K31.2, K31.4, K37	Aortic, pulmonary valve and paravalve procedures	Yes
G11	K25, K31.1, K34.1, K38	Mitral valve procedures	Yes
G12	L05, L06, L07,L08	Closed shunts	No
G13	L23.1, 2 or 3 [- if K code with it, code as K not L]	Coarctation procedures	Yes (simple coarctation)

Table 2.2. Synthesis of Statistical Sources: Primary Procedure Ranking

Rank	Group	Description
1	G 8	Truncus Arteriosus
2	G 9	Fontan type operations
3	G4	TAPVD
4	G 3	Other TGA
5	G 2	Interatrial TGA
6	G 5	AVSD
7	G 11	Mitral valve procedures
8	G 10	Aortic and pulmonary valve procedures
9	G 1	Tetralogy of Fallot
10	G 7	Closure of VSD
11	G 6	Closure of ASD
12	G 12	Closed Shunts
13	G 13	Simple Coarctation

(Note: If any operation features procedures falling into more than one of the consensus groups G1 to G13, the operation is assigned to the highest ranking Group. This table draws on expert clinical advice on the most common combinations of procedures and mortality rates.

Table 2.3. Centres included in the comparative exercise

Code	Centre	Hospital
1	Bristol	Bristol Royal Infirmary
2	Leicester	Glenfield Hospital
3	Leeds	Killingbeck Hospital
4	Oxford	The John Radcliffe Hospital
5	Guys	Guys Hospital
6	Liverpool	Royal Liverpool Children's Hospital
7	Southampton	Southampton General Hospital
8	Great Ormond Street	Great Ormond St Hospital
9	Newcastle	Freeman Hospital
10	Harefield	Harefield Hospital
11	Birmingham	Birmingham Children's Hospital
12	Brompton	Brompton Hospital

(**Note**: Centres were assigned Inquiry codes 2-12 at random. Centres were identified at the BRI Inquiry on November 3rd 1999.)

Table 3.1 Outline comparison of six available sources of data on Bristol’s activity and outcomes.

	PAS **Patient Administration System**	**CCR** **Coded Clinical Records**	**SL** **Surgeons' Logs**	**CHR** **South West Congenital Heart Register**	**HES** **Hospital Episode Statistics**	**CSR** **UK Cardiac Surgical Register**
Purpose	Hospital administration and returns to HES.	Medical records.	Personal record for audit and constructing CSR returns	Epidemiological information and clinical back-up.	National administration system. Now used for DoH performance indicators.	Professional register for comparative anonymous audit.
Completed by:	Coders.	Medical personnel.	Surgeons.	Cardiologists.	Derived from PAS.	Surgical team in Bristol.
‘Activity’	Episodes linked to form admissions/spells.	Operations.	Operations.	Operations.	Episodes linked to form admissions/spells.	Diagnostic group subdivided by ‘corrective’ and palliative’ operations.
Grouping for synthesis.	Existing OPCS4 codes.	Coded into OPCS4 by expert team.	Coded into OPCS4 by expert team.	Mapped by expert consensus.	Existing OPCS4 codes.	Mapped by expert consensus in 13 groups. Open/closed provided on report.
Epochs available	2: 1988 - 1990 3: 1991 - Mar 95 4: Apr 95 - Dec 95	1: 1984 - 1987 2: 1988 - 1990 3: 1991 - Mar 95 4: Apr 95 - Dec 95	1: 1984 - 1987 2: 1988 - 1990 3: 1991 - Mar 95 4: Apr 95 - Dec 95	1: 1984 - 1987 2: 1988 - 1990 3: 1991 - Mar 95 4: Apr 95 - Dec 95	3: Mar 1991 - Mar 95 4: Apr 95 - Dec 95	1: 1985 - 1987 2: 1988 - 1990 3: 1991 - Mar 95 (not Jan – Mar 93) 4: Apr 95 – Mar 96
Age groups	1: 0 – 90 days 2: 90 days – 1 year 3: 1 year +	1: 0 – 90 days 2: 90 days – 1 year 3: 1 year +	1: 0 – 90 days 2: 90 days – 1 year 3: 1 year +	1: 0 – 90 days 2: 90 days – 1 year 3: 1 year +	1: 0 – 90 days 2: 90 days – 1 year 3: 1 year +	1+2: 0 – 1 year 3: 1 year +
Comments.	Considered to be of good quality. Late deaths may be missed.	Usual problems with incomplete medical records. Not all relevant records identified.	Only covers ‘open’ surgery at BRI.	‘Child’ is basis for records. Stable team.	Quality depends on local PAS systems. Missing outcomes on some admissions.	Completed by a range of staff . No validation. Missing years for some centres.

