

Catalogue no. 11-522-XIE

**Statistics Canada International  
Symposium Series - Proceedings**

**Symposium 2006 :  
Methodological Issues in  
Measuring Population Health**



2006



**Statistics  
Canada**

**Statistique  
Canada**

**Canada**

## **Methodological Issues in Measuring the Mental Health of Children and Young People in Great Britain**

Meltzer, H.<sup>1</sup>, Wilmot, A.<sup>2</sup> and Dewar, A.<sup>2</sup>

### **Abstract**

In 1999, the first nationally representative survey of the mental health of children and young people aged 5-15 was carried out in Great Britain. A second survey was carried out in 2004. The aim of these surveys was threefold: to estimate the prevalence of mental disorders among young people, to look at their use of health, social and educational services, and to investigate risk factors associated with mental disorders. The achieved number of interviews was 10,500 and 8,000 respectively. Some key questions had to be addressed on a large number of methodological issues and the factors taken into account to reach decisions on all these issues are discussed.

KEY WORDS: Mental disorders, Children, Epidemiology

### **1. Introduction**

In 1999, the first nationally representative survey of the mental health of children and young people aged 5-15 was carried out in Great Britain. A second survey was carried out in 2004. The surveys were conducted by the UK, Office for National Statistics (ONS) on behalf of the Department of Health. The aim of these surveys was threefold: to estimate the prevalence of mental disorders among young people, to look at their use of health, social and educational services, and to investigate risk factors associated with mental disorders.

The interviews were carried out face-to-face (CAPI<sup>3</sup>) by qualified and experienced interviewers. Since the survey also involved interviewing those children aged 11 or over some additional training for interviewers was provided.

The achieved number of child interviews was 10,500 and 8,000 respectively. Interviews were also conducted with one of the child's parents using CAPI and one of their teachers, who completed a postal questionnaire.

Some key and quite complex questions had to be addressed on a large number of methodological issues. Some relate to the logistics of carrying out a survey of children; others pertain to assessing mental disorders among young people. These are addressed under a typology reflecting stages in the survey process.

### **2. Sample design**

We considered several methods of obtaining a representative sample of children in Great Britain – carrying out a postal sift of the general population to identify households with children, sampling through schools, using administrative databases and follow-up surveys. We decided to draw the sample from administrative records, specifically, the Child Benefit Records held by the Child Benefit Centre (CBC). Parents of each child under 16 living in the United Kingdom are entitled to receive child benefits unless the child is under the care of social services. Using these centralised records as a sampling frame was preferred to carrying out a postal sift of over 100,000 addresses and to sampling through schools. The postal sift would have been time consuming and expensive. Furthermore, response to contact by post can be difficult to obtain. We did not want to sample through schools because we wanted our initial contact to be with parents who could then give signed consent to approach the child's teacher. A two-hour survey was regarded as too long for follow-up on other surveys.

Using centralised records does have some disadvantages as access to the records can be problematic and the frame may not be accurate nor comprehensive. We were aware that some child benefit records did not have postcodes attributed to addresses: 10% in first survey and 5% in the repeat. However, the CBC had no evidence that records with postcodes were different from those without. The addresses with missing postcodes probably

---

<sup>1</sup> Department of Health Sciences, University of Leicester, UK

<sup>2</sup> Office for National Statistics, London, England

<sup>3</sup> Computer Assisted Personal Interview

represent a mixture of people who did not know their postcode at the time of applying for child benefit or simply forgot to enter the details on the form. If there are other factors which differentiate between households with and without post-coded addresses, the key question is to what extent these factors are related to the mental health of children. As these factors are unknown, we do not know what biases may have been introduced into the survey by omitting the addresses without postcodes.

Children in the care of social services only represent 0.5% of the population, but are a vulnerable group with higher rates of mental disorders. To ensure that these children were included in the research, they were surveyed separately (Meltzer et al., 2003)

Also excluded from the original sampling frame were those cases where “action” was being invoked. Examples of such action are: the death of the child and a change of address. They are simply administrative actions as distinct from some legal process concerning the child and hence should not have biased the sample in any way.

## **2.1 Ethical concerns in sampling**

In order to keep within data protection guidelines, we had to go through several steps to obtain our sample of children. First, we supplied the CBC with the list of 475 postal sectors we wanted covered after stratification<sup>4</sup>. We then gave the CBC details of how to calculate the sampling fraction to apply to their records in order to select 30 children whose parents received child benefit, in each postal sector. The CBC then sent a letter on our behalf telling the children’s parents about our survey and giving them an opportunity to opt out of the survey at that stage.

