The Millennium Cohort Study

Emla Fitzsimons

SRA Seminar, Cardiff
12th April 2015

www.ioe.ac.uk
Outline

1. Introduction
2. Why longitudinal data?
3. About the Millennium Cohort Study (MCS)
4. Findings from the Age 11 Survey
5. MCS research and policy impact
6. Accessing the data
**UK Millennium Cohort Study (MCS)**

**Longitudinal birth cohort study** of nearly 19,000 children born in ~400 areas of the UK

- *Longitudinal* - follows the same people through life
- *Birth cohort* - follows people sharing the same birth date (year)

- Provides insight into trajectories over the life course (e.g. in weight, poverty, education) for the population as a whole, and for different sub-populations
- Uniquely placed to allow for an understanding of the relationship between early life events and circumstances and outcomes later on in life
Run by the Centre for Longitudinal Studies, based at UCL Institute of Education, London

Oversamples for:
- Wales, Scotland, Northern Ireland
- Areas with high child poverty
- In England, areas with higher ethnic minority populations

Data publicly available to researchers, policy makers (more on access later)
Core Funding:
Economic and Social Research Council (ESRC)

Main Co-Funding:
Govt departments – Dept of Health; Dept for Education; Dept of Work and Pensions; Welsh Government; Scottish Government; NI Executive; Home Office; Ministry of Justice; Dept for Transport
**British Birth Cohort Studies**

**National Survey of Health & Development** – living in GB born in one week in **1946**

**National Child Development Study (NCDS)** – living in GB born in one week in **1958**

**British Birth Cohort Study (BCS)** – living in GB born in one week in **1970**

*MCS the first national birth cohort in 30 years!*

New birth cohort - **Life Study** - started in **2014**
Why longitudinal data?
Why longitudinal data?

- Policy should be underpinned by solid concrete empirical evidence:
  - Requires careful analysis applied to the best available data

- A variety of data sources:
  - Quantitative
    - Observational (e.g. cross-sectional, longitudinal) – does not involve any intervention/experiment
    - Experimental – entails randomisation of subjects to treatment/control groups
  - Qualitative

  Samples can be representative (quantitative); small-scale/purposeful (qualitative; quantitative)

- Longitudinal data (birth cohort, household, etc.) form a key part of ‘set’ of data available in a country for policy analysis
  - UK unique in richness of longitudinal data it collects, starting 1946
Do longitudinal data have policy impact?

Longitudinal studies are not cheap - tracing expensive - important to be very clear as to when/how better than other sources, e.g. repeated cross sections

**Insights from/absolute advantage of longitudinal data:**

1) Movement/mobility (or lack of it) - poverty, learning, height… Trajectories very important for policymakers to know!

2) Short- and long-run impacts of early childhood shocks/adversities

3) Some specific policy interventions (e.g. of policy events, complex interventions)

→ Clear that longitudinal data sets come into their own *over time*
Do longitudinal data have policy impact?

Longitudinal data provide platform for an unplanned world: often policymakers want to do something now, cannot wait for a few years to find out if good or not

- Longitudinal data allow for analysis of patterns: not causal!
  
  → Longitudinal data very useful for highlighting problems
  
  e.g. areas that may require intervention. For instance, if there us catch-up/reversibility in early mental health problems, versus if there is persistence (if persistence, important to know critical times it happens, so policy makers can focus attention)

- Longitudinal data combined with serendipitous events – can help get closer to causality
  
  e.g. breastfeeding and development
Decision to breastfeed is not random – reflects maternal characteristics and preferences

\[ Y=b_1B+b_2X+u \]

⇒ Difficult to ascertain how much of the effect of breastfeeding on outcomes is causal

Unable to randomise some mothers to breastfeed and others not to!

Instead look for a “natural” randomisation…
Do longitudinal data have policy impact?