Table 3.2 Deaths identified by linkage within 30 days of procedures but NOT captured by HES

	All cases		Open procedures in children under one	
Centre	**'Missed' deaths ***	**%**	**'Missed' deaths ***	**%**
1	3 / 74	4	2 / 43	5
2	3 / 41	7	0 / 25	0
3	13 / 67	19	3 / 27	11
4	6 / 40	15	2 / 24	8
5	2 / 42	5	0 / 24	0
6	3 / 91	3	2 / 43	5
7	1 / 32	3	0 / 20	0
8	12 / 108	11	7 / 59	12
9	6 / 54	11	0 / 25	0
10	2 / 73	3	2 / 27	7
11	8 / 91	9	2 / 57	4
12	8 / 72	11	1 / 32	3
Elsewhere	1 / 21	5	0 / 1	0
Total	68 / 806	8.4	21 / 407	5.2

* 'Missed' in inverted commas since HES is not designed to capture 30-day mortality.

Table 3.3 Comparison of UKCSR returns with HES data for 1991-1994. Admissions are grouped by Surgery, Age, Centre, Consensus Group and Year.

	Number of Cases			Number of Deaths			Ratio of Death Rates
	UKCSR	HES	Ratio	UKCSR	HES	Ratio	
Surgery							
Open	8227	7544	1.09	698	577	1.21	1.10
Closed	2898	2817	1.03	86	100	0.86	0.83
Total	11125	10361	1.07	784	677	1.16	1.07
Age							
Under 1	5360	5078	1.06	500	461	1.08	1.01
Over 1	5765	5283	1.09	284	216	1.31	1.20
Centre							
1	830	750	1.11	79	69	1.14	0.96
2	758	603	1.26	43	37	1.16	0.92
3	556	1068	0.52	50	54	0.93	1.78
4	295	481	0.61	27	35	0.77	1.24
5	664	557	1.19	61	39	1.56	1.30
6	1372	1460	0.94	96	86	1.12	1.19
7	819	639	1.28	40	32	1.25	0.98
8	1187	965	1.23	82	64	1.28	1.04
9	805	609	1.32	49	46	1.07	0.81
10	709	574	1.24	87	70	1.24	1.01
11	1921	1492	1.29	95	84	1.13	0.86
12	1209	1163	1.04	75	61	1.23	1.15
Group							
G1	921	837	1.10	57	45	1.27	1.10
G2	76	158	0.48	15	17	0.88	1.76
G3	685	644	1.06	89	67	1.33	1.13
G4	203	217	0.94	28	27	1.04	1.01
G5	553	749	0.74	65	68	0.96	1.25
G6	1525	1182	1.29	11	18	0.61	0.46
G7	1141	1280	0.89	26	56	0.46	0.50
G8	123	97	1.27	30	30	1.00	0.76
G9	340	620	0.55	42	65	0.65	1.16
G10	827	893	0.93	42	43	0.98	1.03
G11	160	247	0.65	15	27	0.56	0.82
G13	757	632	1.20	12	17	0.71	0.59
Year							
1991	3255	2710	1.20	254	191	1.33	1.09
1992	3403	2944	1.16	245	203	1.21	1.03
1993	2352	2311	1.02	142	142	1.00	0.97
1994	2115	2396	0.88	143	141	1.01	1.14
1995	3509	2178	1.61	195	125	1.56	0.84

For 1991 and 1992 the UKCSR data cover calendar years but the HES data cover financial years.
The HES data for 1995 cover only the nine month period April 1995 to December 1995.
Centre 8 was dropped from the HES data for 1993/4 and 1994/5, as there were no corresponding UKCSR returns for those years.
HES figures include cases where outcome is unknown.
Comparison of mortality rates use a denominator which excludes these cases.
Data for 1995 are only included in the tabulation by year.
Some sub-totals for HES data may disagree slightly with other tables due to small differences in definition.

Table 5.1 UBHT activity and outcomes in Epoch 3 (1991-1995) - all ages, comparison of six different data sources