Using the first national survey as an illustration of using the database of child benefit records, 14,250 letters were despatched by the CBC on behalf of ONS: 30 letters for each of the 475 postal sectors. Nine hundred and thirty one of the sampled addresses (6.5%) contacted ONS via a free phone number to opt out and a further 790 addresses (5.5%) were found to be ineligible according to the survey population definition. The main reason for ineligibility was that the family had moved and could not be traced. This accounted for 629 of the 790 ineligibles (4.4% overall). Other reasons for ineligibility were that the child was deceased, in foster care, outside the age criteria (5-15), or the family had emigrated.

## **2.2 Accuracy of the sampling frame**

When the survey interviewer went to the address of the sampled child, their first task was to find out if the family still lived there and if the name was correct. Prior to the main stage survey, pre-pilot work and a pilot survey had been conducted. This pilot had indicated that about 5% of the addresses from the CBC were out of date because families had changed address and not informed the CBC. In many instances the child benefit was paid directly into a bank account and there was no incentive for the parent to inform the CBC of a change of address. Because the survey had a national coverage, attempts were made to trace the movers, and the family reallocated to another interviewer working in the vicinity of the new address. Interviewers also found that, in many cases, the names of both the selected child and the mother had changed owing to new relationships: remarriage, change of partner, etc. In this regard, the interviewers also checked that they had the right address.

# **3. Interviewing procedures**

## **3.1 Order of interview**

The first stage of the interview involved the completion of a 45 minute face-to-face interview with the parent asking about the child’s emotional state and behaviour. This included a five minute self-completion element comprising the GHQ12 (Goldberg and Williams, 1988) and the General Functioning Scale of the MacMaster Family Assessment Device (Byles et al, 1988). After the parent interview, permission was sought to ask questions of the sampled child. Children aged 11-15 had a face-to-face interview and entered details of their smoking, drinking and drug-taking experiences via a self-completion questionnaire on the laptop using CASI<sup>5</sup> technology. For children in care audio-CASI was used. All children, from 5-15 years, were administered reading and spelling tests.

---

<sup>4</sup> Stratification is a way of improving the precision of a random sample.

<sup>5</sup> Computer Assisted Self Interview.

When the parent and child interviews were completed, parents were asked for their written consent to contact the child's teacher. Parents were asked to nominate the teacher who they felt knew the child best. If the child had been expelled or excluded from school within the last few months, contact names for teachers were still sought. Before the teachers' questionnaire was posted out, various steps were taken to maximise response from them. For instance:

- a paragraph describing the survey was inserted into a journal which is sent to all teachers;
- Chief Education Officers were notified of the plans for the survey and the extent of teachers' involvement;
- a week before any postal questionnaires were sent to teachers, the head teachers in all schools of the sampled children, were notified that some of their staff would be sent a questionnaire to complete;
- the sample design (475 postal sectors and the stratification of the CBC list) was intended to reduce the burden on teachers so that most would not have to complete questionnaires on more than two children.

### **3.2 Choice of parent to interview**

In most interviews (over 95%), the interview with the parent was carried out with the mother as she tended to be available when the interviewer called. In the cases where the father was interviewed, this tended to be because the mother did not speak sufficient English to cope with the interview. The remainder were cases of lone fathers or fathers who were at home by themselves most of the time.

### **3.3 Logistics of arranging interviews**

The unpredictable length of the interview meant that interviewers had to make appointments when their parents would have a clear 90-120 minutes. This was often difficult for those parents who had several children with different 'pick-up' times from school and nursery, and parents who worked. In some areas, this meant that the interviewer might arrange an interview in the morning, but would have to complete it later when all the children were back from school or parents were back from work. Interviewers reported that some of the children had even busier 'social calendars' than their parents requiring a lot of flexibility (on the interviewer's part) to complete both the parent and the child interview.

### **3.4 Privacy**

The need for privacy in the interviews also affected the logistics of appointment making. It was obviously easier for the parent to provide information about their child if they were alone. This was clearly difficult for working parents. Children's interviews, by definition, had to be with the child alone, when the child was home from school, leading to the problems of excluding the rest of the family for what might be a considerable period of time. Some parents were initially taken aback that the interviewer needed to see the child on his/her own, though the great majority were happy with the explanations given. A technique successfully used by interviewers when parents refused to leave the room was to sit side by side with the child, reading out the questions but then asking the child to key their own answers into the laptop computer.

