

---

**ESRC  
End of Award Report  
Full Report**

**The Millennium Cohort Survey:  
First Survey**

---

## Contents

<b>1.</b>	<b>Background</b>	<b>1</b>
	1.1 ESRC funding of the Principal Investigator	2
	1.2 Fieldwork sub-contract	3
<b>2.</b>	<b>Objectives</b>	<b>4</b>
	2.1 Achievement and modification of objectives	4
	2.2 Further objectives	5
<b>3.</b>	<b>Methods</b>	<b>7</b>
	3.1 Sample design	7
	3.2 Sampling procedures	9
	3.3 Fieldwork cycle	10
	3.4 Translation	10
<b>4.</b>	<b>Results</b>	<b>11</b>
<b>5.</b>	<b>Activities</b>	<b>14</b>
<b>6.</b>	<b>Outputs</b>	<b>16</b>
<b>7.</b>	<b>Impacts</b>	<b>17</b>
<b>8.</b>	<b>Future research priorities</b>	<b>18</b>
	<b>Ethical Issues</b>	<b>19</b>
	MREC	19
	Codes of practice	19
	Geo-coded data	19

## 1. Background

The Millennium Cohort Study (MCS) is the fourth of Britain's world-renowned national longitudinal birth cohort studies. Each follows a large sample of individuals born over a limited period of time through the course of their lives, charting the effects of events and circumstances in early life on outcomes and achievements later on. The established studies show how histories of health, wealth, education, family and employment are interwoven for individuals and vary between them. The data collected are used for many scientific and policy purposes. The Millennium Cohort Study is intended to continue the series with lives begun in the twenty-first Century.

The MCS offers large scale information about the New Century's babies and their families for the four countries of the United Kingdom. The First Survey, carried out during 2001-2002 has laid the foundations for a major new longitudinal research resource, taking a new "year long" cohort of around 19,000 babies. In England and Wales they were born over the twelve months, starting in September 2000 in England and Wales, and over 13 ½ months from late November 2000 in Scotland and Northern Ireland. Information for the first survey has been collected from parents when the babies were aged nine months. The sample design allowed for disproportionate representation of families living in areas of child poverty, in the smaller countries of the UK and in areas with high ethnic minority populations in England. The first survey recorded the circumstances of pregnancy and birth, as well as those of the all-important early months of life, and the social and economic background of the family into which the children have been born. This multidisciplinary baseline data will reveal the diversity of starting points from which these 'Children of the New Century' are setting out.

The three original birth cohort studies followed cohorts born in Great Britain in one week each of 1946, 1958 and 1970. For a comparison of some of their findings see Ferri et al, (2003)<sup>1</sup>. During the 1980s and 1990s funding for following up the existing cohorts was precarious and none was forthcoming for starting a new study in the series. However, in 1998, the Economic and Social Research Council (ESRC) established a National Strategy for Longitudinal Data Resources which put the future funding of the 1958 and 1970 cohorts on a firmer footing, alongside the British Household Panel Study. In 1999 the British government decided to mark the passing of the Millennium with a new birth cohort study. The ESRC commissioned a feasibility study from Professor Jonathan Gershuny who held wide consultations with the research community. This resulted in an invitation to tender for the Principal Investigator role in a new Study and to submit expressions of interest to carry out its fieldwork. These

---

<sup>1</sup> Ferri, E. Bynner, J and Wadsworth, W. (eds) (2003) *Changing Britain, Changing Lives*. London: Institute of Education

invitations were issued when the new Millennium was already two months old on February 23 2000.

The original ESRC specification was for a study which in some ways resembled that of the previous cohorts, but which departed from the previous designs in the following respects:

- The sample was to include children born over a full twelve months rather than one week, to include births in all seasons
- The sampled birth dates should include the birth dates in the year 2000, even though it might not be possible to start with births in January
- The first survey was to be carried out when the children were around the same age in months, as near as possible to 6 months old
- The contractors were to consider sample designs allowing for the over-representation of ethnic and national minorities

The geographical coverage of the study was to include the whole of the UK. Its content should emphasize the social and economic circumstances of the families. The target sample size was to be 15000.