		Number of operations								No of deaths							Mortality rates (%)						
		CCR	**SL**	**PL**	**CSR**	**PAS**	**HES***	**SW CHR**	**CV**	**CCR**	**SL**	**CSR**	**PAS**	**HES**	**SW CHR**	**CV**	**CCR**	**SL**	**CSR**	**PAS**	**HES †**	**SW CHR**	**CV**
G1	Tetralogy of Fallot	55	61	62	58	54	51	45	11%	7	6	6	7	5	6	12%	13	10	10	13	11	13	15%
G2	Interatrial TGA	25	27	25	4	29	24	26	37%	2	2	0	4	3	2	61%	8	7	0	14	17	8	59%
G3	Other TGAs (- switch)	28	27	27	45	29	27	25	23%	11	12	10	12	11	11	7%	39	44	22	41	58	44	21%
G4	Repair of TAPVD	24	24	23	19	20	18	21	11%	6	6	6	7	5	9	21%	25	25	32	35	36	43	22%
G5	Repair of CAVSD (complete not partial)	47	56	17	41	39	36		33%	14	15	11	13	12		12%	30	27	27	33	35		11%
G6	Closure of secundum and sinus venosus ASD	89	89	91	126	92	92	92	14%	1	2	2	4	5	12	93%	1	2	2	4	6	13	96%
G7	Closure of VSD	101	101	114	90	115	105	72	15%	4	4	0	4	1	3	66%	4	4	0	3	1	4	66%
G8	Truncus arteriosus	8	10	10	8	9	8		11%	5	4	2	4	3		32%	63	40	25	44	60		32%
G9	Fontan type operations	57	45	47	39	43	42	30	19%	7	8	7	5	5	4	26%	12	18	18	12	13	13	21%
G10	Aortic, pulmonary valve and paravalve procedures	55	45	71	34	57	51	42	24%	3	2	1	4	5	3	47%	5	4	3	7	10	7	39%
G11	Mitral valve procedures	19	6	7	9	23	23	18	50%	3	1	0	3	3	2	63%	16	17	0	13	13	11	52%
G12	Closed shunts	57				66	65		8%	10			7	7		22%	18			11	11		31%
G13	Coarctation procedures	92			61	101	92	80	18%	2		0	2	2	3	61%	2		0	2	2	4	66%
Total		657	491	534	494	677	634	451		75	75	62	76	67	55		11	11	13	7	10	7	
Total of Groups 1-4, 6, 7, 9-11		453	425	424	467	462	433	371	8%	42	44	43	50	43	52	16%	9	10	10	11	10	14	20%
G88	Open	476	454	530¶	563	501	505		8%	68	61	71	65	62		6%	14	13	13	13	14		4%
G99	Closed	136			267	160	245		32%	5		8	7	7		19%	4		3	4	3		16%
Total		**612**			**830**	**661**	**750**		**22%**	**73**		**79**	**72**	**69**		**9%**	**12**		**10**	**11**	**10**		**12%**

* Based on all admissions, whether outcome known or not

† Based only on admissions where outcome known

CV: Coefficient of Variation, defined to be mean / standard deviation.

¶The 530 operations in PL are the total of all those occurring in Epoch 3, not just those classified as open using OPCS-4 code

PAS: Patient Administration system, CCR: Coded Clinical Records, SL: Surgeons' Logs, PL: Perfusionists' logs

SWCHR: South West Congenital Heart Register, HES: Hospital Episode Statistics, CSR: UK Cardiac Surgical Register.
(NB Further examination of data has resulted in small changes from a previous version of this table.)

Table 5.2 Comparison of annual admissions, deaths and mortality rates in Bristol derived from local data sources open operations on under-ones.

Year	No. of admissions					No. of deaths						Mortality rates (%)					
	CCR	PAS	SL	HES	CSR	CCR	PAS	SL	HES	CSR	HES/ONS	CCR	PAS	SL	HES	CSR	HES/ONS
1984	11	-	9	-	-	4	-	3	-	-	-	36	-	33	-	-	-
1985	14	-	15	-	14	2	-	2	-	3	-	14	-	13	-	21	-
1986	28	-	28	-	24	9	-	6	-	6	-	32	-	21	-	25	-
1987	25	-	35	-	25	5	-	7	-	7	-	20	-	20	-	28	-
1988	29	29	36	-	29	10	8	9	-	11	-	34	28	25	-	38	-
1989	40	41	46	-	40	14	14	15	-	15	-	35	34	33	-	37	-
1990	38	33	46	-	39	4	3	7	-	5	-	11	9	15	-	13	-
1991	43	41	48	(32)[1]	46	12	12	16	(7)	14	(8)	28	29	33	25	30	24
1992	45	49	58	50[2]	53	12	12	11	12	8	11	27	24	19	30	15	22
1993	50	54	55	50[3]	(50)	11	12	11	11	(14)	12	22	22	20	29	28	23
1994	32	31	40	32[4]	(32)	11	10	9	10	(7)	10	34	32	23	32	22	31
1995	25	28	14	(8)[5]	(50)	3	3	1	(1)	(3)	(1)	12	11	7	17	6	12

Notes:

1. Brackets indicate figures which are not directly comparable for the following reasons:
HES : figures for 1991 do not include January to March; figures for 1995 include only January to March
CSR : figures for 1993 to 1995 are financial years, so exclude January to March of the current year and include January to March of the following year

2. HES figures include cases where outcome is unknown, as follows:
[1] : 4 [2] : 10 [3] : 12 [4] : 1 [5] : 2
Corresponding mortality rates use a denominator which excludes these cases.

