Data Resource Profile: the Mental Health of Children and Young People Surveys (MHCYP)

Tamsin Ford *, Tim Vizard 2, Katharine Sadler 3, Sally McManus 3, Anna Goodman 4, Salah Merad 2, Maria Tejerina-Arreal 2, Dan Collinson 5 and the MHCYP collaboration

Address for corresponding author *

College House
St Luke’s Campus
Magdalen Road
Exeter
EX1 2LU

1 University of Exeter Medical School
2 Office for National Statistics
3 National Centre for Social Research
4 London School of Hygiene and Tropical Medicine
5 NHS Digital

Word count 2965

Keywords
Children, Young People, Mental Health, Mental Disorder, Well-Being

Topic
UK, Mental health, Children and young people, Surveys
Data Resource basics

There is an increasing international focus on mental health with a particular emphasis on its importance among children and young people (Patel at al. 2018). Large population-based datasets with robust measures therefore present an important resource for policy and practice related researchers, particularly if they incorporate longitudinal data.

The Mental Health of Children and Young People Surveys (MHCYP) comprise three comparable cross-sectional, population-based surveys that were conducted in the United Kingdom in 1999 (n=10,438), 2004 (n=7,797) and 2017 (n=9,117). They aimed to estimate the prevalence of mental health conditions among school-age children to inform policy, service organisation and clinical practice. Each survey involved multi-informant standardised diagnostic assessment of all participating children’s mental health in a large stratified probability sample. The single phase approach in such large samples differentiates these surveys from many other large population-based mental health surveys, in which brief screening questionnaires are used to identify those at risk of disorder, who then proceed to a more detailed diagnostic assessment (McManus et al., 2016). Each survey applied similar methods, which allows comparison across time. All three baseline surveys cover children aged 5-15 years, while 2004 also included 16 year olds and 2017 further extended the age range to include pre-schoolers (aged 2-4 years) and young people aged 17-19 years.

The first two surveys were repeated in the same sample after three years and follow up of the 2017 participants is currently anticipated. An overview of the time-line is provided in Figure 1. Data from the 1999 (https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=4227), 2004 (https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=5269) and 2017 surveys are available via the UK Data Service archive, while data from the 2017 survey requires permission to be granted by NHS Digital before the data can be downloaded (https://digital.nhs.uk/services/data-access-request-service-dars).

England’s Department of Health and Social Care has provided funding for all the surveys; the most recent survey was commissioned by NHS Digital.

These surveys provide an essential resource for researchers interested in the important and topical issue of mental health among children and young people, given the repeated use of the same methods, the universal application of standardised multi-informant assessment of mental disorder and the availability of longitudinal data. They complement the parallel survey programme of adult mental health that preceded the first child survey (McManus et al, in submission).

Insert Figure 1 about here

Data collected


For the 1999 and 2004 surveys, the child benefit register was used to develop a sampling frame of postal sectors from Great Britain. Across the two surveys, approximately 10% (901/9000) of postal sectors were sampled with a probability related to size of sector, and stratified by regional health authority and socio-economic group within it, which represented approximately 90% of all British children once adjusted for inaccurate or ineligible benefit records (Stringaris & Goodman, 2009). The Child Benefit Office contacted parents before their details were passed to the researchers (6% and 9% opted out in 1999 and 2004 respectively; Figure 1 and Figure 2). Productive mental health assessments were completed with 73% and 65% of the original sample in 1999 and 2004 respectively, which equated to 83% and 76% of those approached for interview.
Child benefit was a universal state benefit payable for each child in the family across Great Britain with an extremely high uptake until 2013, when changes made to child benefit eligibility meant it was no longer representative of higher income households. The 2017 survey therefore used the National Health Service (NHS) Patient Register to provide a direct sample of children based on their age, greatly reducing the sampling of ineligible addresses, saving interviewer time and burden on households without children (n=9117). The use of the NHS Patient Register meant children under the care of the local authority were also included in 2017, who comprise 0.6% of the child population. They were previously covered in separate surveys (Ford et al., 2007a). The NHS Patient register offered the most comprehensive and efficient sampling frame for the 2017 survey, but has some coverage issues (Office for National Statistics, 2016). Families may be slow to register with a General Practitioner (GP) when they move or leave the country, and the former is particularly a problem with university students who may have two addresses simultaneously but can only be registered with one GP. Similarly, the armed forces and people using private health care only are under-represented.

