



RESEARCH ARTICLE

# A systematic approach to mapping longitudinal data usage: Reflections on tracking Millennium Cohort Study activity [version 1; referees: awaiting peer review]

Dylan Kneale <sup>1</sup>, Praveetha Patalay<sup>2</sup>, James Thomas<sup>1</sup>, Meena Khatwa<sup>1</sup>, Claire Stansfield<sup>1</sup>, Emla Fitzsimons<sup>2</sup>

<sup>1</sup>Evidence for Policy and Practice Information and Co-ordinating Centre, UCL Institute of Education, University College London, London, WC1H 0AL, UK

<sup>2</sup>Centre for Longitudinal Studies, UCL Institute of Education, University College London, London, WC1H 0AL, UK

**v1** First published: 26 Sep 2018, 7:1559 (doi: [10.12688/f1000research.15990.1](https://doi.org/10.12688/f1000research.15990.1))  
 Latest published: 26 Sep 2018, 7:1559 (doi: [10.12688/f1000research.15990.1](https://doi.org/10.12688/f1000research.15990.1))

**Abstract**

**Background:** The Millennium Cohort Study is the youngest of the UK’s four national birth cohort studies, but the only study (to our knowledge) where a systematic approach to exploring data usage has been undertaken.

**Methods:** In this paper we: (i) explore previous exercises and provide justification for our approach; (ii) share headline findings of our research, (iii) outline the challenges of intersecting systematic review methods with survey design methods; and (iv) discuss the implications for future survey design as well as for future exercises tracking survey data usage. All of the results were obtained through undertaking systematic searches across 30 databases which generated over 4000 results. We then searched these records, first on title and abstract and then on the full text and extracted data on studies that fell within our specific areas of interest.

**Results:** A total of 481 studies were identified as using MCS data in novel analyses. Among these studies, measures that have been collected across sweeps—diet, BMI, SDQ and screen time—are all comparatively well used. Data that were collected from the child’s own reports (e.g. friendships and feelings) have seldom been utilised in comparison to data collected through parental reports and using validated tools (e.g. SDQ). Imposing thresholds on data was found to be problematic in some cases, for example for BMI, where a number of different thresholds for overweight and obesity were in use. The use of different thresholds can lead to substantial differences in the results obtained.

**Conclusions:** Longitudinal consistency in measures is key to identifying change over time, and the review helped map the degree of consistency in measures, and their utility. The findings shaped decisions around inclusion of variables in MCS7 (age 17 years), as well as the way in which existing data were deposited.

**Keywords**

cohort studies, Millennium Cohort Study, systematic review, longitudinal studies, child development, child health

**Open Peer Review**

**Referee Status:** *AWAITING PEER*

*REVIEW*

**Discuss this article**

Comments (0)

**Corresponding authors:** Dylan Kneale ([D.Kneale@ucl.ac.uk](mailto:D.Kneale@ucl.ac.uk)), Emla Fitzsimons ([E.Fitzsimons@ucl.ac.uk](mailto:E.Fitzsimons@ucl.ac.uk))

**Author roles:** **Kneale D:** Conceptualization, Formal Analysis, Methodology, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Patalay P:** Conceptualization, Formal Analysis, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Thomas J:** Conceptualization, Methodology, Project Administration, Writing – Review & Editing; **Khatwa M:** Formal Analysis, Investigation, Writing – Review & Editing; **Stansfield C:** Formal Analysis, Methodology, Writing – Review & Editing; **Fitzsimons E:** Conceptualization, Funding Acquisition, Methodology, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** We acknowledge funding from the Economic and Social Research Council in completing this study (grant numbers ES/K005987/1 and ES/M001660/1).

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2018 Kneale D *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Kneale D, Patalay P, Thomas J *et al.* **A systematic approach to mapping longitudinal data usage: Reflections on tracking Millennium Cohort Study activity [version 1; referees: awaiting peer review]** *F1000Research* 2018, 7:1559 (doi: [10.12688/f1000research.15990.1](https://doi.org/10.12688/f1000research.15990.1))

**First published:** 26 Sep 2018, 7:1559 (doi: [10.12688/f1000research.15990.1](https://doi.org/10.12688/f1000research.15990.1))

## Introduction

### Why review longitudinal data usage?

The UK is home to a number of nationally and geographically representative longitudinal studies that track human development across the life course. Fielding effective instruments to respondents is essential to ensure their continuation. For longitudinal studies, ensuring that respondent burden is kept to a minimum and that respondents feel valued members of the study are particularly important considerations to encourage continued participation. Alongside respondent burden, perhaps the most significant limiting factor is the cost of fielding a question to a large sample of respondents. Metrics of fielding a successful sweep of a longitudinal survey are not formally established but are likely to include measures such as response rates, attrition and representativeness, costs, and whether the data gathered was suitable for addressing pertinent research questions. While proxy indicators are relatively well established for examining the suitability of instruments for respondents (Krosnick, 1999; Presser *et al.*, 2004), less attention has been focussed on how to measure whether questions and data are suitable for meeting researcher and user needs. Suitable data for data users and researchers can only have been collected through questions that elicited reliable responses from respondents; however, not all reliable data will necessarily meet the needs of researchers and conversely researchers may use poorer quality or proxy measures in the absence of measures that directly support their needs. Some data may simply fall ‘under the radar’ of researchers, including those instances where the data are of little research value or policy relevance.

Given the substantial costs in fielding questions, as well as the ethical considerations of collecting but not using data from respondents, there is a need to explore how we can better understand patterns of data usage. In this paper we aim to report on our experience of measuring data usage from the UK’s youngest nationally representative birth cohort study, the Millennium Cohort Study (MCS), and report on some of the methodological choices and issues we encountered, present an overview of our findings, and discuss the implications of our findings and methods for similar exercises and data usage in the future. Here we report on a novel substantive focus for techniques used in systematic mapping and discuss their suitability for exploring survey data usage.