3. HES/ONS figures are HES records confirmed as 30-day deaths by linkage with ONS register of deaths

Table 6.1 Summary of analyses comparing Bristol and elsewhere: all open, case-mix stratified open, and closed operations.

Source	Epoch	Under 1s					Over 1s				
		Mort else.	Mort Bris	Obs	Exp	Excs	Mort else.	Mort Bris	Obs	Exp	Excs
All open operations											
CSR	1: 1984-1987	21%	25%	16	14.0	2.0	8%	8%	24	23.3	.7
	2: 1988-1990	18%	29%	31	22.3	8.7	7%	12%	37	22.4	14.6
	3: 1991-1995	12%	24%	43	24.0	*19.0	5%	7%	28	22.8	5.2
	4: 1995-1996	*12%*	*6%*	*3*	*6*	*-3.0*	*3%*	*1%*	*2*	*4.4*	*-2.4*
HES	3: 1991-1995	12%	29%	41	16.9	*24.1	5%	7%	21	15.0	6.0
	4: 1995	*12%*	*4%*	*1*	*2.8*	*-1.8*	*4%*	*0%*	*0*	*3.7*	*-3.7*
Open operations – case-mix stratified											
CSR	1: 1984-1987			15	13.0	2.0			16	13.7	2.3
	2: 1988-1990			26	19.0	7.0			24	12.7	*11.3
	3: 1991-1995			30	17.1	*12.9			15	12.2	2.8
	4: 1995-1996			*2*	*2.4*	*-.4*			*0*	*1.7*	*-1.7*
HES	3: 1991-1995			42	14.8	*27.2			16	12.0	4.0
	4: 1995			*1*	*2.7*	*-1.7*			*0*	*1.7*	*-1.7*
All closed operations											
CSR	1: 1984-1987	6%	12%	18	9.4	8.6	2%	2%	3	2.0	1.0
	2: 1988-1990	5%	8%	12	7.9	4.1	2%	3%	4	2.6	1.4
	3: 1991-1995	3%	3%	5	6.2	-1.2	3%	3%	3	2.5	0.5
	4: 1995-1996	*3%*	*0%*	*0*	*1.5*	*-1.5*	*1%*	*4%*	*1*	*.3*	*.7*
HES	3: 1991-1995	4%	5%	7	6.9	0.1	2%	0%	0	1.7	-1.7
	4: 1995	*9%*	*0%*	*0*	*2.8*	*-2.8*	*0%*	*4%*	*1*	*0.0*	*1.0*

Epoch 4 (1995) based on simplified analysis.
* indicates > 95% confidence that excess mortality > 0
Obs = Observed deaths
Exp = Expected deaths
Excs = Estimated excess deaths

Table 6.2 Summary of analyses for Epoch 3, April 1991 to Mar 1995. Operations on under-ones.

		HES					CSR				
		Mort. Else.	Mort Bris.	Obs	Exp	Excs	Mort. else.	Mort Bris.	Obs	Exp	Excs
G1	Fallot type	6%	0%	0	0.2	-0.2	8%	0%	0	0.2	-0.2
G2	Interatrial TGA	11%	13%	2	1.6	0.4	*(28%)*	*(0%)*	*0*	*0.8*	*-0.8*
G3	Other TGAs	10%	77%	10	1.5	*8.5	*(13%)*	*(28%)*	*10*	*5.0*	*5.0*
G4	TAPVD	14%	36%	5	2.0	3.0	14%	33%	6	2.6	3.4
G5	AVSD	12%	48%	11	3.0	*8.0	13%	25%	8	4.5	3.5
G6	ASD	7%	50%	5	0.7	*4.3	2%	40%	2	0.1	*1.9
G7	VSD	6%	0%	0	2.7	-2.7	3%	0%	0	1.4	-1.4
G8	Truncus	32%	75%	3	1.3	1.7	25%	29%	2	1.9	0.1
G9	Fontan type	17%	50%	2	0.7	1.3	33%	100%	1	0.4	0.6
G10	Aortic, pulm	10%	50%	2	0.4	1.6	13%	100%	1	0.2	0.8
G11	Mitral valve	24%	67%	2	0.7	1.3	14%	0%	0	0.2	-0.2
G12	Closed shunts	10%	8%	3	3.8	-0.8					
G13	Coarctation	4%	3%	2	2.6	-0.6	2%	0%	0	0.6	-0.6
G1-11	Stratified open	10%	30%	42	14.8	*27.2	10%	19%	30	17.1	*12.9
	Open	11%	29%	41	16.9	*24.1	12%	24%	43	24.0	*19.0
	Closed	4%	5%	7	6.9	0.1	3%	3%	5	6.2	-1.2

* indicates > 95% confidence that excess mortality > 0
Obs = Observed deaths,
Exp = Expected deaths,
Excs = Estimated excess deaths.