A further contrast with the earlier surveys was that this initial data collection was to be carried out by a professional fieldwork agency, rather than by health visitors. Constraints on the available numbers of professional interviewers and the consequent need to spread interviewing over time were also considerations favouring a wider span of birth dates than a single week.

## **1.1 ESRC funding of the Principal Investigator**

The Principal Investigator contract for the First Survey of the Millennium Cohort Study (MCS1) was awarded in May 2000 to a consortium based in three institutions:

1. Institute of Education (IOE): Centre for Longitudinal Studies (CLS) (John Bynner);
2. University College London: Department of Community Epidemiology and Public Health, (Michael Marmot), and Institute of Child Health, (Catherine Peckham);
3. City University: Department of Psychology, (Dermot Bowler).

The scientific leader of the project was Heather Joshi, based at CLS in the IoE, along with the internal management team. She was confirmed as Director of the Millennium Cohort Study in October 2001 when the Principal Investigator contract for Sweep 2 was awarded to the CLS-led consortium through the Joint Centre for Longitudinal Research. At that time the study was taken into the umbrella of the portfolio of longitudinal data resources overseen by the ESRC's National Strategy Committee for Longitudinal Data. The management structure

for the MCS1 is given in Appendix 1 of Nominated Publication 1. This also shows the full complement of research staff at CLS including those supported by HEFCE and ONS funding. ESRC supported staff are identified on the End of Award form.

## **1.2 Fieldwork sub-contract**

The National Centre for Social Research (NatCen) was appointed to carry out the fieldwork at the end of September 2000 after the first cohort children had already been born.

## 2. Objectives

The objectives of the first MCS survey were laid down in the CLS proposal to the ESRC in March 2000:

1. To chart the initial conditions of social, economic and health advantages and disadvantages facing new children in the new century, capturing information that the research community of the future will require.
2. To provide a basis for comparing patterns of development with the preceding cohorts.
3. To collect information on previously neglected topics, such as the fathers' involvement in the children's care and development.
4. To focus on the children's parents as the most immediate elements of the child's 'background', charting their experience as mothers and fathers of this year's babies, to record how they (and any other children in the family) are adapting to the newcomer, and what their aspirations for her/his future may be.
5. To emphasise intergenerational links including those back to the parents' own childhood.
6. To investigate the wider social ecology of the family, including social networks, civic engagement and community facilities and services, splicing in geo-coded data when available

### 2.1 Achievement and modification of objectives

1. *A multipurpose-public use data set*  
The first and second objectives have been achieved within the constraints of what is feasible to collect in a multi-purpose survey without excessively burdening respondents. A public use dataset has been deposited at the UK Data Archive.
2. *The establishment of birth cohort comparisons*  
Studies in other countries: Australia, US, Denmark, Norway, Ireland, Canada, India, Peru, Vietnam and Ethiopia etc, created the possibility of international as well as inter cohort comparison
3. *Fathers*  
The topic of fathers' involvement has been addressed in interviews with 72% of the families who responded, and has also been the subject of questions to mothers where fathers were absent.

4. *Range of topics*  
Of all the topics it was originally planned to include on parents' experiences, some had to be dropped for want of space in the questionnaire, notably their aspirations for the child's future and details of their own occupational history. It was felt that it would be possible, and in the case of aspirations, more appropriate to ask these questions at a later sweep.
5. *Intergenerational links*  
There are some questions about grandparents' current interaction with the cohort family and about the parent's own childhood. The question on grandparents social class was held over until sweep 2.
6. *Locality*  
The wider social ecology of the sampled locations is being investigated by splicing in data from ONS Neighbourhood Statistics and from a survey of Health Visitors in these localities. These enhancements are both funded by the ONS consortium. Data about localities has not been deposited in the Data Archive as yet, until safeguards of respondents' anonymity can be ensured, but the ecological material can be, in principle, integrated for analysis in the safe setting of CLS.