While the 1999 and 2004 surveys included children from Scotland and Wales as well as England, the 2017 survey sampled children from England only. Children were selected from a stratified probability sample of approximately 6.3% of postcode sectors (460/7275). These were stratified by Government office region, selected relative to the size of the region, and sorted on factors associated with mental health disorders that were derived from 2011 Census data, such as tenure, employment status and economic activity. This created a sample frame 18029, of whom 17326 were eligible (Figure 3; Vizard et al., 2018). Productive mental health assessments were conducted with 52% of those eligible and 59% of those approached. The decline in response rates over time across the three surveys is in line with other similar studies over this time (Bolling and Smith, 2018).

Parents and children were interviewed face to face using computer assisted interviews by trained lay interviewers, including sections administered by self-completion for some more sensitive topics. If the family agreed, a teacher was mailed a questionnaire. The parent was the primary contact for children up to the age of 16 years, with young people invited to interview if their parent agreed and if they were aged 11 or over. Young people aged 17 or older were the primary contact whose consent was required to involve others. Parents and young people who agreed to researchers contacting the school nominated a teacher they thought knew the child well.

Follow up surveys for 1999 and 2004
To aid sample retention, brief questionnaires were sent to parents between the first two surveys and the three year repeat of the survey, which were completed after three years, in 2002 and 2007 respectively (Figure 1 and Figure 2). The schedule and coverage of these follow ups differed according to policy interests and funding, but on each occasion included the Strengths and Difficulties Questionnaire (SDQ), a dimensional measure of mental health completed by parents (Table 1).
The follow up surveys had different sample frames. All children meeting diagnostic criteria for psychiatric disorder (911) in the 1999 survey and a random sample of those who did not (2334) were invited to the 3-year follow-up survey in 2002 (Figure 1). After exclusion of the 9% who were ineligible as they had died or had been taken into the care of the local authority, the response rate was 88% of those approached (2,586; Meltzer et al., 2003). The 2007 follow up included as many of the children who participated in 2004 as possible (Figure 2; total invited=7329, 90%) (Parry-Langdon 2008); 8% were lost to follow up and 12% refused, while 343 children were ineligible; the final sample of 5326 children represented 78% of those approached.

Assessment of mental health
Table 2 illustrates the mental health measures collected in each of the baseline and follow up surveys as well as which informant provided data for which condition.

The assessment used (Development and Well-Being Assessment www.dawba.info; Goodman et al, 2000) combines the advantages of highly structured questions about a large range of mental health conditions, with semi-structured probes for all informants when difficulties were reported. The nature and time frame of the highly structured questions relate directly to the diagnostic criteria in the Diagnostic and Statistical Manual (DSM; American Psychiatric Association, 1994; American Psychiatric Association, 2013) and the International Classification of Disease (ICD; World Health Organisation, 1992), so the measure generate point prevalence. The DAWBA was developed for the 1999 survey and has been updated since, so that the version used in 2017 included DSM 5 (American Psychiatric Association, 2013) as well as ICD 10 diagnoses. The semi-structured probes seek to elicit a detailed account of any problems reported in the closed questions in terms of their nature, frequency and severity. The interviewers record the comments verbatim for subsequent review by clinical raters (Goodman et al. 2000).

For the MHCYP surveys, a small team of experienced child psychiatrists reviewed the structured and qualitative information at symptom level, presented simultaneously for each disorder from all informants, to assign diagnoses according to both DSM and ICD criteria (American Psychiatric Association, 1994; American Psychiatric Association, 2013; World Health Organisation, 1992). Clinical rating permits the detection of misunderstandings by informants that invalidate their answers to the structured questions, the management of conflicting reports from informants as they would be assessed in the clinic and the assignment of “not otherwise specified diagnoses” for children with clinically significant difficulties that fail to fit the current diagnostic criteria but would warrant clinical intervention. The kappa statistic for chance-corrected agreement between two clinical raters who independently rated 500 children in the 1999 survey was 0.86 for any disorder (standard error SE 0.04), 0.57 for internalising disorders (SE 0.11), and 0.98 for externalising disorders (SE 0.02) (Ford et al., 2003).

In addition, the Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001) was completed by parents, young people and teachers (Table 2). The SDQ (www.sdqinfo.com; Goodman, 2001) comprises 25 items divided between five sub-scales, which generate scores for conduct problems, hyperactivity-inattention, emotional symptoms, peer relationships and prosocial behaviours. The first four are summed to generate a total difficulties score, which has good internal consistency for a UK population (Cronbach’s alpha of 0.73; (Goodman, 2001)). Items are phrased as either a positive or negative statement and informants select between the following responses; not true, somewhat true or certainly true, scored 0, 1, 2 or reversed with positive statements, so that a high score indicates difficulty. The exception is the prosocial scale in which higher scores indicate better skills.
The SDQ impact supplement generates a score based on ratings of the impairment in friendship, classroom learning (all informants), home life, and leisure activities (children and parents only). The final question asks about the burden of difficulties to the informant (rated not at all, only a little, quite a lot, a great deal).