Table 6.3 Annual mortality rate for open operations on under-ones: Bristol compared to 11 other centres.

CSR Data

	No. of admissions				No. of deaths				Mortality rate			
	Bristol	11 other centres			*Bristol*	11 other centres			*Bristol*	11 other centres		
		Min	**Med**	**Max**		**Min**	**Med**	**Max**		**Min**	**Med**	**Max**
1984	-	-	-	-	-	-	-	-	-	-	-	-
1985	*14*	5	38	82	*3*	1	7	15	*21*	5	20	46
1986	*24*	14	35	79	*6*	2	8	23	*25*	6	21	48
1987	*25*	17	36	96	*7*	4	9	22	*28*	14	23	47
1988	*29*	5	39	125	*11*	1	8	23	*38*	11	20	28
1989	*40*	19	46	146	*15*	3	11	21	*37*	11	21	43
1990	*39*	21	48	179	*5*	3	9	18	*13*	8	14	43
1991	*46*	13	55	173	*14*	1	10	29	*30*	6	10	29
1992	*53*	44	65	221	*8*	4	9	28	*15*	5	13	18
1993	*50*	42	48	188	*14*	3	5	14	*28*	4	10	15
1994	*32*	42	75	207	*7*	6	9	26	*22*	9	12	23
1995	*50*	11	95	101	*3*	0	10	22	*6*	0	10	22

HES data

	No. of admissions (HES data)				No. of deaths (HES data)				Mortality rate			
	Bristol	11 other centres			*Bristol*	11 other centres			*Bristol*	11 other centres		
		Min	**Med**	**Max**		**Min**	**Med**	**Max**		**Min**	**Med**	**Max**
1991	*33*	22	41	87	*7*	4	6	14	*24*	7	16	31
1992	*50*	30	62	158	*13*	3	9	22	*32*	7	13	33
1993	*52*	30	60	180	*11*	2	6	18	*27*	6	12	22
1994	*32*	27	66	160	*10*	3	7	15	*32*	4	11	19
1995	*8*	5	19	41	*1*	0	2	6	*17*	0	11	22

Table 6.4. Results for open operations, under one year of age, 1991-1995, excluding switch (group 3) and AVSD (group 5) operations.

Source	Mortality elsewhere	Mortality in Bristol	Estimated excess deaths	Simple p-value
HES	248/2201 = 11 %	21/130 = 16 %	6.4	.12
CSR	279/2257 = 12 %	25/111 = 22 %	11.3	0.003
Additionally excluding inter-atrial repairs (group 2)				
HES	237/2103 = 11 %	19/115 = 17 %	6.0	.12
CSR	265/2207 = 12 %	25/108 = 23 %	12.0	0.001

The full analysis has not been re-run for this particular subset of patients. A simple comparison has been made between the overall mortality rate elsewhere and that in Bristol. The ‘p-value’ is the chance of observing such a difference by chance alone, and is based on a standard ‘chi-squared test’.

Table 6.5. Impact of including all HES data for Bristol with missing outcomes, and assuming they all were survivors.

Age group	Number of missing outcomes in Bristol for open operations	Mortality elsewhere for open operations	Number of additional deaths expected if Bristol were ‘typical’	Reduction in excess number of deaths
< 90 days	7	16%	1.1	1.1
90 days – 1 year	22	7%	1.5	1.5
> 1 year	19	5%	1.0	1.0
Total	**48**		**3.6**	**3.6**

Table 7.1 Comparison of mortality rates elsewhere and in Bristol with published literature for the period 1991- 1995.

	% Mortality rates					
Procedure	Vardulaki *et al*	Hannan *et al*	HES elsewhere	CSR elsewhere	HES Bristol	CSR Bristol
Switch (G3)	8 – 12	10*	10	*12*	58	*22*
TAPVD (G4)	10 – 30	18	12	12	36	32
AVSD (G5)	8 – 12	10	8	11	35	27
Truncus (G8)	10 – 25	22	31	24	60	25
Fontan (G9)	10 – 20	14 (8**)	11	12	13	18
All surgery						
< 90 days		15	11		19	
90 days – 1 year		7	6		14	
< 1 year		11	9	9	16	13
> 1 year		3	4	5	5	7

