# The use of sample weights in the Clinical Case Note review

## Background

1. The Inquiry's terms of reference meant they had to "examine the care of children receiving complex cardiac surgical services at the Bristol Royal Infirmary between 1984 and 1995" and "to make findings as to the adequacy of the services provided".

2. There were over 1800 sets of medical case notes that related to children who had received the cardiac surgical services in the time period. It would not be possible for careful review of every detail of every child's set of notes to be reviewed by experts in cardiac surgical care in any reasonable timescale. It was decided to review a sample of the notes.

3. The Inquiry's approach to selecting cases for expert clinical review was based on a set of key guiding principles:

   - selected cases must be representative of all children falling within the Inquiry's terms of reference, as identified by the United Bristol Healthcare NHS Trust [UBHT] through a formal discovery process;
   - selected cases must reflect those concerns that led to the Inquiry;
   - cases must be selected in a way that is fair, transparent, rigorous, statistically valid, and feasible.

4. On the basis of these guiding principles and expert statistical advice, the Inquiry decided to select a stratified random sample of eighty cases, weighted preferentially towards children who:

   - were under one year at time of their first procedure;
   - received higher risk open heart procedures;
   - died within thirty days of their last procedure.

5. When due account is taken of the preferential weighting, it is possible to generalise from the sample of 80 cases to the full group of paediatric cardiac patients in Bristol. This is described as "re-weighting" the sample, and may be seen as reversing the process of sampling. There are obvious limitations in the precision of the extrapolation to the full group because only 80 and not over 1800 cases have been examined. In order to see how the re-weighting is done, the process of selecting the 80 cases must be reviewed in detail.

**Details of the cases selected**

6. Eighty cases were selected for review comprising forty children who had died within thirty days of their last surgical procedure and forty children who were alive at thirty days after their last surgical procedure. In both the 40 who had died and the 40 who were alive at 30 days there were six who had closed heart surgery and thirty four who had open heart surgery.

7. Table 1 provides a summary of the types of case in the group of 1827 as a whole, and the types of case in the sample. There are some notes giving further explanation as a footnote to Table1.

**Explanation of sampling weights and re-weighting**

8. If the first row of numbers in table 1 is examined, it is seen that there were 28 children in the youngest age group (<29 days) who had higher risk open heart surgery and who died. There were 97 in this category who were alive at 30 days. From each of these two categories there were 6 children chosen. Thus the sampling ratio for the first is 6/28 = 21.4%; while for the "Alive" category the ratio is 6/97 = 6.2%. A similar process is repeated for each of the other categories of risk, age and deaths/alive. The lowest sampling ratio is for the oldest (1-15 years) group who were alive and had closed heart surgery. Here 1/230 was chosen, a sampling ratio of 0.4%.

9. When it comes to re-weighting the sample of 80 in order to estimate the results for adequacy of care the process is reversed. Each child in the youngest age

group (<29 days) who had higher risk open heart surgery and who died will represent 28/6= 4.66 children. This is referred to as a *weight* of 4.66. Each child in the oldest (1-15 years) group who was alive and had closed heart surgery will have a weight of 230/1, 230. The same process is repeated for each of the 80 children in the sample. In order to be able to use the weights in analysis they are rounded to the nearest whole number.

10. In the CCNR report at Tables 3a and 3b the results of the reweighting are shown. These are reproduced below. In order to see how the re-weighting works it is easiest to examine the category of adequacy 2/3 since it has the fewest children. From the third row of table 3a, with adequacy 2/3 there are four children in the sample. In fact they have weights 4,4,5 and 7. We then take these children to represent 4+4+5+7 =20 children in the full group of over 1800 children. A similar process is done for each of the children in the other categories of adequacy. It is clear that the process is both slightly inaccurate because of the rounding of the weights in order to have whole numbers in all tables, but also the numbers of sets of case notes that have been studied in detail is only 80. Statistical methods are able to allow for this in estimating the uncertainty in the derived percentages for the whole group. These have not been used extensively.

11. Because the stratified sample of 80 was chosen randomly with known sampling ratios, it is possible to draw some wider conclusions about the adequacy of care given to all 1827 children. But any such conclusions must take full account of the weighting in the sample. As those at highest risk and those who died were included preferentially, then it is important to realise that it leads to distortion if this is not taken into account. The reweighting of the sample is designed to take this into account.