## **2.2 Further objectives**

Further objectives were added with the co-funding provided by government departments.

1. *Controls for Sure Start*  
The MCS1 dataset has been handed over to the National Evaluation for Sure Start (NESS), who will have to determine which MCS locations or individuals came closest to the notion of being free of "treatment". The instruments used in NESS and MCS1 (and MCS2) have been developed in parallel, although they are not duplicates because of the wider interests of the MCS survey.
2. *Sample sizes in Celtic countries large enough for in-country analysis*  
The original design allowed for 1,500 in each of the three Celtic countries. After booster funds contributed by the devolved administrations, the target sample sizes rose to 3000 in Wales, 2500 in Scotland and 2000 in Northern Ireland. The samples eventually achieved were 2,761, 2,300 and 1,955 children respectively, which means these countries are all over-represented in the unweighted sample. It is particularly pleasing to have made a good start in Northern Ireland which is not covered by the earlier 'National' birth cohort studies.

3. *Enhancement of health data*

A postal survey is following up some of the 450 mothers whose cohort child was born through Assisted Reproduction techniques is being undertaken. The routine hospital statistics about the cohort child delivery are being added to the dataset. Both of these exercises are funded by the Department of Health through ONS and are being carried out in collaboration with the National Perinatal Epidemiology Unit in Oxford.

4. *Timescale*

The original timescale was clearly not feasible. The target age for the interviews was set at 9 months (and the first birthday in the cohort at September 2000). Fieldwork started in June 2001.



### 3. Methods

#### 3.1 Sample design

It was decided after the award of the contract to CLS, that the sample of a year's births should be tightly clustered geographically, disproportionately stratified to over-represent areas with high proportions of: ethnic minorities in England, areas of high child poverty and the three smaller countries of the UK respectively. The geography of electoral wards provided the sampling frame as the Child Poverty component of the Oxford Index of Deprivation 2000 (Noble et al, 2000)<sup>2</sup> provided as up to date as possible indicators of localities with high proportions of children in families receiving means tested benefits.

The sample was drawn from babies born between 1 September 2000 and 31 August 2001 in England and Wales. In Scotland and Northern Ireland the start date of the birthdays was delayed to 23 November, 2000 in order to avoid an overlap with an infant feeding survey being carried out in September and October of this year. In these countries the sampled cohort was extended to 59 weeks of births to make up for a shortfall in numbers which became apparent during fieldwork. The last eligible birth date in these countries was 11 January 2002.

Children with sample birth dates were eligible for the survey if they were living in one of approximately 400 electoral wards across the whole of the UK when they were 9 months old.

The breakdown of the target sample for Sweep 1 is outlined in Table 1 below.

---

<sup>2</sup> Noble, M. Smith, G. Penhale, B. Wright, G. Dibben, C. Lloyd, M. (2000), *Measuring multiple deprivation at the small area level: The indices of deprivation*. Final report for the DETR.

**Table 1: Millennium Cohort sample structure by stratum and country:  
wards and babies aged 9 months (numbers of children in italics)**

<b>Country</b>	<b>Total wards target no's of children</b>	<b>(of which, boost from original number)</b>	<b>Advantaged wards expected sample</b>	<b>Disadvantaged wards expected sample</b>	<b>Ethnic wards expected sample</b>
<b>England</b>	<b>200</b> <b>13,146</b>	(2,646 disadvantaged babies in 35 wards for Sure Start)	110 5,511	71 5,258	19 2,606
<b>Wales</b>	<b>73</b> <b>3,000</b>	(1,500 babies in disadvantaged wards)	23 897	50 2,219	Na
<b>Scotland</b>	<b>62</b> <b>2,500</b>	(1,000 babies)	32 1,243	30 1,285	Na
<b>Northern Ireland</b>	<b>63</b> <b>2,000</b>	(500 babies in disadvantaged wards)	23 762	40 1,322	Na
<b>Total</b>	<b>398</b> <b>20,646</b>	<b>5,700</b>	<b>188</b> <b>8,413</b>	<b>191</b> <b>10,084</b>	<b>19</b> <b>2,606</b>

