

TECHNICAL NOTES

Combining data from four sources

The statistics presented in this publication are based on four returns made by local councils in England:

- form CLA100, on which councils provide a range of aggregate figures on their looked after children; (introduced in 1997/8)
- form SSDA 903, on which they provide detailed data on a 'one-third' sample of their looked after children, covering children born on a day of the month divisible by 3. This sampling approach was introduced from 1 April 1997;
- form OC1, on which authorities provide detailed data on the educational qualifications of all young people aged 16 or over who ceased to be looked after; (introduced for 1999/2000)
- form AD1, on which councils provide information on all children adopted from care during the year (introduced 2000/01)

Figures from the CLA100, AD1 and OC1 collections for the year ending 31 March 2001 were summarised in Department of Health Statistical Bulletins published in October 2001¹. Since then additional information has been received and as a result some figures have been revised in this publication.

Of the four sources of data used in this publication (see above), it is generally felt that the SSDA 903 is the more accurate. There are three reasons for this:

- it is a straightforward survey of all episodes of care and there is no bias in the collection towards any particular child or data item
- all SSDA 903 records are subjected to a rigorous validation to ensure that all reported episodes are fully consistent one with another. Failure to pass any validation test renders the entire care-history of the child invalid, and these records are excluded from all analyses.
- efforts are made to ensure that for every child who is looked after at 31 March, a further episode for that child is submitted in the following year. No child can simply go "missing" from the database without this being reported back to the appropriate council.

The principal disadvantage of the SSDA 903, being a one-third sample survey, is that it is unable to produce reliable estimates when sample numbers are small.

Estimation methodology

Figures presented here for years before 1997/98 are based entirely on the detailed SSDA 903, which for those years covered all looked after children. For 1997/98 onwards data from the two collections, SSDA 903 and CLA100, have been combined to provide the best estimates available of the figures at a national level and at a local council level.

For the years 1997/98 and 1998/99, the methodology in the main has been to use a range of multiplying factors which are calculated by comparing CLA 100 totals for the numbers of children who

- started to be looked after during the year
- ceased to be looked after during the year
- looked after at 31 March

with corresponding counts of valid SSDA 903 records. For the year 2000/01, the count of the number of children starting to be looked after was dropped from the CLA100, and for 2001/02, the count of the number ceasing to be looked after during the year will also be dropped. These changes have meant a revision to the current grossing methodology, and it was decided that for the years 1999/2000 and 2000/01, only one multiplying factor based on the number of children looked after at 31 March should be prepared for each local council. This change, which would in any event need to be introduced with effect from 2001/02, was brought forward to 1999/2000 following a close examination of the CLA100 and SSDA 903 data, which were found to deviate one from the other in a number of key areas in recent years. (See section on Data Quality for explanations of these differences). By using just one factor, the results obtained will follow the pattern as revealed by the SSDA 903 sample, rather than the CLA100. For the reasons stated above, the SSDA 903 is considered to be the more accurate.

The result of this change is that the numbers of children both starting and leaving during the year 1999/2000 as shown in this publication differs from the results published last year, and is, in some

¹ www.doh.gov.uk/public/stats3.htm

respects, at odds with the CLA 100 aggregate data as reported on last year's returns.

Additionally, OC1 and AD1 data has been used to prepare tables H, Xi, Xii, 16, 17 and 18. As the coverage of the OC1 is 100%, the age and sex breakdown of care leavers as shown on this return differs slightly from the results obtained by grossing sample numbers from the SSDA903.

In some tables that present data on children being looked after at 31 March, figures from 1997/98 are based entirely on the aggregate data from the CLA100 return. These include Tables A - C, F, 3 and 5 - 7.

Figures from the two sources, SSDA 903 and CLA100, are presented on a consistent basis wherever possible, so that the sum of figures from the sample equals the total figure from the CLA100. However in some cases this has not been possible – and footnotes have been added to explain any inconsistency. The act of sampling invariably reduces the reliability of some of the breakdowns presented. This should be borne in mind when interpreting figures for individual authorities that are derived from the SSDA903, particularly where the figures for individual categories are relatively small.

Missing data for 2000/01

SSDA903 'one-third' sample

South Tyneside was the only local council who were unable to submit a SSDA 903 return, although a CLA100 return was received. Hence, data have been shown as being "not available" for this council on all tables for which SSDA 903 input is required. The SSDA 903 data received from Kent did not include any children who were Unaccompanied Asylum Seekers; and the submission from Hillingdon was never fully validated.

OC1 and AD1 returns

All councils in England managed to submit both an OC1 and an AD1 submission in time for this publication.