### Previous exercises in reviewing longitudinal data usage

Reviewing the contribution of longitudinal survey data is not a new science and a body of literature is emerging that summarises the findings of longitudinal studies. These forms of enquiry can be divided into three categories. The first category includes cohort profiles that describe the development of longitudinal data sources and showcase their main objectives, strengths and findings of studies and describe the breadth of data collected; they may also review some of the main contributions to knowledge that these studies have offered (see Connelly & Platt, 2014 for an example). While not always explicitly stated, an underlying motivation of these cohort profiles is to publicise the existence and encourage the usage of the data through advising potential (and existing) users of the content, design and access to the data (Joshi & Fitzsimons, 2016;

Power & Elliott, 2006). A second group of studies examines the contribution of different longitudinal studies to advancing knowledge around a given topic or research question; an example includes Joshi’s (2014) non-systematic literature review, which examines findings from the 1958- and 1970-born cohort studies around non-cognitive development among children as part of a study examining the intergenerational transmission of social advantage, which was later extended to other longitudinal sources (Joshi *et al.*, 2016). As well as selectively reviewing substantive contributions of longitudinal data, one of the latent objectives of this form of review is to support users of longitudinal data in designing their own studies through highlighting research gaps and potential approaches that could be adopted to address these (for example Corden & Millar, 2007; Joshi, 2014). A third subset of studies utilises systematic review techniques to examine a tightly defined subject or question based on studies published from longitudinal data. An example using MCS data is Twamley & colleagues’ (2013) review of the evidence of how the involvement of fathers influences child and maternal mental health during early years. The prime aim of this form of study is to make a substantive contribution to the body of evidence through utilising systematic review techniques. Commonly the synthesis involves narrative synthesis of the results of studies, but evaluating the utility of different variables is often only a secondary consideration. In this paper we report on a fourth approach to reviewing longitudinal data usage, where the aim is to utilise a systematic approach to reviewing the literature, and to apply this to appraise the utility of different question areas and scales in the MCS.

The MCS is a longitudinal interdisciplinary study following the lives of just over 19,500 children born in the UK in 2000/2001. The study recruited families of children born in randomly selected electoral wards, disproportionately stratified to boost representation in England of children from disadvantaged and ethnic minority families; and with oversamples also from Scotland, Wales and Northern Ireland. Information has been collected at 9 months, 3, 5, 7, 11 and 14 years, with the next sweep of data collection at age 17 years being fielded at the time of writing. Over the course of the first two waves, approximately 19,000 households were recruited into the study; by age 7 the number of participating families had dropped to 13,800 and at age 14, just under 11,800 families were contacted. A wide range of data have been collected from children, parents and guardians, the partners of parents/guardians, older siblings and teachers, as well as sub-studies that collected data from health visitors; these include self-reported and objectively measured/verified as well as linked data from administrative records. The remainder of the paper focusses on the methods we used and our overall substantive and methodological learning from applying this approach to studying the MCS. The work presented in this paper is based on a previously published, non-peer-reviewed report available from the EPPi Centre website (<http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3502>) (Kneale *et al.*, 2016).

## Methods

### Mapping systematically and clarifying the scope

Systematic reviewing involves conducting a through an explicit, rigorous and accountable process of discovery, description

and assessment of literature according to defined criteria, followed by a synthesis of the cumulative evidence around a given condition or intervention (Gough *et al.*, 2012), with methods developed for statistical synthesis (meta-analysis) of the evidence across studies (Borenstein *et al.*, 2011). As an often undefined stage of systematic reviewing, but also as an independent exercise in its own right, producing a systematic map of the literature involves summarising the topography of the evidence landscape around a given issue. Systematic mapping can be considered a more appropriate research tool in the presence of a broad research question and can be used to develop a narrower research question for a systematic review. Producing a systematic map of the literature follows many of the same stages as a systematic review in the formation of a research question, identification and clarification of key concepts for use in the search strategy and in defining the inclusion criteria, and some degree of data extraction. However, a systematic map may differ in the rigidity of the inclusion/exclusion criteria employed (for example, greater inclusivity in research design), in the narrative synthesis methods employed to summarise the map rather than to address tightly defined research questions, and in the absence of formal method of quality assessment.

One of the first stages for producing a systematic map of MCS data usage was to clarify the aims. Clearly the MCS is a large study, with over 3,500 variables deposited in the age 11 standard dataset, for instance. In addition, there were already indications that a large body of evidence had accumulated. The Centre for Longitudinal Studies (CLS) has maintained a bibliography of cohort study publications that is populated through user notification and supplementary web searching. It represented a bibliography of publications, as opposed to studies, meaning that the same study could appear multiple times, for example as a conference paper, working paper and journal article. Furthermore, not all the included studies directly reported on new empirical analyses of the MCS and some were reports of MCS analyses published in other papers. Therefore, one of the aims of this work was to establish the number of unique studies that were reporting on primary analyses of MCS data. However, in order to further understand patterns of data usage and to inform the design of future sweeps of the MCS and other child cohort studies, there was a need to (i) identify where potentially under-explored areas of data may lie for MCS users, and (ii) highlight examples where detailed response categories are rarely used. Mapping out the totality of MCS data usage and meeting these objectives was an undertaking beyond the scope of the project and priority areas of research were identified based on specific topic areas, questions or scales. This meant that the study would be able to identify the total number of studies using MCS data through systematic methods, but that a systematic map of *how* all MCS data are used across different topic areas was focussed only on a core set of questions (see Table 1 for a list of measures identified).

The ten areas for in-depth mapping were selected to represent:

- (i) Allied topics/scales where usage could be contrasted by the type of scale used (e.g. Strengths and Difficulties

Questionnaire (SDQ) and Child Social Behaviour Questionnaire (CBQ) for dimensions of child behaviour);

- (ii) Topics where usage could be contrasted in terms of respondent (parent/teacher reports (SDQ; CBQ) vs child report (feelings, school dis/like, friends));
- (iii) Allied topics where usage could be contrasted in terms of whether they are usually specified as outcomes or as antecedents (outcomes (e.g. Body Mass Index (BMI)) and antecedents (e.g. diet, screen time));
- (iv) Topics of high policy relevance (arguably all fall within this category, but immunisation was selected as representative here).

This meant that some important areas, notably cognitive development, were sacrificed in order to conduct a more thorough examination of these chosen constructs.

#### Systematically searching for and identifying evidence

Our strategy was first to systematically identify MCS studies through implementing a search across databases, and then secondly to search within these studies for those that focussed on subject areas in Table 1 using specialist systematic review software (EPPI-Reviewer 4 (see Thomas *et al.*, 2010)). We tested a search strategy that was based on variants of MCS and was implemented across a number of datasets. For an indication of the comprehensiveness of the search, we were able to compare our results against the CLS bibliography. Specifically, we tested whether a simple search based on 'Millennium Cohort Study' and variants (see Supplementary File 1) would be sufficient to capture studies or whether a more in-depth search strategy was necessary. We conducted preliminary searches based on the simpler set of search terms in Supplementary File 1 and compared these to a snapshot of 60 publications in the CLS bibliography (approximately 15% of records held for CLS publications).