('wards' counts amalgamations of small wards ('superwards') as units. See Technical Report on Sampling)

## **3.2 Sampling procedures**

### **3.2.1 The use of Child Benefit Records**

We had originally assumed that birth registration records would be used as a sampling frame, but there was concern about non-response to the postal opt-in operated by ONS. The representation of socially excluded groups in the sample could be compromised, especially those with poor literacy or poor grasp of English. By contrast the use of Child Benefit Records, made newly accessible by DSS/DWP involves asking people to opt-out, which is known to be more inclusive of marginal and low literacy cases. The arrangement to use the DWP system was confirmed in September 2000

There was concern about DWP's withdrawal of 'sensitive cases' from the issued sample. But in the event these exclusions were kept to a very modest level (see Nominated Publication 2). The sampling frame also omitted families not entitled to Child Benefit such as recent international in-migrants.

### **3.2.2 Recruitment via Health Visitors**

Because the Child Benefit Records did not reveal all families who had moved into the sample wards as the child approached 9 months of age, we attempted to find movers-in with the help of Health Visitors. These local community health professionals were expected to be aware of families transferring into their areas in the course of their duties. They were asked to see if eligible families moving into survey wards were willing to be recruited. They were also asked to inform and reassure other families who might have received the opt-out invitation about the survey.

It is difficult to quantify how important this reassurance may have been, but we do know that health visitors reported 220 in-moving families with children over 6 months of age. However only 56 had not also been found by DWP. The possible explanations for this rather disappointing result are set out in nominated publication 2. We also used our Health Visitor contacts to complete a postal survey on local services in the sample areas.

### **3.3 Fieldwork cycle**

The fieldwork was carried out in 17 consecutive waves. Each wave comprised babies born in a 4-weekly birth cycle. This rhythm of recruiting the sample was dictated by the four weekly cycle of DWP scans of the Child Benefit database, done when the cohort child was around 7 months old. DWP then contacted the selected families asking if they wanted to opt out, before releasing their names and addresses. Interviewers were to reach the families while the baby was as close as possible to 9.5 months of age. Interviews with partners could be delayed until the child's first birthday.

Waves 1-13 of fieldwork took place in England and Wales from June 2001 to August 2002. The last wave in Scotland and Northern Ireland, wave 17, was the extended sample spanning 7 weeks of births. It finished before the end of 2002 in Scotland and on 11 January 2003 in Northern Ireland

75% of main interviews took place while the baby was aged 9 months and most of the remainder at 10 months.

### **3.4 Translation**

On the recommendation of the Multi-Centre Research Ethics Committee (MREC), a simplified leaflet was produced for interviewers to give to respondent families on the doorstep. This leaflet, the advance letter and the thank-you letter were translated into the most common non-English languages spoken in the 19 selected 'ethnic' wards: Bengali, Gujarati, Kurdish, Punjabi, Somali, Turkish and Urdu. Some interviews were carried out in verbal translation (in these and other languages) by relatives or friends. In certain circumstances where no one was available to translate into English, NatCen provided translator interviewers. Other languages encountered in non-trivial numbers included Arabic, Hindi and Tamil. 226 (1%) main interviews were carried out in a non-English language and a further 547 (3%) were done in a mix of English and another language. For partners the corresponding figures were 306 (2%) and 94 (1%).

## 4. Results

Through face-to-face interviews the study has found out about the start in life of nearly 19,000 babies born during 2000 and 2001 in all four countries of the UK. To details of their birth we add early health, household structure, the socio-economic circumstances of the parents, neighbourhoods and so on.