In this way, can compare outcomes of children breastfed and not breastfed, and interpret differences as *causal*, as mothers are on average “identical”
Getting the most out of longitudinal data

High quality data key: Design and measurement are crucial
- Design studies to best pick up things the data have absolute advantage in
- Need to be clear about important areas to cover and make tough choices
- Use well-validated questions to obtain best possible measurements

Make use of the best analytical skills

Political support important
Main issue is attrition

People cannot be replaced – or lose all the advantages of longitudinal data set

Tracing is key – cohort maintenance team dedicated to this at CLS
Back to MCS…
Objectives of MCS

- To chart initial conditions of the social, economic and health advantages and disadvantages facing new children in the new century and their consequences

- To generate insights that will help to improve the health, development and wellbeing of individuals in future generations

- To compare patterns of development with other cohorts (earlier and later)

- To collect information on previously neglected topics, such as father’s involvement, child care

- To investigate the wider social ecology of the family, including community and services, splicing in geo-coded data
In **England** and **Wales** born between:
  01/09/2000 and 31/08/2001

In **Scotland** and **Northern Ireland** born between:
  24/11/2000 and 10/1/2002
Sample Design

Stage 1: selection of electoral wards (398)

Ethnic

Disadvantaged

Advantaged

Stage 2: selection of children (27201)

Sample of children drawn from child benefit records held by Department of Work and Pensions
MCS Surveys

Surveys timed to take place at key developmental, social and educational stages

• 9 months: 2001/02
• Age 3: 2003/04
• Age 5: 2006
• Age 7: 2008
• Age 11: 2012
• Age 14: 2015
• Age 17: 2018

• Plan to follow through life
• MCS provides a ‘Wales Longitudinal Study’

• Large initial samples including Welsh oversample of ~2700

• Largest fall off between wave 1 and 2 (as common for longitudinal surveys), but subsequently wave on wave response remains high

• Sample in Wales ~ 1800 now, but still suitable for much analysis and power for cross-country comparisons
<table>
<thead>
<tr>
<th>COUNTRY</th>
<th>Number of sampled wards</th>
<th>MCS1 Achieved Sample</th>
<th>MCS2 Achieved Sample</th>
<th>MCS3 Achieved Sample</th>
<th>MCS4 Achieved Sample</th>
<th>MCS5 Achieved Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENGLAND</td>
<td>200</td>
<td>11532</td>
<td>10050</td>
<td>9717</td>
<td>8839</td>
<td>8618</td>
</tr>
<tr>
<td>WALES</td>
<td>73</td>
<td>2761</td>
<td>2261</td>
<td>2181</td>
<td>2018</td>
<td>1881</td>
</tr>
<tr>
<td>SCOTLAND</td>
<td>62</td>
<td>2336</td>
<td>1814</td>
<td>1814</td>
<td>1628</td>
<td>1480</td>
</tr>
<tr>
<td>N. IRELAND</td>
<td>62</td>
<td>1923</td>
<td>1465</td>
<td>1534</td>
<td>1372</td>
<td>1308</td>
</tr>
<tr>
<td>TOTAL UK</td>
<td>398</td>
<td>18552</td>
<td>15590</td>
<td>15246</td>
<td>13857</td>
<td>13287</td>
</tr>
<tr>
<td>Data Collection Method</td>
<td>9m</td>
<td>Age 3</td>
<td>Age 5</td>
<td>Age 7</td>
<td>Age 11</td>
<td>Age 14</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td>----</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Interview (and self-completion with both resident parents)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cognitive assessments</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Physical measurements</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Child self-completion</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Older siblings</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interviewer observations</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Teacher survey</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Consent to data linkage</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Saliva samples</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Time use record</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Activity monitors</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
Welsh translations

Main and partner interviews
  Interviewer translation to Welsh

Young person questionnaire
  Interviewer translation to Welsh (going through the interviewer-administered script to avoid the most sensitive questions)

Cognitive assessments
  Translation to Welsh (pre-age 14); English only (age 14)

Survey materials translated to Welsh
  e.g. parent advance letter/booklet/appointment leaflet; YP advance letter/booklet/appointment leaflet; consent forms; showcards; thank you letters/postcards; further information leaflet
Content of MCS Parent Interviews