Professor Stephen JW Evans
October 2000

## Table 1 BRISTOL ROYAL INFIRMARY INQUIRY:

## Children Receiving Heart Surgery at BRI and BCH, 1984 to 1995

| | 30 day mortality - CCR database | | | | | Samples and sampling ratios | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | Deaths | "Alive" | % deaths | | Sample deaths | Sampling ratio | Sample "Alive" | Sampling ratio |
| Age groups | | | | | | | | | |
| Higher risk Open | | | | | | | | | |
| <29 days | 125 | 28 | 97 | 22% | | 6 | 21.4% | 6 | 6.2% |
| 29 days – 1yr | 146 | 43 | 103 | 29% | | 6 | 14.0% | 6 | 5.8% |
| 1-15 yrs | 104 | 21 | 83 | 20% | | 3 | 14.3% | 3 | 3.6% |
| Total | 375 | 92 | 283 | 25% | | 15 | 16.3% | 15 | 5.3% |
| | | | | | | | | | |
| Other open | | | | | | | | | |
| <29 days | 99 | 34 | 65 | 34% | | 9 | 26.5% | 9 | 13.8% |
| 29 days – 1yr | 281 | 42 | 239 | 15% | | 6 | 14.3% | 6 | 2.5% |
| 1- 15yrs | 535 | 28 | 507 | 5% | | 4 | 14.3% | 4 | 0.8% |
| Total | 915 | 104 | 811 | 11% | | 19 | 18.3% | 19 | 2.3% |
| | | | | | | | | | |
| Closed heart surgery | | | | | | | | | |
| <29 days | 128 | 36 | 92 | 28% | | 4 | 11.1% | 4 | 4.3% |
| 29 days – 1yr | 176 | 8 | 168 | 5% | | 1 | 12.5% | 1 | 0.6% |
| 1-15 yrs | 233 | 3 | 230 | 1% | | 1 | 33.3% | 1 | 0.4% |
| Total | 537 | 47 | 490 | 9% | | 6 | 12.8% | 6 | 1.2% |
| | | | | | | | | | |
| | | | | | Sample numbers | | | | |
| Totals | | | | | | Deaths | | "Alive" | |
| Open | 1290 | 196 | 1094 | 15% | Open | 34 | | 34 | |
| Closed | 537 | 47 | 490 | 9% | Closed | 6 | | 6 | |
| Total cases | 1827 | 243 | 1584 | | Total | 40 | | 40 | |
| | | | | | | | | | |
| Sample cases available = | | 80 | | | Grand Total | 80 | | | |

**Notes on the sampling method**

- classification of coded procedures as higher risk open, other open, or closed was based on clinical advice to the Inquiry, and inevitably reflects an element of clinical judgement;
- For sampling purposes, and to give emphasis to the whole child rather than to individual procedures, the age of the child was set as the age at first procedure (and not necessarily as the age at defining operation). One consequence of this is an apparent – but not real – exaggeration of the numbers of neonates in the sampling frame;
- Deaths are defined as deaths occurring within 30 days of the last operation received by the child;
- The group of children who were alive 30 days after their last surgical procedure includes three children who died much later. For the purposes of this report, these children continue to be considered in the category "alive 30 days after last surgical procedure";
- The target sample size (eighty cases) is not statistically determined, but reflects the maximum number of cases that the Inquiry - in the light of clinical advice - considered feasible to submit to in-depth expert clinical review within the short timescale available.

**Distribution of overall grades for adequacy of care, (a) for the sample (b) extrapolated to all patients known to the Inquiry to have received paediatric heart surgery between 1984 and 1995**

**Table 2a. for the sample of 80 cases**

| Adequacy (Overall Grade for Adequacy) | Frequency (No. of children in sample assigned this grade) | Percent (% of children assigned this grade) | Cumulative % |
|---|---|---|---|
| 1 | 13 | 16 | 16 |
| 2 | 11 | 14 | 30 |
| 2/3 | 4 | 5 | 35 |
| 3 | 13 | 16 | 51 |
| 4 | 39 | 49 | 100 |
| Total | 80 | 100 | |

**Table 2b. Extrapolated to the Bristol paediatric heart patients as a whole**

| Adequacy (Overall Grade for Adequacy) | Frequency | Percent | Cumulative % |
|---|---|---|---|
| 1 | 101 | 5.5 | 5.5 |
| 2 | 69 | 4 | 9.5 |
| 2/3 | 20 | 1 | 10.5 |
| 3 | 345 | 19 | 29.5 |
| 4 | 1294 | 70.5 | 100 |
| Total | 1829 | 100 | |

Note to Table 2b: The numbers in the column entitled "frequency" should not be taken as precise values, but as estimates. They have been calculated by re-weighting each of the eighty cases individually according to the original probability of that case being selected. The final report on the Clinical Case Note Review [Annex C] makes it clear how this was done. This process can result in an estimate not being a whole number. For presentation, the numbers are rounded to the nearest whole number, hence the total of 1829 is an approximation to the original total of 1827.