### **3.5 Use of self-completion to ask sensitive questions**

Although the presence of an interviewer is important for maintaining data quality and response to surveys their presence can influence respondents' answers to survey questions, particularly when asking questions about sensitive topics. Respondents exhibit what is known as a 'social desirability bias' providing answers which they consider to be concurrent with the beliefs of the interviewer or the norms of society in general (Dillman 2000). As well as the interviewer, the presence of other people can also affect the responses given (Bajekal and Purdon, 2001). Scott (1997) notes that children are just as likely as adults to exhibit social desirability bias.

In order to overcome potential social desirability effects, self-completion methods were used when asking for sensitive information. In this respect, CASI works much better than paper self-completion methods since research has shown that children have more trouble than adults in following routing patterns through questionnaires (Zukerberg and Hess, 1996).

Although the administration of CASI using laptop computers to ask sensitive questions - awkward and troublesome behaviour, smoking, drinking and drug taking - of children aged 11-15 appeared to work well, the

reporting of all types of substance use and abuse was under-reported compared with the national surveys of smoking and drinking carried out by group administration in school settings (Goddard and Higgins, 1999)

### **3.6 Language difficulties**

In some circumstances, neither parent had a sufficient grasp of English to be interviewed. Specific sections in the questionnaire dealing with the mental health of children, e.g. obsessions and compulsions, were quite difficult to formulate in English. To overcome this difficulty, the two-page, Strengths and Difficulties Questionnaire was made available in approximately 40 languages. This was used, in a self-completion format, instead of the face-to-face parent interview.

### **3.7 Incentive payments**

During the pilot stage of the survey, the need for incentive payments was investigated. Most teenagers agreed to be interviewed without such a payment. The lack of an incentive payment did not appear to affect response in the main stage of the survey, as shown by the high overall response rate, 96% of co-operating families. A small number of young people may have agreed to an interview with an incentive payment but this was not sufficiently widespread to recommend the universal use of incentive payments.

## **4. Survey design**

Two key decisions had to be made in deciding how to measure the prevalence of mental disorders of children and adolescents. The first related to whether to adopt a one- or two-phase design, i.e. ask all questions of all respondents or start off with a short screening instrument applicable to all children followed up with a detailed assessment with all screen positives and a sample of screen negatives. The second crucial question was who to interview. We decided on adopting a one-stage design with a questionnaire for parents, children aged 11 or over and teachers.

### **4.1 One- versus two-phase designs**

About half of the national surveys that have been carried out in other countries have used the multimethod-multistage approach of Rutter et al. (1970) to ascertain potential cases. In this approach, rating scales completed by children above a certain age and/or parents and/or teachers are used as first stage screening instruments. Subjects with scores above the cut-off score are identified as potential cases and further evaluated. A small sample of individuals with scores below the cut-off threshold is also selected for interview to assess the frequency of false negatives, i.e., those who have problems but whose rating scale scores were below the cut-off score. In the second stage, children with scores above the cut-off score and a sample of those with scores below this value are interviewed using semi-structured or structured psychiatric interview instruments. At this stage categorical diagnoses are made. The overall prevalence of disorder is determined at the conclusion of this two-stage process.

The other method does not base caseness upon the multimethod-multistage approach. All children and adolescents identified through the initial sampling procedure are eligible for diagnostic assessment. There are many advantages of such an approach:

- Detailed information is collected on all children. A sample distribution can be produced on all subscales even though only those with above-threshold score will have psychopathology.
- Because the survey aims to investigate service use, social disabilities, stressful life events, risk factors and the use of tobacco, alcohol and drugs, it is also important to have this information for all children for comparative purposes
- With the possibility of a longitudinal element in the survey, there is a large pool of children from which to select controls who could be matched on several characteristics to the children who exhibit significant psychiatric symptoms during the first stage interview.
- A one-phase design is likely to increase the overall response rate compared with a two phase (screening plus clinical assessment) design.
- A one-stage design reduces the burden put on respondents. Ideally, a two-phase design would require a screening questionnaire to be asked of a parent, a teacher as well as the child, followed up with an assessment interview administered to the child and the parent. A one-stage design only requires an interview with the parent and child and, if possible, the administration of a teacher questionnaire.

- One of the advantages of a one-phase over a two-phase design is that it can be carried out in a far shorter timescale.

The main disadvantage of a one-stage design is cost: the administration is far cheaper in two-stage designs, although the latter is likely to elicit more biases and less precision.

## 5. Questionnaire design

Following the pre-pilot and piloting stages of the survey testing, data analysis and interviewer debriefings helped to inform the final design of the questionnaire with respect to the question wording, response options and the order in which the questions were asked.