Of these 60 studies, 14 were identified as problematic as they did not appear in our initial set of studies. When we examined these records further we found that six would not meet our inclusion criteria as they did not use MCS data directly but instead reported on the results of MCS data published elsewhere (see details in Kneale *et al.*, 2016). Of the remaining eight studies identified, we found that a search that included the terms in Supplementary File 1, which looked for their occurrence anywhere in the document (as opposed to title and abstract only) and implemented through Scopus and Science Direct, located seven of these studies. The remaining study was a CLS working paper and was not indexed in these sources; as a result, CLS working papers were added as a specific source. This testing was used to justify our approach of implementing a small search across a large number of databases to locate studies using MCS data, and then to screen the results for inclusion across any one of the chosen subject areas. One deviation from this was in our search for economic literature, where the search on EconLit was expanded to include terms

**Table 1. Priority study areas by topic for in-depth systematic mapping.**

Topic area	Description and rationale
Strengths and Difficulties Questionnaire (SDQ)	The SDQ measures children and young people's behaviour, emotions and relationships. It is composed of five domains: (i) emotional symptoms, (ii) conduct problems, (iii) hyperactivity/inattention, (iv) peer problems and (v) prosocial behaviour. Four domains (i-iv) are summed together to calculate a total difficulties score (Goodman, 1997). The SDQ was selected as a question area that was thought to be well utilised.
Child Social Behaviour Questionnaire (CBQ)	The CBQ is an extension of the Adaptive Social Behaviour Inventory (Hogan <i>et al.</i> , 1992), and has been implemented in different forms across different sweeps of the survey (see later section). The CBQ is distinguished from other measures, including the SDQ, through including a focus both on pro-social and anti-social behaviours (Warden <i>et al.</i> , 2003). Three domains are covered in the CBQ: independence and self-regulation; emotional dysregulation and cooperation (Johnson, 2012). The focus here was on understanding how CBQ is used in studies compared with well-known alternatives such as the SDQ.
Diet	Several questions have been developed and fielded around children's diet and nutrition in the MCS. However, the comprehensive nature of the questions – which include more general fields such as frequency of consuming portions of fruit to more specialist data such as whether children eat fish on Fridays for religious reasons – mean that it is likely that some questions or areas may rarely be used in published studies. Our focus here was on distinguishing which questions are commonly used in sweeps 2, 3 and 4 and which are rarely used.
BMI	Data collected on children's BMI are thought to feature in several publications using MCS data. However, of particular interest here is the way in which BMI data are used. Grouped data have historically been derived and deposited using International Obesity Task Force thresholds for identifying obesity/overweight among children (based on the gender and age of the child); the UK90 thresholds are an alternative set of 'rules' to categorise children as being obese/overweight.
Immunisation	It is hypothesised that immunisation data collected at 9 months has been used in several MCS-based studies. However, the extent to which 'completed' immunisation histories have been used in studies is unknown and the focus here is on establishing the way in which immunisation information collected at ages 3 and 5 years have been used in studies
Hobbies and Interests	Two of the earlier British birth cohort studies (1958 and 1970) have fielded extensive questions on surveys (albeit at later ages), which have been used in different disciplines, for example in examining resilience among young people (Schoon & Bynner, 2003). Age 7 saw the inclusion of a self-completion module that asked children about their hobbies, and we investigate the use of these data in published studies.
Children's feelings	One of the weaknesses of measures such as the SDQ is that it is reliant on the observations of parents/teachers and not a reflection of the child's own perceptions. Through asking about children's feelings at age 7 via a self-completion module, the MCS attempted to capture the child's voice, and we investigate the use of these data in published studies.
School dis/like	Several 'objective' and established measures of cognitive development and school readiness are collected in the MCS. Additional questions capturing the child's own perceptions of school may provide additional insight and may be predictive of future indicators of school performance and adjustment. School dis/like has been used in studies examining the outcomes of older children and young people in terms of engagement in risk behaviours (Bonell <i>et al.</i> , 2007), and the focus here is on how these data are used to examine the outcomes of younger children in the MCS.
Friends	Through asking about children's friendships at age 7 via a self-completion module, the MCS attempted to capture the child's perception of their peer group, and we investigate the use of these data in published studies.
Screen time	BMI and diet are issues that are thought to be well used in studies using MCS data. However, another important predictor of BMI is thought to be physical activity and sedentary behaviour. Here the focus is on examining how widely used indicators of sedentary behaviour at ages 3, 5 and 7 are used, and how these data are used in studies.

reflecting 'birth cohort' (and UK geography) as well as those in [Supplementary File 1](#); this did not yield additional results after screening. Therefore, a simple search strategy conducted across a wide range of sources (29 in total (see details in [Kneale \*et al.\*, 2016](#))) was deemed to be an efficient way of identifying studies using MCS data, albeit with the caveats outlined in the conclusion.

All records were inputted into EPPI-Reviewer 4 for further screening (4,329 records). Records were first screened for

duplicates, with just under half of records identified as duplicates and excluded (2,056 records). All remaining records were screened on the basis of title and abstract by two reviewers (DK and MK); any disagreements that could not be resolved were to be referred to other team members (although this did not prove to be necessary). Initial title and abstract screening mainly focussed on whether the data being used were MCS data. This involved excluding studies using data from a US-based Millennium Cohort Study (a study of military veterans) and the Gateshead Millennium Cohort Study. Studies that used MCS

data, but clearly were not using the variables in [Table 1](#) were excluded but marked separately from others (and rescreened) in order that we could accurately obtain a complete list of MCS studies. Full texts of records that were deemed to be using MCS data and were focussed on any one of the variables in [Table 1](#) on the basis of title and abstract were retrieved and subject to a second round of full text screening by two reviewers (DK and MK). Both reviewers used a list of questions and potential synonyms for the terms used in questions to establish eligibility. We retrieved the full record for 224 publications to examine their relevance at this stage.

Studies were deemed eligible for in-depth analyses if they used MCS data from one of the variables in [Table 1](#) as a main dependent or independent variable in their analyses. ‘Main’ variable was defined on the basis of the scope of the study as outlined in the aims/objectives or research questions. Where studies did not clearly specify an independent variable of interest in the aims/objectives—for example, if the study explored which of a range of factors predicted a specific outcome of interest—then we examined whether there was a focus on the question areas of interest in the literature review or conceptual framework. We aimed to exclude studies where the question area in scope was being used only as a background control variable as we were unlikely to be able to systematically identify this occurrence across all studies. This was often made apparent in studies when parameter estimates in models were not published or discussed in the write up. Studies could be included as being relevant across multiple areas of interest.