**Table 2: The achieved responses for Sweep 1 of the Millennium Cohort Survey**

	Children	Families	Partners	Single Parents
<b>UK</b>	18,819	18,553	13,599	3,194
<b>ENGLAND</b>	11,695	11,533	8,558	1,853
<b>WALES</b>	2,799	2,761	1,957	590
<b>SCOTLAND</b>	2,370	2,336	1,758	375
<b>N IRELAND</b>	1,955	1,923	1,326	376
<b>ADVANTAGED</b>	7,429	7,317	6,139	646
<b>DISADVANTAGED</b>	8,974	8,842	5,935	2,064
<b>ETHNIC</b>	2,416	2,394	1,525	484
	Includes twins and triplets	Almost all mothers	Almost all fathers	Almost all mothers

The achieved sample shown in Table 2 above represents an overall response rate of 68%. This is good for a sample heavily weighted towards 'difficult areas'. It is a little, but not much, below the assumptions made on little evidence on drawing up the original design. Potential sample numbers also shrank because fertility fell to an all-time low during the period the cohort members were born. Details of the responses at the two stages are given in the nominated publications. The analysis of potential response bias is still under way.

There has not yet been time for us, or the user community for whom we have produced this data set, to draw out detailed scientific or policy analysis.

There are 3 topics to which the survey design paid special attention, on which preliminary findings can be illustrated:

Firstly, the survey is an innovation in that it notes the fathers' perspective on parenthood as well as the mothers'. There are questions for the partner as well as the main respondent. Among the 13,359 two-parent families 99.7% of the fathers were natural fathers. Most of the children of absent fathers had some degree of contact with them. Of those fathers living with their children, 57% reported changing nappies at least once a day. Over 10% of resident fathers reported at least one child living elsewhere, and a slightly greater proportion of cohort children had half-siblings living in the same household. Allowing for sample weighting, over half of the cohort children had some kind of older sibling living with them. 14% of the families were lone parent families, in 3% there was at least one grandparent living with a single parent. These data can now be used to study the changing British family itself in depth, and to control for circumstances when looking into other questions such as health, housing, employment, childcare, etc, and to compare the consistency of responses from mothers and fathers.

Secondly, the survey design also sought to over-represent ethnic minorities by over-sampling areas expected to contain a high proportion of such communities. The result of this strategy was successful in that as many as 86% of the babies in the 19 'ethnic' wards were non-white, whereas our national estimate is 13% of babies who are not white. It remains to be seen whether the characteristics of members of different ethnic groups vary by whether or not they live in an area where minorities concentrate, and to explore the diversity across groups. Pakistanis are the largest group in this sample. It can already be noted that more than one in four of the 'non-white' children nationally were reported as 'mixed' and that the data set should provide a valuable resource for the study of ethnic diversity. 69% of the families in the ethnic wards spoke a language other than English at home (10% nationally). This presented a challenge to fieldwork which NatCen and the informants took in their stride. Although, as expected, it was more difficult to find respondents in these areas, some of those who did participate gave us very long interviews to cope with language difficulties.

A third feature of the sample was that, to ensure adequate coverage of poor families, and also to collect information about neighbourhoods in which they live, electoral wards were over-sampled in which the proportion of families claiming certain means tested benefits was high. This evidence came from 1998 and 1999 and included two in-work benefits then on offer: Family Credit, Disability Working Allowance along with two means test benefits: Income Support and Job Seekers Allowance. Taking as a preliminary, and not quite comparable indicator the proportion of respondents receiving either of the last two benefits, Working Families Tax Credit or Disabled Persons Tax Credit, we found that 53% of the families in the disadvantaged wards received these benefits, compared to 59% in the ethnic areas, and 24% in the rest. These families are not as concentrated in particular localities as the non-whites. Our estimate is that 41% of the benefit recipients nationally live in the areas labelled 'advantaged'. Further work, and further data collections, will help answer whether the consequences of poverty differ by where a family lives.

Differences between 'advantaged' and 'disadvantaged' strata are generally more marked for economic and employment variables than on information relating to health. For example, lone parents are more common in disadvantaged areas, where planned pregnancy and breastfeeding are less common. Access to health services does not seem to differ greatly between area types, as confirmed by our survey of Health Visitors.