Historically, in terms of the measurement of psychiatric morbidity, researchers felt that they had to make a choice between a full-structured interview carried out by lay interviews or a more free-flowing, semi-structured interview carried out by a clinician. The questionnaire designed for the present study was intended to combine some of the best features of structured and semi-structured measures. Using existing semi-structured measures for a large national survey would have been impractical and prohibitively expensive since it would have required recruiting a team of several hundred clinically trained interviewers or providing prolonged additional training and supervision to lay interviewers.

Given the practical and financial imperative to use lay interviewers with relatively little additional training, it was clear that the main interviewing would need to be fully structured. The disadvantage of relying entirely upon existing structured interviews is that the results are far less clinically convincing than the results of surveys based on semi-structured interviewing. When informants answer fully structured interviews, they often over-report rare symptoms and syndromes because they have not really understood the questions. (Brugha et al., 1999) To circumvent this problem, the questionnaire used structured interviewing supplemented by open-ended questions. When definite symptoms were identified by the structured questions, interviewers used open-ended questions and supplementary prompts to get parents to describe the problems in their own words. The wording of the prompts was specified.

### 5.1 Clinical input

Answers to the open-ended questions and any other respondent comments were transcribed verbatim by the interviewers but were not rated by them. Interviewers were also given the opportunity to make additional comments, where appropriate, on the respondents' understanding and motivation. A small team of experienced clinicians reviewed the transcripts and interviewers' comments to ensure that the answers to the structured questions were not misleading. The same clinical reviewers could also consider clashes of information between different respondents, deciding which account to prioritise. Furthermore, children with clinically relevant problems that did not quite meet the operationalised diagnostic criteria could be assigned suitable diagnoses by the clinical raters. There are no existing diagnostic tools that combine the advantages of structured and semi-structured assessments in this sort of way, which is why a new set of measures were specifically designed for this survey. The new measures and their validity are described in more detail elsewhere. (Goodman et al., 2000).

### 5.2 Single versus multiple respondents

While single-respondent investigation characterised nearly all of the early epidemiological studies, more recent studies (within the multi-method multistage approach) have broadened data collection to include information gathered from parents, teachers, and the children themselves. Hodges (1993) has pointed out that children and adolescents can respond to direct questions aimed at enquiring about their mental status and that there is no indication that asking these direct questions has any morbidity or mortality risks. However, information from many sources is a better predictor of disorder than just one source. Many experienced clinicians and researchers in child psychiatry believe that information gleaned from multiple respondents facilitates the best estimate of diagnosis in the individual case. Agreement between child and parent has varied depending on type of pathology; but at the population level, information from multiple respondents enhances the specificity of prevalence estimates. (Young et al., 1987; Angold 1989; Hodges, 1993).

## 6. Analysis

### 6.1 Combining information from various sources

One of the problems of collecting information from various sources in different formats, both qualitative and quantitative, is finding the best way to integrate the information which may show a lack of agreement. One method has been to accept a diagnosis irrespective of its source (Bird et al., 1992). Others have promoted “case vignette” assessments where clinical judgements are made on detailed case histories from several sources (Goodman et al., 1996).

This case vignette approach for analysing survey data uses clinician ratings based on a review of all the information on each respondent. This information includes not only the questionnaires and structured interviews but also any additional comments made by the interviewers and the transcripts of respondents’ comments to open-ended questions particularly those which ask about the child’s significant problems. The case vignette approach was extensively tested among community and clinical samples in the pre-pilot and pilot phases of the survey.

Using this approach the clinical raters performed four major tasks. First, they used the transcripts to check whether respondents appeared to have understood the fully structured questions. This was particularly valuable for relatively unusual symptoms such as obsessions and compulsions - even when parents or young people said “yes” to items about such symptoms, their own description of the problem often made it clear that they were not describing what a clinician would consider to be an obsession or compulsion.

Second, the clinical raters considered how to interpret conflicts of evidence between respondents. Reviewing the transcripts and interviewers’ comments often helps decide whose account to prioritise. Reviewing all of the evidence, it may be clear that one respondent gives a convincing account of symptoms, whereas the other minimises all symptoms in a defensive way. Conversely, one respondent may clearly be exaggerating.

Third, the clinical raters aimed to catch those emotional, conduct and hyperactivity disorders that slip through the ‘operationalised’ net. When the child has a clinically significant problem that does not meet operationalised diagnostic criteria, the clinician can assign a ‘not otherwise specified (NOS)’ diagnosis such as ‘anxiety disorder, NOS’ or ‘disruptive behaviour disorder, NOS.’

Finally, the clinical raters relied primarily on the transcripts to diagnose less common disorders such as anorexia nervosa, Tourette syndrome, autistic disorders, agoraphobia or schizophrenia. The relevant symptoms are so distinctive that respondents’ descriptions are often unmistakable.