The Warwick Edinburgh Mental Well-Being Scale (WEMWBS; Clarke et al., 2011) was completed by young people in the 2017 survey. The WEMWBS is a 14 item scale with 5 response categories (none, rarely, some, often, all of the time; www.warwick.ac.uk/fac/sci/med/research/platform/wemwbs/), summed to provide a single score ranging from 14-70. The items are all worded positively and cover feeling and functioning aspects of mental wellbeing. The measure was designed to enable monitoring of wellbeing in the adult general population, but a validation study with 13-16 year old English and Scottish school children demonstrated construct validity and revealed strong internal consistency (Cronbach’s alpha of 0.87; Clarke et al, 2011).

Other measures in the baseline and follow up surveys
In addition to the comprehensive assessment of mental health, a wide range of child, parent, family and neighbourhood characteristics were assessed in each survey, as outlined in Table 3 and described in more depth in the 2017 survey methods report (Vizard et al, 2018). In 2017 we obtained consent to link to NHS health records and education records (such as the Medical Research Information Service Integrated Database & Administration System (MIDAS), Hospital Episodes Statistics, ONS mortality data, the Mental Health Services Dataset and the National Pupil Database; (https://files.digital.nhs.uk/23/FC1278/MHCYP%202017%20Appendix%20G%20-%20Fieldwork%20Documents.pdf))

Data resource use
The data from the earlier surveys has been highly influential for mental health policy in relation to children and young people. The survey reports (Meltzer et al., 2000; Green et al., 2004) are commonly cited in policy documents. One example of an influential finding from the 1999 survey is that only a quarter of children with impairing mental health conditions reach child and adolescent mental health services, while the commonest professional source of mental health support were teachers (Ford et al. 2007b; Newlove-Delgado 2015). The latter finding had major resource implications for the education system (Snell et al., 2013; Knapp et al., 2015) and is one of the drivers of a recent policy focus in England on mental health in schools (Department of Health and Social Care and Department for Education, 2017).

The longitudinal elements of the data have allowed the study of how mental health changes over time and what influences this (Ford et al., 2007c; Ford et al, 2017), as well as the study of how parents, young people and teachers report difficulties differently (Collishaw et al, 2009). Similar analyses have allowed areas of particular policy interest to be explored in more depth, such as irritability and bipolar disorder (Stringaris and Goodman, 2009; Stringaris et al., 2010) and the relationship between callous and unemotional traits and behaviour disorder (Rowe et al., 2010).

There remains a wealth of questions that could still be addressed from these data, as well as from the planned data linkage.

Currently we are aware of plans to study trends in mental health and service use, and the links between poor mental health and function at school.
Strengths and weaknesses

These data provide large, nationally representative samples in which to study the association between mental health in children and many background variables. Survey weights have been constructed to account for non-response and response bias. Mental health was assessed using a validated, multi-informant and standardised diagnostic instrument with all participating children and measures were kept as consistent as possible to allow the examination of trends over time. There is also longitudinal data and the prospect of data linkage to routine health and education datasets for the 2017 survey. By covering both treated and untreated conditions, these surveys present unique data for assessing the extent of the treatment gap, as well as trends and inequalities in treatment use.

In such large samples, measures of background variables are inevitably brief, although validated and robust questionnaires have been applied wherever possible. Inevitably there have been some changes in measures; the DAWBA assessment has been updated over the years, so that not all concepts have been measured at all time points. DSM was revised, so the 2017 survey has DSM 5 and not DSM IV diagnoses, but ICD 10 diagnoses are available across all the surveys.

Response rates were higher in the earlier surveys, but have declined in similar surveys over the last two decades, so the 2017 data are not likely to be particularly unreliable in relation to other recently gathered samples (Bolling and Smith, 2018). The follow up studies demonstrate higher response rates among those in good initial mental health, than amongst children meeting diagnostic criteria for a mental disorder (Meltzer et al., 2002; Parry-Langdon, 2008). This is also a common finding, and while likely to lead to underestimation of incidence or prevalence, does not necessarily invalidate associations between predictors and mental health (Wolke et al, 2009). Response rates for 17-19 year olds were lower than for other groups (40% c.f. 52% overall). Finally, the samples were carefully selected to represent the population from which they were drawn, which means that they provide relatively few children from Black and Minority Ethnic backgrounds. Analyses of mental health by ethnicity can have limited power once other confounders have been adjusted for so these data are not ideal for exploring or comparing mental health among children from different ethnicities. The similarity between the different surveys allows the combination of datasets to offset this problem (Goodman et al., 2010). While a similar analysis suggests that parents, young people and teachers from deprived neighbourhoods were 8% less likely to participate than those from the most affluent area (Goodman & Gatward, 2008). This difference in participation seemed largely to be explained by difficulties in making contact rather than refusal and in contrast, opt out from initial contact letters was highest among the most affluent neighbourhoods. While these datasets contain sufficient numbers of children and families from deprived circumstances to support most analyses, researchers should be aware of the selective response, which is echoed in differential retention in the longitudinal studies.