## 5. Activities

The first task of the team was to set the main parameters of the survey design in order to issue an invitation to tender for the fieldwork, which was announced in July 2000. Three bids were submitted and the contract awarded in late September 2000 to the National Centre for Social Research (NatCen), subcontracting to Northern Ireland Statistics and Research Agency (NISRA) in Northern Ireland. A consultative conference was held in October 2000 during the process of questionnaire development. Potential users from the scientific and policy community attended

The first (paper) pilot survey in January 2001 gave an indication of the cuts that would be necessary to bring the survey down even to the 105 minutes for two-parent families, extended from 90 minutes by government funding. This pilot was debriefed on the 23 January 2001 and collaborators and government departments attended a meeting to discuss these cuts on 24 January 2001.

While the instrument was being converted into CAPI and CASI by NatCen, CLS were applying for ethical clearance from the MREC in February 2001. The Southwest MREC considered it in March. They were unhappy about recruiting via opt-out, but eventually agreed that we could proceed, subject to strengthening our efforts to obtain informed consent on the first visit, sending an advance letter, re-writing our information sheet in simpler language, and having the material available and translations. The ESRC provided an additional £55,000 towards the cost of these measures.

Negotiations with the MREC coincided with the Dress Rehearsal Pilot Study which took place in April 2001. This was followed by a series of briefings and the main fieldwork started in England and Wales in mid June 2001.

It became apparent during the April 2001 Dress Rehearsal that the plan to contact health visitors to help find mobile families missed from the Child Benefit Records was not going to be straightforward. Attempts to locate a health visitor responsible for families in each of the survey wards continued well into 2002. (see 1.3.2 of Nominated Publication 1)

During fieldwork, response rates at DWP and in the field were closely monitored. DWP increased their efforts to detect moving families after November 2001 and CLS continued attempts to find health visitors. It was eventually decided (in July 2002) to extend sample birth dates, and hence the period of fieldwork in Scotland and Northern Ireland by six weeks.



Coding frames were agreed with NatCen before fieldwork ended, but coding did not start until November 2002. Two variables offered particular challenges. One was the question about what each parent found best and most difficult about having the cohort child, the coding of which was delayed until after the first deposit. The other was the International Classification of Diseases on responses about long-term illness, which was done in-house at CLS.

NatCen delivered a dataset for Great Britain on 21 March 2003 and a (near) complete dataset for the UK on 8 April 2003. CLS and collaborators spent a short time cleaning the data and checking its consistency. After removing fields which might identify individuals CLS made the first deposit of the dataset on the Data Archive on May 23, 2003 with a revised version, containing more derived variables and what will probably not be the last installment of corrections following on 23 June 2003. The Archive had the dataset (and its documentation) ready for public use by the time of the launch of Millennium Cohort First Survey on 24 June 2003.

During the cycle of the First Survey, planning for the follow-up was also part of the team's remit. There was a meeting with collaborators and users (notably NESS) to discuss the timing of the follow-up, on 17 May 2001, and in October 2001 CLS were awarded another contract (starting 1 December 2001 from proposal) to conduct a follow-up survey at age 3. While details of this operation are outside the scope of this report, it should be noted that the launch and deposit of Sweep 1 coincided with the final stages (Dress Rehearsal Pilot) of developing the Sweep 2 Instrument.

The existence of a follow-up also means that our operations have expanded to cover cohort maintenance, i.e. keeping track of our informants and keeping them informed. They have been sent a thank-you letter after the interview, and a feedback document in June 2003 (the latter funded by another ESRC grant on Cohort Studies infrastructure and ONS contributions to the MCS).

The analysis of Sweep 1 has hardly started and is beyond the remit of this award. CLS hopes that academic collaborators and other users will soon be making use of this data resource, but it also has its own plans for a 'source book' funded from the ONS contribution to the survey.

## 6. Outputs

The principal outputs of the project have been the creation of a major, large national dataset, deposited for public use in the UK Data Archive. It contains 1,876 variables on 18,533 families and 18,819 children. It is also accompanied by a substantial set of documentation among which are the two Nominated Publications.