Information was extracted on: the country and institution of the lead author; study sweep(s) of data used; other data sources analysed in study; questions used in analysis; aims/objectives of study; analytical methods used in analysis; additional study design notes; whether measures were used as outcome variable or main predictor of interest; findings/results; strengths of the data/measures; difficulties reported in using data/measures and/or study limitations; recommendations for future research/data collection; journal discipline; citations of study (based on those listed on Google Scholar). Data extraction forms were piloted first before being completed for each study. Where a reviewer was unable to populate a particular field, the advice of a second reviewer was sought. The results are presented in full elsewhere ([Kneale et al., 2016](#)), and here we focus on the summary points that represent both the substantive and methodological learning we uncovered.

## Results

### Overall patterns of data usage

The total number of unique MCS studies identified was 481. This was a higher number of records than found on the CLS Bibliography at the time of the search (481 vs 440); however, the results represented a greater volume of studies (as opposed to publications) as we did not include duplicates, and did not include reviews, reports or news of other MCS studies that did not include primary analyses of MCS data (including, for example, the review of fatherhood studies discussed earlier ([Twamley et al., 2013](#))). We observed that a systematic approach to discovering MCS studies results in a substantially higher volume of

studies being identified than was the case through methods that rely on researcher cooperation and were supplemented through non-systematic web searches. Again, a relatively simple search strategy implemented across a comprehensive range of data sources was found to yield efficient results.

### Patterns of data usage among priority areas

A number of measures that had been collected across different sweeps—diet, BMI, SDQ and screen time—were comparatively well used and featured as a focus in 11, 49, 121 and 16 studies, respectively. Those measures that started to be collected at age 7 (and first made available in 2010: hobbies, feelings; school dis/like; friends) had a substantially lower usage and each featured in a maximum of two studies; furthermore many of the studies using these data were descriptive reports published to coincide with the depositing of MCS data in data archives. Overall, [Table 2](#) clearly shows that data that are collected through a recognised and well-validated scale with defined thresholds or cut-off points for identifying constructs of interest and/or data that can provide a unique insight into a policy-relevant issue, were those most widely used among the topics selected for in-depth mapping.

Strengths and Difficulties Questionnaire (SDQ) data (collected in three surveys to MCS4) had by far been the most widely used of the 10 topic areas in focus here, and featured as a main independent or dependent variable in 121 studies (out of the total 481 MCS studies). Study authors identified the strengths of SDQ measures in MCS as including their repeated collection which enabled the implementation of longitudinal modelling strategies and contributed to the understanding of developmental trajectories (for example [Dillenburger et al., 2015](#); [Hartas, 2012](#); [Jokela, 2010](#); [Pronzato & Arnstein, 2013](#)) and identification of some of the moderators (and mediators) of these trajectories. For example, the repeated nature of SDQ observations was used by [Midouhas & colleagues \(2013\)](#) to examine how trajectories of psychopathology were moderated by family-level circumstances among children with autism. Other strengths of SDQ identified by authors included that the data were collected from different informants, parents and sometimes teachers ([Hartas, 2012](#); [Kelly et al., 2013](#); [Zilanawala et al., 2015](#)), which allowed for a degree of validation between reports, as well as the availability of data across different SDQ domains, which allowed one example study to explore differential impacts of contextual risk factors across the different domains ([Flouri et al., 2010](#)). Almost two-fifths of studies using SDQ (39%) relied solely on the total domains score, while most other studies examined one or more subscales, often alongside the total difficulties score. We found only one example where a single question was used as the basis for analysis, in a study focussed children’s subjective well-being which used information on whether parents viewed their children as ‘often unhappy’, an item from the emotional symptoms subscale ([Chanfreau et al., 2013](#)). We also found other studies that used items from SDQ outside the SDQ scoring framework. For example, [Delaney & Doyle \(2012\)](#) used items from the hyperactivity/inattention scale in combination with two items from the Child Social Behaviour questionnaire to derive three factors (inhibition, compulsivity, impulsivity) in their examination of socioeconomic differentials of ‘time discounting’.