Data resource access

These surveys have comprehensive reports that can be accessed via NHS Digital (https://digital.nhs.uk/pubs/mhcypsurvey17), which include access to the survey documents. Data from the 1999 (https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=4227), 2004 (https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=5269) and 2017 surveys are available from the UK Data Service (https://digital.nhs.uk/services/data-access-request-service-dars). However, for the 2017 dataset only, permission for use will need to be granted before the data can be downloaded. Applications for data use permission will need to be made to the Data Access Request Service (DARS) at NHS Digital (https://digital.nhs.uk/services/data-access-request-
Researchers should be aware that the application process has to comply with the new General Data Protection Regulation, which was enacted in May 2018. Staff at DARS are working hard to streamline the process but researchers would be wise to consult them early and work with them closely to avoid frustrating delays in access.

There is at the time of writing no official acknowledgement paragraph required in papers reporting analyses of these data.

**Funding**

The 1999 survey was commissioned by the English Department of Health, Scottish Health Executive and National Assembly for Wales, as were the follow ups. Data collection was led by the Office for National Statistics (ONS) with input to the clinical rating from Kings College London. The 2004 survey was commissioned by the Department of Health and the Scottish Executive Health Department, as were the follow ups, and data collection was led as before. The 2017 survey was funded by England’s Department of Health and Social Care, commissioned by NHS Digital and conducted by a consortium from National Centre for Social Research (NatCen), ONS, University of Exeter Medical School and YouthinMind.

**Acknowledgements**

We would like to thank the children, their parents and their teachers, as well as our colleagues on the steering groups, and from England’s Department of Health and Social Care, ONS, NatCen, University of Exeter Medical School, and YouthinMind for their contributions to these data. In particular, we are also indebted to team of ONS and NatCen interviewers who ensured participants were able to take part in this survey. Howard Meltzer, Bob Jezzard and Rachel Jenkins were instrumental to the success of the initial survey programme (1997-2007). Robert Goodman designed the mental health and many other measures, and led the clinical rating in all the surveys conducted to date. Support from Alison Neave at NHS Digital in the conduct and reporting of the 2017 survey was invaluable. The CYPMH collaboration includes Tamsin Newlove-Delgado, Eva Wooding, Dhriti Mandalia, Franziska Marcheselli, Ellie Brodie, Si Ning Yoeh, Nicola Pearce, Jodie Davis, Nick Forbes, Laura Brown and Vahé Nafiliyan.

**References**


Office for National Statistics. Patient Register: quality assurance of administrative data used in


Profile in a nutshell

The Mental Health of Children and Young People surveys estimated the prevalence of clinically relevant mental health conditions and trends in mental health and risk factors over time. Data were collected in 1999 (n=10,438; England, Wales and Scotland), 2004 (n=7,977; England, Wales and Scotland) and 2017 (n=9,117; England).

These cross-sectional population-based single phase surveys drew separate stratified probability samples of school-age children who all received a multi-informant, standardised, diagnostic mental health assessment. The two earlier surveys were followed up after three years, and data linkage to the health and education administrative data is planned for the 2017 survey.

Data from the 1999 (https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=4227) and 2004 (https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=5269) surveys are available from the UK Data Service (UKDS), while data from the 2017 survey will be available from the archive in 2019 with NHS Digital permission. Applications for data use permission will need to be made to the Data Access Request Service (DARS) at NHS Digital (https://digital.nhs.uk/services/data-access-request-service-dars). Access is free for those intending to use the data for non-commercial research, teaching or training purposes and belonging to a research institution.
### Table 1 Content of the interim postal surveys