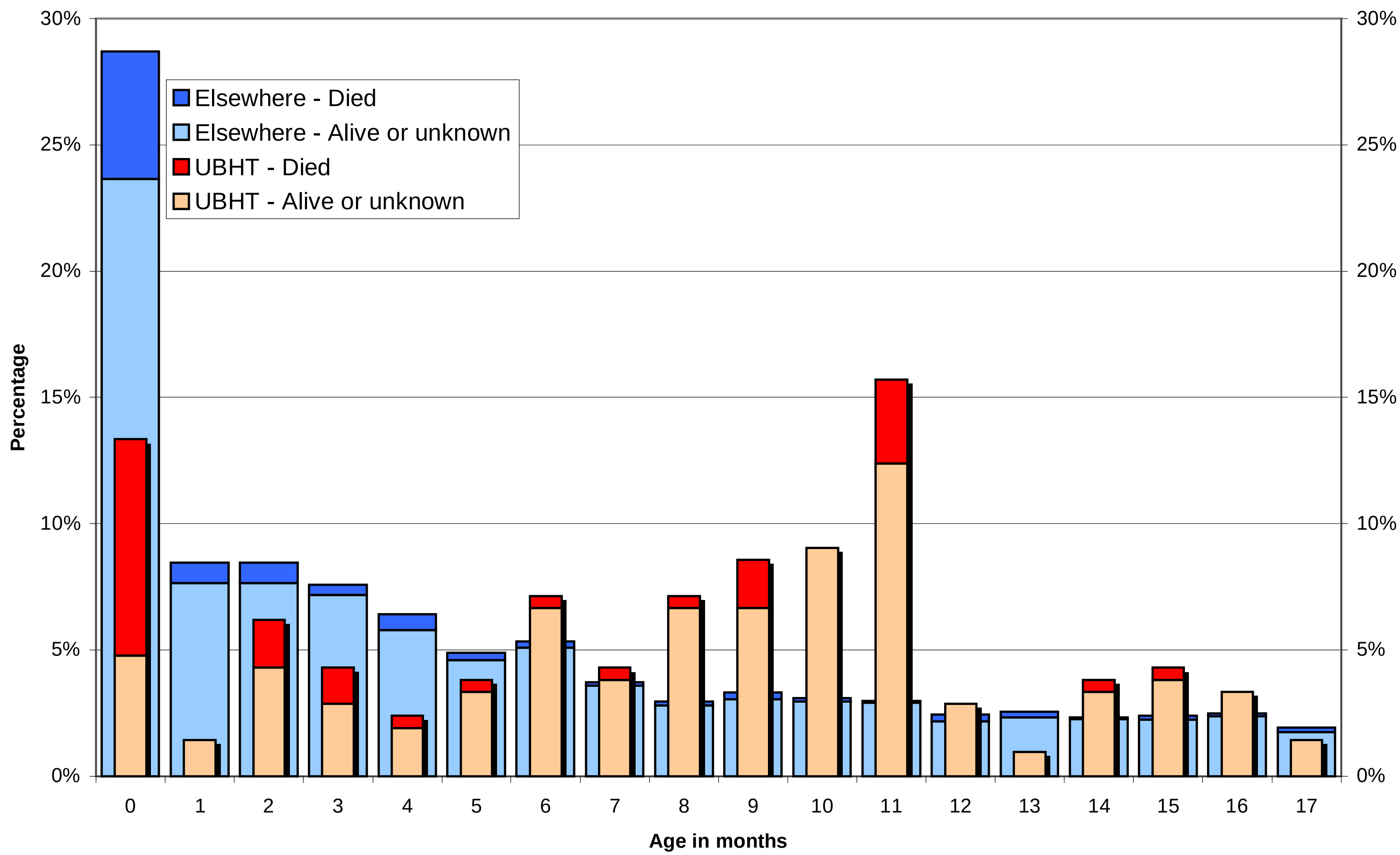
* Includes Rastelli repair / intraventricular tunnel repair
** Including bidirectional Glenn

Table 8.1 Summary table of status at admission and comorbidity for open operations, epoch 3

		UBHT			Elsewhere		
		n	%	Mortality*	n	%	Mortality*
Age	under 90 days	37	7%	63%	1,696	22%	16%
	90 days – 1 year	135	27%	19%	1,641	21%	7%
	1 - 15 years	333	66%	7%	4,408	57%	5%
Down's Syndrome	Mention of Down's syndrome in any diagnosis field	52	10%	14%	539	7%	8%
	No mention of Down's syndrome	453	90%	14%	7206	93%	7%
Transfers	Transfers from other units	32	6%	58%	1707	22%	14%
	Non-transfers	473	94%	11%	6038	78%	5%
Emergencies	Emergency admissions	36	7%	46%	783	10%	12%
	Non-emergency admissions	469	93%	11%	6962	90%	7%