England estimates

All figures for England include estimates for missing data.

Data Quality

The Department goes to great lengths to ensure that the quality of its data relating to looked after

children are as accurate as can be achieved within a reasonable time scale with the resources available.

Both the SSDA 903 and the CLA 100 are subjected to validation checks, as well as being compared one with the other. Inconsistencies between the two returns are raised with the authorities for checking and correction.

The introduction of the AD1 this year (and the OC1 last year) has both required extra validation checks to be carried out across all four returns. This highlighted many inconsistencies that were addressed, and this led to a number of changes being made to local councils' SSDA 903 submissions for 1999/2000. The nature of many of these changes and alterations were such as to indicate that a number of corresponding changes were also required to the 1999/2000 CLA100 returns. Unfortunately, with the exception of two local councils, no changes were submitted. It is felt that much of the divergence between the two returns is rooted in this problem, and it was this that led to the Department deciding to change its estimation methodology (see above).

Table A2 at Appendix A gives a comparison of the CLA 100 and SSDA 903 databases as at the time they were "frozen" in order to prepare this publication. In this table ratios are given for each local council in respect of the number of children looked after at 31 March for the past four years. Taking into account leap years, the expected value of all ratios is 3.07, although exactly this figure is unlikely to be achieved very often because of the chance effects of sampling. Ratios which indicate a significant difference between the CLA 100 and the SSDA 903 submission have been underscored, and any estimate in this publication that depends on an underscored ratio should be treated with caution.

It will be noticed that the number of councils with underscored ratios have remained fairly constant at around 25 for the past three years.

Various validation checks were also carried out between the OC1 and SSDA 903 returns, to ensure that the child details were consistent. Checks were also done to help distinguish records where an OC1 record was submitted for a child who should have fallen in the one-third sample, but no SSDA 903 record was received and vice versa.

Powers of Criminal Courts (Sentencing) Act 2000

For reasons of simplicity, for the period 1997/98 to 1999/2000 local councils were asked to include in their CLA100 figures children under Sections 90 -

92 of the Powers of the Criminal Courts Act 2000, (formally Section 53 of the Criminal & Young Persons Act 1933) even though these are generally not looked after. This requirement was however dropped for 2000/01. This means that estimates for the number of children looked after at 31 March 1997/98 to 1999/2000 are inflated by about 100 children compared with figures for other years.

Local Government Reorganisation

Local councils affected by reorganisation at the end of March 1998 supplied the Department with information indicating which successor council took over responsibility for which looked after child. This information covered all children looked after by these authorities since the implementation of the Children Act. It has therefore enabled us to present estimated data in this publication for the new authorities for years before they came into existence.

Rounding and Confidentiality

To ensure that no individual child can be identified from statistical tables, we have introduced new conventions for the rounding and suppression of very small numbers. All figures which appear in national tables (Tables A to Y) have been rounded to the nearest hundred if they exceed 1,000 and to the nearest ten otherwise; and (for the first time), all figures at local council level have been rounded to the nearest five. In addition, all figures at local council level of five or less have been suppressed, and all information for three very small authorities, City of London, Isles of Scilly and Rutland, has been suppressed totally.

It has on occasion been necessary to suppress other data whenever it would be possible to calculate missing data by means of simple arithmetic. The rule applied in these circumstances has been to suppress the next smallest data item. As a result of all these changes, detailed figures may not sum to the totals presented.

For further details of roundings and suppression, see Appendix D.

Rates and percentages

Rates per 10,000 population have been calculated using the appropriate mid-year estimates of the number of children in England aged under 18, supplied by the Office for National Statistics. A small number of young people aged 18 and over are included in the figures, but this is not thought to seriously distort the rate. The mid-year population estimates for 1998 to 2000 are presented in Appendix A. These have been used to derive estimates of rates for the years 1998/99 to 2000/01 respectively.

Percentages are presented in the tables in italics. They are shown as whole numbers, and as a result may not sum to 100 due to rounding. All rates and percentages, where either the denominator, or the numerator, is less than five, have been suppressed for reasons of client confidentiality.

Figures on the Internet

The Department of Health's web site contains a copy of this publication, with all the tables in a suitable format for downloading into Microsoft Excel. It can be found under the heading Social Care at:

<http://www.doh.gov.uk/public/stats3.htm>

Symbols used in the tables

The following conventions have been used in the tables:

- .. not available
- data suppressed for reasons of confidentiality. Normally this will refer to numbers of five or less, or percentages where either the numerator or the denominator is less than five