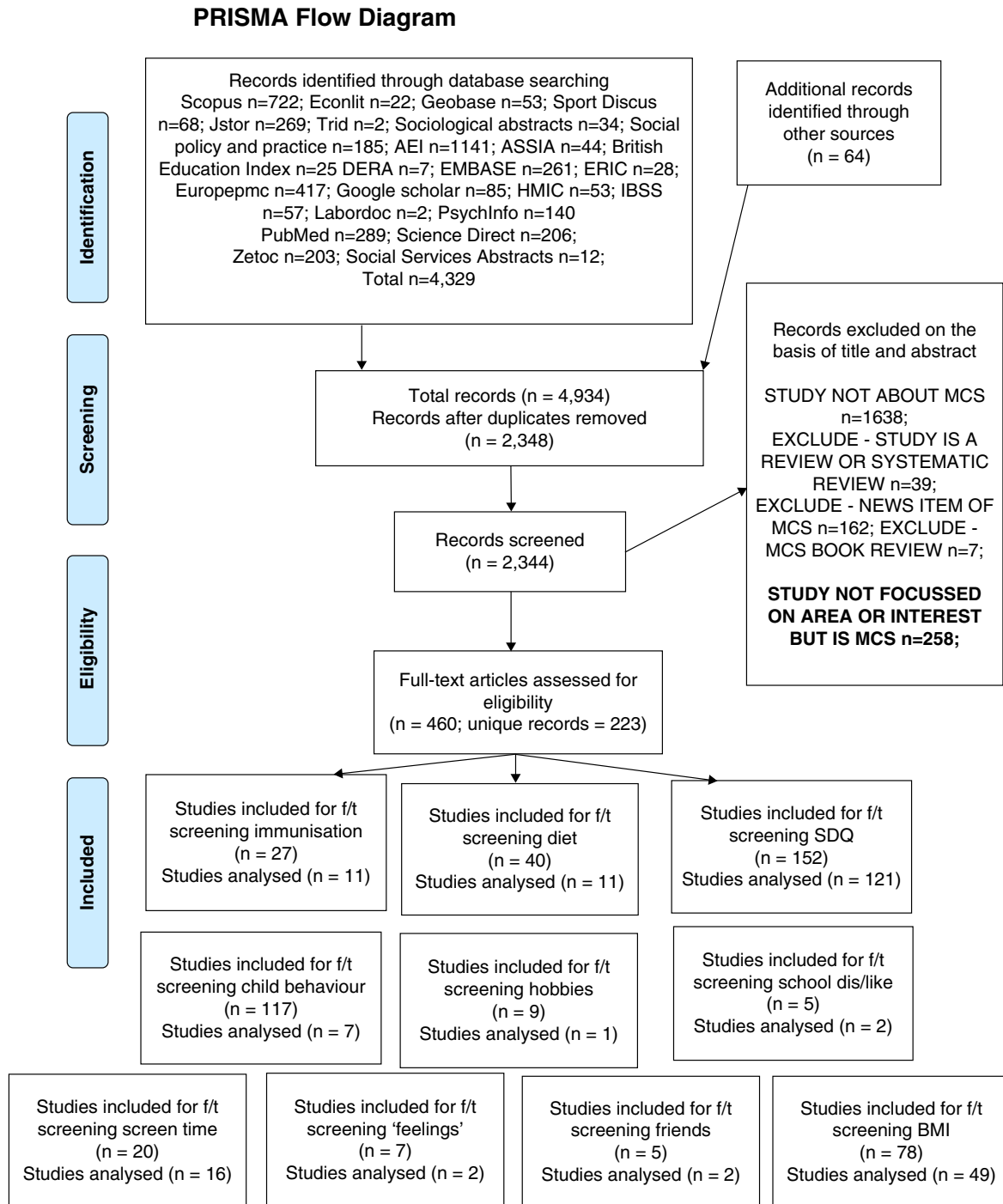


Figure 1. PRISMA diagram of the flow of studies through the review.

The popularity of SDQ in the MCS follows its status as a recognised scale, collected at different time points and from different informants (parents, and at age 7 years, teachers), and with defined thresholds for identifying problem behaviour.

In contrast, an allied measure of child behaviour, the CBQ, was collected solely from parents' reports and was developed as part

of a longitudinal study examining the Effective Provision of Pre-school Education (EPPE) in the UK. For the CBQ, clear thresholds or cut-offs for identifying constructs of interest are not widely reported. Several of the studies identified as using CBQ data did not clearly report the exact questions that were being used, and there was even ambiguity as to *how* to refer to the CBQ scale in terms of nomenclature. In the absence of

**Table 2.** Topic areas by number of studies using the data and characteristics.

Topic	Collected at multiple points at ages 9 months, 3 years, 5 years or 7 years ****	Recognised scale or measurements used	Recognised thresholds for scale available	Multi-informant	Number of studies
SDQ (strengths and difficulties)	✓	✓	✓	✓	121
CBQ (child social behaviour)	✓	✓	-	-	7
Hobbies (self-reported)	(first repeated at 11)	-	-	n/a	1
BMI	✓	✓	✓**	n/a	49
Feelings (self-rated)	(first repeated at 11)	-	-	n/a	2
Diet	✓	-	-	-	11
Feelings about school (self-rated)	(first repeated at 11)	-	-	n/a	2
Friends (self-rated)	(first repeated at 11)	-	-	n/a	2
Immunisations at age 3 and 5	✓	-	-	-	11
Screen time	✓	-	***	-	16

\*\*Different alternatives available; \*\*\*A (US) threshold for recommended maximum hours is available but not calculable in the data; n/a, not applicable; single informant construct. \*\*\*\* shows if these indicators were collected at multiple points during the first four sweeps of data collection; first repeated at age 11 shows that these indicators were only available from a single point in the first four sweeps

publishing full details of the questions used, many authors referred to a technical report from the EPPE project on the usage of the CBQ measures. However, this report in itself does not clearly provide technical guidance on how to construct measures and whether thresholds for underlying constructs should be imposed (as in the case for SDQ) (Sammons *et al.*, 2003). Nevertheless, the CBQ was used in seven studies, and did capture some domains that would be otherwise unavailable, including, for example, self-regulation (Flouri *et al.*, 2014).

BMI data were also widely used in the literature, reflecting concerns about increasing rates of childhood obesity, which was substantiated in one paper through comparing levels in MCS with previous cohorts (Johnson *et al.*, 2015). BMI data also shared many of the same properties as SDQ in terms of being a measure collected in a similar way across waves with defined thresholds for identifying overweight and obese children (albeit with different thresholds in use in the literature, see below), and consequently featured as a main variable of interest (either continuously or in categories) in 49 studies.

Unlike BMI, data on children's diets were utilised less frequently as the focus of a study, appearing in 11 studies. This

may be due to the quality of the data, and some authors reported the need for objective measures of diet and for better measures of the frequency of consumption of different foods. This would have included collecting objective data through tools such as food diaries (Brophy *et al.*, 2009). The lack of nuanced objective data on children's nutritional intake was thought to undermine some of the observed associations between children's diet and other outcomes including BMI. For example, the association between irregular breakfasting and higher BMI uncovered in Brophy & colleagues' (2009) study may be an artefact of irregular dietary intake and compensatory snacking, although this cannot be investigated further as measures of nutritional intake are not collected. Similarly, others have highlighted that accurate measures around the frequency of intake of snacks are not collected (Sullivan & Joshi, 2008), as well as a more broadly, a detailed inventory of what the children eat, how frequently, and in what quantities (Connelly, 2011).

Data on immunisations at age 3 and 5 years did not feature in many publications. However, those using MCS data were highly cited. One of the unique strengths of the MCS data is that they were able to directly reflect and address the research needs of policy-makers in terms of understanding antecedents



of MMR uptake. For example, the study of [Pearce & colleagues \(2009\)](#) examined children who had not been vaccinated against MMR, and uncovered that for around three-quarters of children this was through conscious choice, highlighting the level of misinformation around MMR combined vaccines that was prevalent at the time at which MCS data were collected. Crucially, MCS data were able to provide a unique insight into uptake of single vaccines as well as combined vaccines; these data were not readily available elsewhere ([Anderberg \*et al.\*, 2011](#)).

Screen time data were collected in a diffuse way across different sweeps, although data on the frequency of television viewing was collected consistently across all three sweeps of interest (age 3, 5 and 7 years). Screen time data featured as a focus in 16 studies and the MCS was viewed as one of the few studies that allowed for examination of patterns of screen entertainment while controlling for a broad range of sociodemographic factors ([Griffiths \*et al.\*, 2010](#)). It is also one of the few studies that allows for longitudinal analysis of relationships between screen time and outcome measures ([Parkes \*et al.\*, 2013](#)).