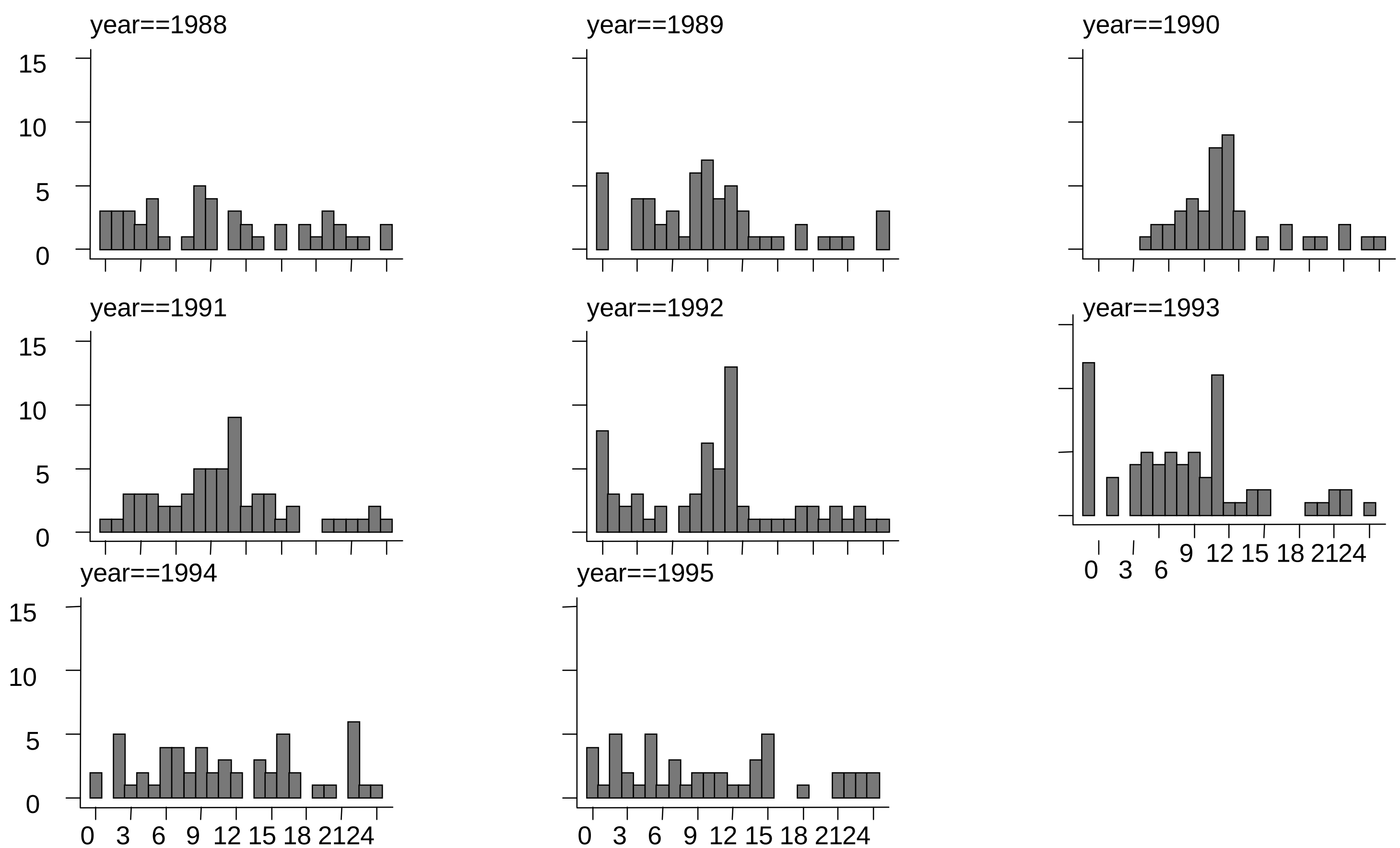
*based on spells with known outcome

Figure 8.2 Comparison of percentage of Open operations including outcome (death, alive or unknown) by age at admission (in months) between UBHT and elsewhere in England in epoch 3 (HES 1 April 1991 to 31 March 1995) aged under 18 months.



UBHT represented by narrow bars

Figure 8.3 PAS data: Age at open operations by calendar year – Bristol.



Age in months for those 2 years & under

Figure 8.4 Comparison of percentage of open operations including outcome (death, alive or unknown) by age at admission (in months) between UBHT and individual centres between April 1991 to 31 March 1994, aged under 18 months.