<table>
<thead>
<tr>
<th></th>
<th>Baseline Survey and months from baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1999: 6 months follow up</td>
</tr>
<tr>
<td>Strengths &amp; Difficulties Questionnaire (SDQ)</td>
<td></td>
</tr>
<tr>
<td>parent/YP/teachers</td>
<td>parent/child</td>
</tr>
<tr>
<td>Service provision through school (e.g.: school counselling, educational psychologist advice)</td>
<td>parent/teacher</td>
</tr>
<tr>
<td>General Health Questionnaire (GHQ12)</td>
<td></td>
</tr>
<tr>
<td>Size of primary support group</td>
<td></td>
</tr>
<tr>
<td>Level of perceived social support</td>
<td></td>
</tr>
<tr>
<td>Follow up version of SDQ with additional questions about change in symptoms and impairment</td>
<td></td>
</tr>
</tbody>
</table>

¹ children aged 11 or over who were originally interviewed
<table>
<thead>
<tr>
<th></th>
<th>Survey year</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Development and Well-Being Assessment (DAWBA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body dysmorphic disorder</td>
<td>P,C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Attachment disorder</td>
<td>P,C</td>
<td></td>
<td></td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>Disruptive Mood Dysregulation Disorder</td>
<td>P,C,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pervasive developmental disorders</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eating disorder</td>
<td>P,C</td>
<td>P,C</td>
<td>P,C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other less common disorders</td>
<td>P,C</td>
<td>P,C</td>
<td>P,C</td>
<td>P,C</td>
<td>P,C</td>
</tr>
<tr>
<td><strong>Rosenberg Self-esteem Scale</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Warwick Edinburgh Mental Wellbeing Scale (WEMWBS)</strong></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
</tr>
<tr>
<td><strong>Moods and Feelings Questionnaire</strong></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
</tr>
<tr>
<td><strong>Parental mental health (General Health Questionnaire 12)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P Parent, C Child or young person aged 11+, T teacher
**Figure 1. Timeline of the 1999, 2004 and 2017 Mental Health of Children and Young People baseline and follow-up surveys**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline survey</td>
<td>Follow up SDQ</td>
<td>Follow up survey</td>
<td>Base line survey</td>
<td>Follow up SDQ</td>
<td>Follow up SDQ</td>
<td>Follow up survey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SDQ, Strengths and Difficulties Questionnaire mailed to participants as a keeping in touch exercise
Figure 2 - Participant recruitment to the Child and Young People’s Mental Health Survey 1999 the subsequent follow ups
(Full survey involves face to face interview with parents and young people aged 11 years and a postal questionnaire to teachers with a standardised psychiatric diagnostic assessment to generate ICD 10 and DSM IV diagnoses)

1999
Sample frame = Child
Benefit Records
n=14,250
Opt-outs: 931,
Moved no trace: 629
Ineligible: 161
All not approached for interview: 1721

Approached for interview: n=12,529
Refusals: 1774
Non-contact: 317

Baseline sample n=10,438
Full survey

1999 six months postal SDQ questionnaire to wave 1
(1/3 of sample)
n=2,344

2000 Postal Parental and young people SDQ and service use
Sent to 3992 (all with disorder, 1/3 without disorder at baseline)
n=2944 (2348 no disorder, 596 with disorder)

Approached for follow up interview: n=2,938
All refusals: 306
Non-contact 45
Other non-response: 1

2002 Follow up sample n=2,586
Full Survey
Figure 3– Participant recruitment to the Child and Young People’s Mental Health Survey 2004 and the subsequent follow ups
(Full survey involved face to face interview with parents and young people aged 11 years+ and a postal questionnaire to teachers with a standardised psychiatric diagnostic assessment to generate ICD 10 and DSM IV diagnoses)

2004
Sample frame = Child Benefit Records
n= 12,294

Opt-outs: 1,085,
Moved no trace: 631
Ineligible: 82,
All not approached for interview- 1,798

Approached for interview-
n=10,496

All refusals- 2,183
Non-contact- 313
Insufficient information for diagnostic classification- 23

Baseline sample n= 7,797

2004 (6 months post baseline for 1/3 data collection) Postal Parental SDQ n=1,174

2005 Postal Parental SDQ from all 2004 participants consenting to re-contact n= 6,375

2006 Postal Parental SDQ from all 2004 participants consenting to re-contact n= 5,794

Approached for follow up interview
n=7,329

2007 Follow up sample n = 5,326
Figure 4 Participant recruitment to the Child and Young People’s Mental Health Survey
2017

Sample Frame
19320
42 per PSU

Addresses issued
18029

Ineligible = 393
Refusals = 4956
No contacts = 2194
Other Unproductive = 1369

Achieved interviews with one or more participants in a household
9,117

Main Sample = 15960
postcode sectors (PSUs) = 380
Reserve Sample = 3360

Main Sample = 15690
Reserve Sample = 2399