### Specific issues around granularity and data usage

A question that we wished to address in this research was to identify where granularity was lost in the data. That is, where detailed data are collected from respondents, but where such granularity is obscured by the need to collapse response categories to achieve a workable sample size for that category. Contrary to our expectations, we saw little evidence of granularity being ‘lost’ in this way, although this is likely a reflection of these data being underutilised. Nevertheless, two examples were identified where grouping data seemed to be somewhat problematic. The first was in terms of screen time, where data on TV viewing and computer usage were collected in bands, but where these bands did not correspond to the American Academy of Paediatrics (AAP) recommendation<sup>1</sup> that screen time be limited to 1–2 hours per day. This meant that authors were not directly able to measure whether MCS children exceeded the AAP limits, although some did attempt to impose thresholds regardless. A further potential mismatch between the recommended thresholds is also observed to some extent in the case of fruit consumption, where data are collected on the number of fruit portions consumed, but the UK guidance around minimum consumption refers to fruit *and* vegetable consumption ([NHS Choices, 2015](#)). Therefore, it was not possible to measure whether MCS children were consuming the recommended number of portions of fruit or vegetables per day.

The second example where grouping data were found to be problematic was in the case of BMI, where a number of different thresholds for overweight and obesity were in use. The use of different thresholds can lead to substantial differences in the proportions of children classified as overweight/obese; for example, a Colombian study of children aged 5–18 years found differences of almost six percentage points in the prevalence of overweight/obese children when applying different thresholds of overweight/obesity ([Gonzalez-Casanova \*et al.\*, 2013](#)). Most

MCS users classified overweight/obesity using International Obesity Taskforce (IOTF) thresholds (29/49 studies); less commonly researchers used Centres for Disease Control (CDC) thresholds (6/49 studies), the World Health Organisation thresholds (4/29) and the UK90 thresholds (4/49). MCS data have traditionally been deposited with pre-constructed variables reflecting International Obesity Taskforce thresholds for obesity<sup>2</sup>. Meanwhile the National Obesity Observatory (at the time) recommended that in England, the British 1990 (UK90) growth reference charts should be used to determine the weight status of an individual child and population of children<sup>3</sup>, although with the caveat that other thresholds may be more appropriate dependent on the research question ([National Obesity Observatory, 2011](#)). Perhaps most concerning was that some users failed to report which definition was used (6/49 studies), impeding the comparison of results entirely.

### Discussion

1998 saw the announcement that funding would be provided for a new cohort study tracking the development of individuals born in the new millennium. [Joshi & Fitzsimons \(2016\)](#) outline some of the founding principles of the MCS including that the study should ‘capture as much detail on the child’s origins that may later turn out to be relevant’ to explain differentials in life course trajectories and outcomes. Meeting the needs of diverse groups of end users of a multipurpose study, including policy-makers, third sector organisations, academics, and ultimately the wider public, is not without its challenges. The properties of individual instruments may be appraised through measuring their reliability and validity, as well as establishing their responsiveness to change longitudinally, and determining the substantive focus of such instruments is usually dependent on the research question. Beyond measuring the scientific properties of the questions, there is no (known) standard method for evaluating the content of a survey, or more importantly for measuring the impact or success of fielding different instruments.

In the current mapping exercise, a simple search strategy that was implemented across a number of different databases significantly outperformed the existing methods of identifying studies using MCS data, identifying over 481 studies in total. Through systematically mapping the usage of ten different areas of questions or instruments in the published grey and peer-reviewed literature, we also confirmed that data that are collected through a recognised and well-validated scale with defined thresholds for identifying constructs of interest and/or data that can provide a unique insight into a policy-relevant issue, are those most widely used. Unusually, data collected from children themselves (at age 7 years) were not well utilised, although this may reflect the quality of the instruments used to collect these data, as well as the domains covered by these instruments themselves. Nevertheless, collecting self-reported information on domains that are meaningful to children themselves, such as their hobbies

<sup>2</sup>At Age 14, data were deposited that included derived variables for overweight/obesity using both IOTF and UK90 thresholds.

<sup>3</sup>Also featured here ([http://www.noo.org.uk/NOO\\_about\\_obesity/measurement/children](http://www.noo.org.uk/NOO_about_obesity/measurement/children)) (Accessed 07/03/16)

<sup>1</sup>No known UK equivalent exists

and friendships, may be of greater substantive interest in future longitudinal studies and may also serve as a means of engaging cohort members' future participation.

This study was one of the first to map systematically how data from a longitudinal survey are used in the literature. To fit within the resources for the exercise, the remit was restricted to ten priority areas which were selected in conjunction with the study management team. This means that while we were able to create a count of MCS studies through systematic means (481 studies), further mapping was more focussed, resulting in limitations in terms of coverage of topic areas (e.g. cognitive development and measures on parental characteristics). We also excluded studies that analysed data collected at age 11 from the systematic review, as the data had only been deposited a short time before conducting the review (a total of three studies were identified as using these data; none falling within our priority areas of interest). There were further limitations to our approach. Firstly, some databases only allow for title and abstract searching. Therefore we were dependent on users including mention of the study somewhere in a word-restricted abstract. We were concerned that this was unlikely to be common practice in economic literature in particular and expanded the search parameters, although this produced no additional results after screening. Relying on title and abstract is also likely to mean that we have undercounted working papers and conference papers, where the abstract is often unavailable or is not indexed. Furthermore, use of MCS data by third sector organisations as part of reports or briefings is also likely to be underrepresented. Encouraging authors to name the data source in the title/abstract would increase the likelihood of discovery in future studies, and is a recommendation that has implications beyond the MCS. A second limitation is that our conclusions around the utilisation of different topics was based on identifying these as the focus of a paper. Often this status can be hard to ascertain and is accompanied by a degree of subjectivity. While we did employ a standard definition in our screening, this may still have been open to interpretation, particularly in terms of studies testing a range of different predictors simultaneously with only a broad research question guiding variable selection. A third limitation was that our conclusions around utilising data are based on studies publishing their findings. Very few studies reported results that were not statistically significant for their variable of focus; Kelly and colleagues' study provided one of the few examples where indicative although statistically insignificant associations were the focus of the paper (Kelly *et al.*, 2013).