Both before, at and beyond the launch, considerable efforts have been made to disseminate the study to academic users, policy audiences and internationally. CLS is in the early stages of planning an International Conference of Longitudinal Studies of Children in the New Century, so that these studies may learn from each other and comparative analysis be facilitated.

Although analysis is beyond the agreement of the ESRC contract, ONS co-funding has enabled the team to prepare a preliminary Descriptive Report, due out within a few months of the launch and a Sourcebook of more considered analysis to be produced over the coming year. Over the same period it is planned to produce the third edition of the Technical Report on Sampling including an assessment of response bias, which, like other analysis requires the data collection to be completed, and to report on other enhancements by the ONS, the Health Visitor Survey, Linkage of Neighborhood Statistics with Hospital Episode and Birth Registration data.

## **7. Impacts**

The launch of the First Survey on June 24 2003 at the Institute of Education was part of the ESRC Social Science Week. Around 500 people were alerted to the survey's launch through invitation, and more than 100 attended to hear Minister Yvette Cooper praise the cohort studies as important social documents, relevant to policy as well as science. Preliminary results were displayed on posters and released to the press. The study received widespread and positive media coverage.

The first survey was already sufficiently successful for the second survey to have been commissioned. The government departments in the consortium have already commissioned analyses, whose results are not yet available. Another department, the Office of the Deputy Prime Minister, has enquired about an analysis of the housing data. The University of Strathclyde has commissioned an analysis of social and economic policies in Scotland using MCS data as evidence. The CLS has participated in a bid to DEFRA to provide longitudinal data on life in rural areas, to which MCS would contribute.

## 8. Future research priorities

The Millennium Cohort was designed to be a longitudinal survey, and its first follow-up is already being implemented. To continue this series it is necessary to plan the funding and co-funding of the third and further surveys. It has already been agreed in the National Strategy Committee for Longitudinal Data that the next sweep should be at age 5, so that material from schools can be included. It is important to explore the terms on which schools can be expected to co-operate. It is also important to ensure that the funding continues at a level which does not require either the length of the questionnaire or the size of the sample to be cut. It is also vital to plan for the resources to maintain contact with informants between surveys.

It may also be worth flagging up that our sampling strategy, namely using an opt-out from the Child Benefit register might no longer be available. The Inland Revenue has taken over Child Benefit records and are reported to have suspended the opt-out facility for new research attempting to reach disadvantaged families, who would be less likely to opt-in in writing. The ESRC may wish to point to the success of MCS as an argument in favour of the opt-out in correspondence with the government and to the Multi-Centre Research Ethical Authorities.

## **Ethical issues**

### **MREC**

The process of gaining medical research ethical approval adds to the time needed to develop the survey and proved a major hurdle to the MCS. As had been the practice with the previous cohort studies, medical research ethical clearance was sought in February 2001. This was as a general precaution for future health data collection. It was specifically required because of the proposal to involve Health Visitors and for permission to make links with vital registration and routine health data. (Any research using NHS staff needs to be given such clearance.) We were directed to the South West Multi-Centre Research Ethics Committee in March 2001, who felt that opt-out sampling could be coercive and would fail to obtain properly informed consent. They did however accept that written opt-ins would tend to exclude vulnerable people and so procedures were devised in consultation with the Committee to give potential respondents more information before they committed themselves for interview. Advance letters introducing the interviewer were sent shortly before her/his first visit and they were asked to arrange interviews, generally after their first visit, whose main purpose should be to give information. A simplified information sheet was produced and translated into several languages. These suggestions were helpful.

### **Codes of practice**

In order to support our assurances of confidentiality to informants, ethics committees, and government agencies to whose records links are being made, the CLS extended the Cohort Studies Code of Practice to cover all those working with MCS data and developed a Data Security Policy, setting out the secure, isolated computing environment which handles any named data files within CLS.

### **Geo-coded data**

One of our assurances to informants and the MREC was that to protect the anonymity of individuals, the exact locations of sample wards would not be published. We await guidance as to how to explore the geographical potential of the dataset without releasing ward identities or postcodes.