- Household questionnaire  (Either Parent)
- Family Context  (Main)
- Early Education, schooling & childcare  (Main)
- Child & Family Activities & child behaviour  (Main)
- Parenting Activities  (Main & Partner)
- Child Health  (Main)
- Parent’s Health  (Main & Partner)
- Employment, Education and Income  (Main & Partner)
- Housing & Local Area  (Main)
- Social and Community Context  (Main & Partner)
- Self-completion  (Main & Partner)
Data collected directly from cohort member
1. Cognitive assessments since age 3

2. Physical measurements since age 3

3. Self-completion questionnaires since age 7
1. Cognitive assessments since age 3
## Child cognitive assessments

<table>
<thead>
<tr>
<th>Age 3</th>
<th>Age 5</th>
<th>Age 7</th>
<th>Age 11</th>
<th>Age 14</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAS Naming Vocabulary</td>
<td>The story of Sally &amp; Ann</td>
<td>The story of Sally &amp; Ann</td>
<td>BAS Verbal Similarities</td>
<td>Vocabulary knowledge</td>
</tr>
<tr>
<td>Bracken School Readiness Assessment:</td>
<td>BAS Naming vocabulary</td>
<td>Word Reading</td>
<td>CANTAB Spatial Working Memory</td>
<td></td>
</tr>
<tr>
<td>Colours Letters Sizes Comparisons</td>
<td>BAS Picture Similarities</td>
<td>Progress in Maths</td>
<td>CANTAB Cambridge Gambling Task</td>
<td>CANTAB Cambridge Gambling Task</td>
</tr>
<tr>
<td></td>
<td>BAS Pattern Construction</td>
<td>BAS Pattern Construction</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1. Cognitive assessments since age 3

2. Physical measurements since age 3
# Child physical measurements (& samples)

<table>
<thead>
<tr>
<th>Age 3</th>
<th>Age 5</th>
<th>Age 7</th>
<th>Age 11</th>
<th>Age 14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height</td>
<td>Height</td>
<td>Height</td>
<td>Height</td>
<td>Height</td>
</tr>
<tr>
<td>Weight</td>
<td>Weight</td>
<td>Weight</td>
<td>Weight</td>
<td>Weight</td>
</tr>
<tr>
<td>Oral fluid sample</td>
<td>Waist circumference</td>
<td>Waist circumference</td>
<td>Body fat</td>
<td>Body fat</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Body fat</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Activity monitor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Shed milk teeth</td>
<td></td>
</tr>
</tbody>
</table>
Data from cohort members

1. Cognitive assessments since age 3
2. Physical measurements since age 3
3. Self-completion questionnaires since age 7
Age 11 child self-completion

28 page paper self-completion, covering:

- Family & Friends
- School
- Things done outside school
- The area they live in
- Things they may have done – anti social behaviour
- Things they may have tried – incl. alcohol & smoking
- How they feel
- Growing up
Age 11 child self-completion

1. How often do you listen to or play music, not at school?

2. How often do you draw, paint or make things, not at school?

3. How often do you play sports or active games inside or outside, not at school?
40 min computer self-completion, covering:

- Things that they do
- Their views
- School and their future
- About them
- Their family
- Their friends
- Relationships
- Risky behaviours:
  - Things they may have tried
  - Things they may have experienced
  - Things they may have done
- Their health
- Their body
- How they feel
MCS survey data can be enriched by administrative data: consents for linkage to admin data asked at various sweeps

• Linked **maternity and birth records** data

• **Hospital records** (cohort members, parents, siblings) – Welsh data nearly ready for deposit; application in process covering other 3 countries of UK; covers parents and siblings also

• **Education records** – National Pupil Database, Key Stage 1 (England and Wales); further education linkages in process

• Consent gained for linkage to **parental economic records** (DWP and HMRC)
A look ahead

- The next sweep will be at age 17 → consultation starting very soon

- Timing of future surveys after that remains open to discussion

- Potential for research and policy impact will only grow as new sweeps collected

- The childhood sweeps will lay an incredibly detailed context for understanding long term trajectories and outcomes, as well as continuing to provide new evidence about the childhood and adolescent years themselves
MCS5 Initial Findings
Inequalities in cognitive development persist for the millennium generation

- Parents’ qualifications remain the most powerful predictor of cognitive development across the board.
- Indian children scored highest on all assessments, with White children also performing well. Pakistani and Bangladeshi 11-year-olds had the lowest overall performance.
- At age 11, boys are outperforming girls in verbal ability for the first time.
One in five children of the new century are obese

Significant variation in overweight/obesity by country:
  40% in Wales and NI
  35% in England
  33% in Scotland

→ Clear links between weight and parental level of education

→ Strong associations with mother’s weight
Poverty has touched more than half of the UK’s millennium generation

- **More than half** of the MCS children were living in poor families in at least one of the five MCS surveys to date.