Recommendations, which are also applicable to other longitudinal studies, can be made around how future data usage mapping exercises could be facilitated through the further development of a community of MCS users. Establishing a searchable database of MCS users could help to foster a community of users. The database could hold a short entry with users' contact details, topic areas of interest and key variables of interest. This would allow MCS users to develop links with others with similar interests, and potentially foster collaborations between users and across institutions. This database could also be used as the

basis of future work in contacting users for consultations for future sweeps and other forms of user engagement. Participation in such a database would be voluntary although it could be encouraged when users obtain the data. Similarly, enhancing the functionality of the existing library of publications could allow for the recording of a greater number of study level data. For example, users notifying CLS of new publications could be invited to complete a template of meta-data about their publication including, for example, keywords and key variables used in the analysis. This enhanced functionality would assist in future exercises aimed at tracing MCS data usage and would also be beneficial to future researchers to identify where data have been used previously and where they are underutilised. Further guidance or emphasis of the importance of naming of the MCS in publications' titles, abstracts or keywords when users obtain data may facilitate future reviews of data usage, and may give additional prominence to the study in the literature. Finally, most variables included in MCS surveys go through a process of consultation which involves a written case being made for their inclusion. Publishing a record of this case for inclusion for new variables could allow other users to understand why variables have been included. For example, in the case of hobbies data, which are not widely used, publishing this information could allow users to understand the rationale underlying new questions and may stimulate further use of the data.

The mapping exercise showed that a systematic approach to obtaining counts of overall study data usage is feasible. Detailed exploration of individual variable usage for a study as large as the MCS required limits to be placed on the scope. Nevertheless, over 150 unique studies were profiled further and the exercise confirmed the properties of variables that are highly utilised, and those whose usage remained relatively dormant. It also uncovered specific issues (not insurmountable) and incompatibilities between the way in which MCS data were collected and deposited and the wider practices and recommendations of the research community.

The systematic mapping approach exhibited strengths in being able to build a detailed depiction of published variable usage that allowed for the understanding of levels, patterns and results of usage, and facilitators and barriers to variable usage. We would welcome further exploration in terms of how a systematic approach to discovering, mapping and synthesising literature could be integrated with the further analysis of MCS and other longitudinal data. An example might be the investigation of the relationship between BMI and behavioural outcomes. A systematic review could be conducted of studies using MCS data on BMI and child behaviour to synthesise the conceptual frameworks and to help design a model to be tested in the data, with covariates selected based on the results and/or recommendations of previous studies. This synthesised model (based on the synthesis of theory and previous results) could then be tested on MCS data, blending both the systematic review approach and new analysis of the data.

Longitudinal consistency in measures is key to identifying change over time, and the review helped map the degree of

consistency in measures, and their utility. This shaped decisions around inclusion of related variables in MCS7 (age 17 years). Proportions classified as overweight and obese were calculated and deposited at UKDS for the first time using both the UK90 and IOTF thresholds for the age 14 years data, with the review results prompting this decision, and providing an impetus for researchers to consider and report the choice of threshold used. Systematic reviewing techniques are a relatively new, although flourishing, approach to the synthesis of research evidence; by contrast, longitudinal studies, such as the 1958-born cohort, have made significant contributions to the advancement of social and medical sciences for decades. Further intersection of both approaches is likely to lead to substantive and methodological innovations, and the results of the current mapping exercise show one of many potential approaches that could be taken to blending both disciplines.

## Supplementary material

### Supplementary File 1. Search syntax used in the current study.

[Click here to access the data.](#)

## Data availability

All data underlying the results are available as part of the article and no additional source data are required.

## Grant information

We acknowledge funding from the Economic and Social Research Council in completing this study (grant numbers ES/K005987/1 and ES/M001660/1).

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

## Acknowledgments

We would like to acknowledge the valuable comments made by Professor Heather Joshi on an earlier version of this paper.

## References

- Andersberg D, Chevalier A, Wadsworth J: **Anatomy of a health score: education, income and the MMR controversy in the UK.** *J Health Econ.* 2011; **30**(3): 515–530. [PubMed Abstract](#) | [Publisher Full Text](#)
- Bonell C, Fletcher A, McCambridge J: **Improving school ethos may reduce substance misuse and teenage pregnancy.** *BMJ.* 2007; **334**(7594): 614–6. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Borenstein M, Hedges LV, Higgins JPT, et al.: **Introduction to meta-analysis.** New York: John Wiley & Sons. 2011. [Reference Source](#)
- Brophy S, Cooksey R, Gravener MB, et al.: **Risk factors for childhood obesity at age 5: analysis of the millennium cohort study.** *BMC Public Health.* 2009; **9**: 467. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chanfreau J, Lloyd C, Byron C, et al.: **Predicting wellbeing.** 2013. [Reference Source](#)
- Connelly R: **Drivers of unhealthy weight in childhood: analysis of the Millennium Cohort Study.** (Scottish Government social research: public services and government) (9781780453095). Retrieved from Edinburgh, 2011. [Reference Source](#)
- Connelly R, Platt L: **Cohort profile: UK Millennium Cohort Study (MCS).** *Int J Epidemiol.* 2014; **43**(6): 1719–1725. [PubMed Abstract](#) | [Publisher Full Text](#)
- Corden A, Millar J: **Time and change: A review of the qualitative longitudinal research literature for social policy.** *Soc Policy Soc.* 2007; **6**(4): 583–592. [Publisher Full Text](#)
- Delaney L, Doyle O: **Socioeconomic differences in early childhood time preferences.** *J Econ Psychol.* 2012; **33**(1): 237–247. [Publisher Full Text](#)
- Dillenburg K, Jordan JA, McKerr L, et al.: **The Millennium child with autism: early childhood trajectories for health, education and economic wellbeing.** *Dev Neurorehabil.* 2015; **18**(1): 37–46. [PubMed Abstract](#) | [Publisher Full Text](#)
- Flouri E, Midouhas E, Joshi H: **Family poverty and trajectories of children's emotional and behavioural problems: the moderating roles of self-regulation and verbal cognitive ability.** *J Abnorm Child Psychol.* 2014; **42**(6): 1043–1056. [PubMed Abstract](#) | [Publisher Full Text](#)
- Flouri E, Tzavidis N, Kallis C: **Area and family effects on the psychopathology of the Millennium Cohort Study children and their older siblings.** *J Child Psychol Psychiatry.* 2010; **51**(2): 152–161. [PubMed Abstract](#) | [Publisher Full Text](#)
- Gonzalez-Casanova I, Sarmiento OL, Gazmararian JA, et al.: **Comparing three body mass index classification systems to assess overweight and obesity in children and adolescents.** *Rev Panam Salud Publica.* 2013; **33**(5): 349–355. [PubMed Abstract](#)
- Goodman R: **The Strengths and Difficulties Questionnaire: a research note.** *J Child Psychol Psychiatry.* 1997; **38**(5): 581–586. [PubMed Abstract](#) | [Publisher Full Text](#)
- Gough D, Oliver S, Thomas J: **Introducing systematic reviews.** In D. Gough, S. Oliver, & J. Thomas (Eds.), *An Introduction to Systematic Reviews.* London: Sage. 2012. [Reference Source](#)
- Griffiths LJ, Dowda M, Dezateux C, et al.: **Associations between sport and screen-entertainment with mental health problems in 5-year-old children.** *Int J Behav Nutr Phys Act.* 2010; **7**: 30. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hartas D: **Children's language and behavioural, social and emotional difficulties and prosocial behaviour during the toddler years and at school entry.** *British Journal of Special Education.* 2011; **38**(2): 83–91. [Publisher Full Text](#)
- Hartas D: **Children's social behaviour, language and literacy in early years.** *Oxf Rev Educ.* 2012; **38**(3): 357–376. [Publisher Full Text](#)
- Hogan AE, Scott KG, Bauer C: **The Adaptive Social Behaviour Inventory (ASBI): A new assessment of social competence in high risk three-year-olds.** *J Psychoeduc Assess.* 1992; **10**: 230–239. [Publisher Full Text](#)
- Johnson J: **Millennium Cohort Study: Psychological, Developmental and Health Inventories.** Review of. *Centre for Longitudinal Studies.* 2012. [Reference Source](#)
- Johnson W, Li L, Kuh D, et al.: **How Has the Age-Related Process of Overweight or Obesity Development Changed over Time? Co-ordinated Analyses of Individual Participant Data from Five United Kingdom Birth Cohorts.** *PLoS Med.* 2015; **12**(5): e1001828; discussion e1001828. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Jokela M: **Characteristics of the first child predict the parents' probability of having another child.** *Dev Psychol.* 2010; **46**(4): 915–926. [PubMed Abstract](#) | [Publisher Full Text](#)
- Joshi H: **'Non-cognitive' skills: What are they and how can they be measured in the British cohort studies.** Retrieved from London. 2014. [Reference Source](#)