### 6.2 Differential non-response among sources

In most cases, data were available from both the parent and child where appropriate, but for a substantial number of cases (20%), the teacher did not provide any information.

It can be assumed that, given a complete set of data from both home and school, the psychiatrists would on average, make the right assessment. Therefore, if they were able to use both home and school information to assess all children in the population and we were able to average the results of this census over a large number of repeats under identical circumstances, they would arrive at the prevalence level in the population. Therefore if this complete information were available for all sampled children in the survey, the estimated prevalence level would be unbiased for the actual prevalence level, differing only through sampling error and response error. On the other hand, if repeated censuses were taken but collecting only the home information, another average prevalence would be measured. The question is: would the prevalence level measured with the school and home information taken together be the same as with just the home information? If the measures were the same, then we would not need to carry out adjustments for the missing teacher data. If the measures were different, some adjustment factor would need to be incorporated into the reported data. The evidence for the need to carry out this readjustment comes from looking at the ratio of parent-based to clinical-based diagnoses in no teacher information ( $t=0$ ) and with teacher information ( $t=1$ ) groups. If the ratio is the same, it suggests that having the teacher report doesn’t make a significant contribution. If the ratio is higher for the  $t=1$  group, it suggests that clinical diagnoses underestimate prevalence in the absence of teacher reports.

The prevalence of the three broad categories of disorder (emotional disorders, conduct disorders, hyperactivity) and for any disorder were calculated for  $t=0$  and  $t=1$  conditions, followed by the clinical-parent ratios. The results fitted in with clinical experience: teacher reports contribute little if anything to the diagnosis of emotional disorders but make a substantial difference to the diagnosis of conduct and hyperkinetic disorders, though for rather different reasons. As far as conduct disorder is concerned, the key issue is that there are a lot of children who are oppositional, aggressive and antisocial at school but not at home. Psychiatrists rarely get to know about these children unless they have a teacher report. As far as hyperkinetic disorders are concerned, there are many children where the parental evidence is inconclusive and where the teacher report tips the balance (Goodman, 1999). Hence disorder adjustment factors (emotional disorders (1.00); conduct disorders (1.12); hyperkinetic disorders (1.09) and any disorder (1.06) were applied to the tables on prevalence.

## References

- Angold A (1989) Structured assessments of psychopathology in children and adolescents, in Thompson C (ed), *The Instruments of Psychiatric Research*, John Wiley & Sons Ltd.
- Bajekal M, and Purdon S. (2001) *Social capital and social exclusion: development of a condensed module for the Health Survey for England*, London: The Stationery Office.
- Brugha TS, Bebbington PE and Jenkins R (1999) A difference that matters: comparisons of structured and semi-structured psychiatric diagnostic interviews in the general population, *Psychol Med.* **29(5)**:1013-20
- Byles J, Byrne C, Boyle M H and Offord D R (1988) Ontario Child Health Study: Reliability and validity of the General Functioning Scale of the MacMaster Family Assessment Device. *Family Process* **30(1)**, 116–23.
- Dillman D. A., (2000), *Mail and Internet Surveys: The Tailored Design Method*, 2<sup>nd</sup> Edition, New York: John Wiley & Sons.
- Goddard E and Higgins V (1999) *Smoking, drinking and drug use among teenagers, Volume 1, England*, London: The Stationery Office.
- Goldberg D and Williams P (1988) *A User's Guide to the General Health Questionnaire*, NFER-NELSON
- Goodman R (1999) Personal communication.
- Goodman R, Ford T, Richards H, Gatward R and Meltzer H (2000) The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology, *J Child Psychol Psychiatry.* **41(5)**:645-55.
- Goodman R, Yude C, Richards H and Taylor E (1996) Rating child psychiatric caseness from detailed case histories. *Journal of Child Psychology and Psychiatry* **37**, 369–379.
- Hodges k (1993) Structured Interviews for Assessing Children, *Journal of Child Psychology.* **34**, 49-68
- Meltzer H, Gatward R, Corbin T, Goodman R and Ford T (2003) *The mental health of young people looked after by local authorities in England*, London: The Stationery Office.
- Scott, J. (1970) *Children as respondents: Methods for improving data quality*. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz and D. Trewin (Eds), *Survey measurement and process quality* (pp.331-350). New York: Wiley-Interscience.
- Young J G, O'Brien J D, Gutterman E M and Cohen P (1987) Research on the clinical interview. *J. Am. Acad. Child Adol. Psychiatry* **26,5**, 613–620.
- Zuckberg, A. L., and Hess, J. (1996). Uncovering adolescent perceptions: Experiences conducting cognitive interviews with adolescents. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 950-955.