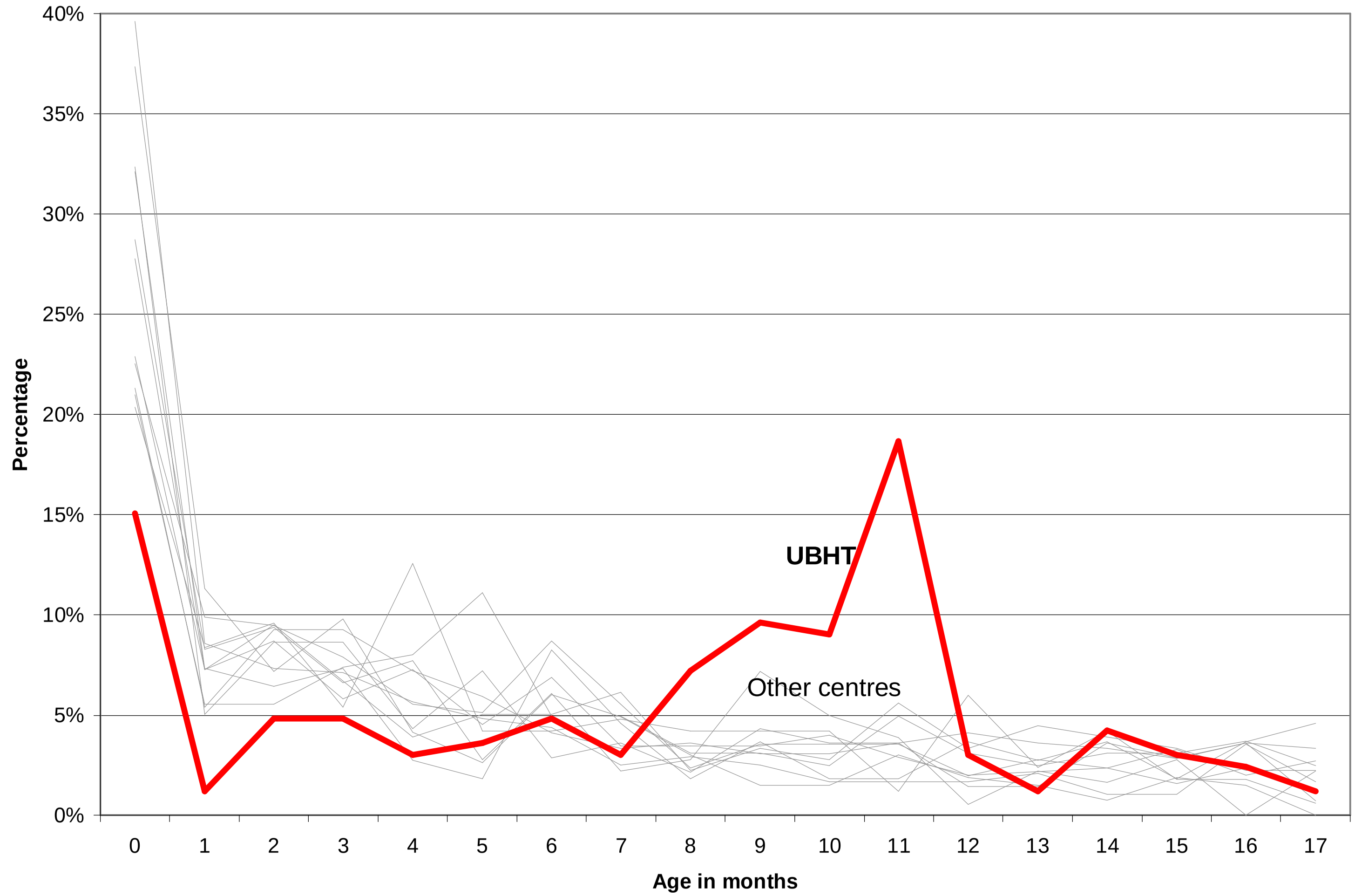


Table 8.2 Age at which open operations took place: distribution elsewhere and in Bristol, mortality elsewhere and in Bristol and relation to excess mortality. HES data Epoch 3: April 1991 – March 1995.

Age group. Mths	Percentage operations carried out by end of age-group.		Mortality rate				
	Elsewhere	Bristol	Else-where.	Bristol	Bristol deaths / num ops	Bristol expect deaths	Bristol excess deaths
0 - 2	23	9	15 %	65 %	22 / 34	5.2	16.8
3 - 5	32	13	7 %	28 %	5 / 18	1.3	3.7
6 - 8	39	21	5 %	9 %	3 / 34	1.7	1.3
9 - 11	43	35	5 %	18 %	11 / 60	3.2	7.8
12 - 14	47	38	8 %	6 %	1 / 16	1.3	-.3
15 - 17	50	42	7 %	6 %	1 / 17	1.2	-.2
18 - 20	54	45	3 %	21 %	3 / 14	.4	2.6
21 - 23	57	49	4 %	6 %	1 / 18	.7	.3
2 yrs +	65	57	4 %	10 %	4 / 40	1.6	2.4
3 yrs +	72	66	4 %	7 %	3 / 41	1.6	1.4
4 yrs +	78	73	4 %	3 %	1 / 35	1.4	-.4
5 yrs +	93	91	4 %	6 %	5 / 84	3.4	1.6
10 yrs +	100	100	4 %	4 %	2 / 46	2.0	0.0
Total							34.1