Joshi H, Fitzsimons E: **The Millennium Cohort Study: the making of a multi-purpose resource for social science and policy.** *Longit Life Course Stud.* 2016; **7**(4): 409–430.

[Publisher Full Text](#)

Joshi H, Nasim B, Goodman A: **The Measurement of Social and Emotional Skills and their Association with Academic Attainment in British Cohort Studies.** *Non-cognitive Skills and Factors in Educational Attainment.* Springer. 2016; 239–264.

[Publisher Full Text](#)

Kelly Y, Becares L, Nazroo J: **Associations between maternal experiences of racism and early child health and development: findings from the UK Millennium Cohort Study.** *J Epidemiol Community Health.* 2013; **67**(1): 35–41.

[PubMed Abstract](#) | [Publisher Full Text](#)

Kelly Y, Iacovou M, Quigley MA, *et al.*: **Light drinking versus abstinence in pregnancy - behavioural and cognitive outcomes in 7-year-old children: a longitudinal cohort study.** *BJOG.* 2013; **120**(11): 1340–1347.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Kneale D, Patalay P, Khatwa M, *et al.*: **Piloting and producing a map of Millennium Cohort Study Data usage: Where are data underutilised and where is granularity lost?** Retrieved from London. 2016.

[Reference Source](#)

Krosnick JA: **Survey research.** *Annu Rev Psychol.* 1999; **50**(1): 537–567.

[PubMed Abstract](#) | [Publisher Full Text](#)

Midouhas E, Yogaratnam A, Flouri E, *et al.*: **Psychopathology trajectories of children with autism spectrum disorder: the role of family poverty and parenting.** *J Am Acad Child Adolesc Psychiatry.* 2013; **52**(10): 1057–1065.e1.

[PubMed Abstract](#) | [Publisher Full Text](#)

National Obesity Observatory: **A simple guide to classifying body mass index in children.** Retrieved from London. 2011.

[Reference Source](#)

NHS Choices: **Why 5 A DAY?** 2015.

[Reference Source](#)

Parkes A, Sweeting H, Wight D, *et al.*: **Do television and electronic games predict children's psychosocial adjustment? Longitudinal research using the UK Millennium Cohort Study.** *Arch Dis Child.* 2013; **98**(5): 341–348.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Pearce A, Elliman D, Law C, *et al.*: **Does primary immunisation status predict MMR uptake?** *Arch Dis Child.* 2009; **94**(1): 49–51.

[PubMed Abstract](#) | [Publisher Full Text](#)

Power C, Elliott J: **Cohort profile: 1958 British birth cohort (National Child Development Study).** *Int J Epidemiol.* 2006; **35**(1): 34–41.

[PubMed Abstract](#) | [Publisher Full Text](#)

Presser S, Couper MP, Lessler JT, *et al.*: **Methods for testing and evaluating survey questions.** *Public Opin Q.* 2004; **68**(1): 109–130.

[Publisher Full Text](#)

Pronzato C, Arnstein A: **Marital breakup and children's behavioural responses.** *IDEAS Working Paper Series from RePEc.* Turin: University of Turin. 2013.

[Reference Source](#)

Sammons P, Sylva K, Melhuish E, *et al.*: **Measuring the impact of pre-school on children's social/behavioural development over the pre-school period.** Institute of Education. 2003.

[Reference Source](#)

Schoon I, Bynner J: **Risk and resilience in the life course: implications for interventions and social policies.** *J Youth Stud.* 2003; **6**(1): 21–31.

[Publisher Full Text](#)

Sullivan A, Joshi H: **Child Health.** In K Hansen & H Joshi (Eds.), *Millennium Cohort Study Third Survey: a user's guide to initial findings.* London: Centre for Longitudinal Studies, Institute of Education, University of London. 2008.

[Reference Source](#)

Thomas J, Brunton J, Graziosi S: **EPPI-Reviewer 4.0: software for research synthesis.** Retrieved from London. 2010.

Twamley K, Brunton G, Sutcliffe K, *et al.*: **Fathers' involvement and the impact on family mental health: evidence from Millennium Cohort Study analyses.** *Community, Work & Family.* 2013; **16**(2): 212–224.

[Publisher Full Text](#)

Warden D, Cheyne B, Christie D, *et al.*: **Assessing Children's Perceptions of Prosocial and Antisocial Peer Behaviour.** *Educational Psychology.* 2003; **23**(5): 547–567.

[Publisher Full Text](#)

Zilanawala A, Sacker A, Nazroo J, *et al.*: **Ethnic differences in children's socioemotional difficulties: Findings from the Millennium Cohort Study.** *Soc Sci Med.* 2015; **134**: 95–106.

[PubMed Abstract](#) | [Publisher Full Text](#)

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**