- **Around 17 per cent are persistently poor**, that is they have been poor at four or five MCS surveys.
  - Children who are persistently poor also face **more material deprivation** and have **lower subjective wellbeing**.

- Children living in **Wales and Northern Ireland** were more likely to be persistently poor than those living in England and Scotland (21% and 19% compared to 16% and 13%).
Family structure

Family instability affects 4 in 10 children of the millennium generation

- Nearly **four in ten children** had experienced some form of family disruption by age 11.
- The number of families that included both natural parents decreased from 85% at age 9 months to 61% at age 11.
- By age 11, 50% of children were living with two natural parents who were married.
- Families that have always included two natural parents are more likely to have older mothers and more educated parents. Children in these families are the least likely to experience poverty at age 11.
- Children from two-natural-parent families are the least likely to have behavioural problems, judging by parent reports. Such problems are more common for children in lone parent and step-parent families.
21st century childhoods may be very different but they still seem largely enjoyable

- **Poverty:** 1 in 6 MCS children (17%) were persistently poor during their first 11 years. Children in such families were less happy with their lives than other cohort members.

- **Divorce and separation:** 4 in 10 11-year-olds experienced some family disruption by age 11 compared with only 1 in 10 in 1969. Nevertheless, three quarters of children were completely happy with their families.

- **School:** Although most MCS children were facing the stresses of primary to secondary school transition, only 6% said they did not like school.

- **Safety:** Almost 9 in 10 felt their home area was ‘safe’ or ‘very safe’.

- **Bullying:** Just over half of the 11-year-olds said they were hurt or picked on by other children.

- **Drinking/smoking:** Only 3% of 11-year-olds had smoked a cigarette; 13% had tried an alcoholic drink.
Evidence from ESRC report on impact of the MCS

(1) **Childcare**

-The last Labour government improved Sure Start Children’s Centres in the most disadvantaged parts of England and funded more outreach workers and training “as a result of research using MCS data”

-Coalition government has adopted a different strategy regarding children’s centres but David Willetts has indicated that the MCS is a “very valuable research tool” and is continuing to inform political thinking on this issue
Evidence from ESRC report on impact of the MCS

(2) Breastfeeding

- Study based on the MCS data 9 months showed that breastfeeding protects against hospitalisation for diarrhoea and lower respiratory tract infection (Quigley et al, 2006, Pediatrics).

- Other studies show strong links with later cognitive development

→ Research has influenced National Institute for Health and Clinical Excellence (NICE) guidance on Maternal and Child Nutrition Guidance issued by the Department of Health/Department for Children, Schools and Families, ‘Commissioning local breastfeeding support services’
Irregular bedtimes may affect children's brains (Guardian, July 2013)

New Study Shows Why You Should Get the Kids to Bed on Time (Wall Street Journal, July 2013)

Not having a regular bedtime could affect children’s behaviour by disrupting circadian rhythms, and by the harm that sleep deprivation causes to the developing brain...

As the study shows the effects of inconsistent bedtimes are reversible, health care providers could check for sleep disruptions as part of routine health care visits.
• The couch potato seven-year-olds who sit for half the day: Only half exercise for a minimum 10 minutes a day with girls the least active (Daily Mail, Aug 2013)

• Doctors sound alarm on child fitness and health (Guardian, Aug 2013)

• Summer-born pupils ‘should have exam scores boosted’ (BBC News, May 2013)
Accessing the MCS Data
Resources available via CLS website: www.cls.ioe.ac.uk

- Questionnaires and CAPI Documentation
- Technical Reports – e.g. on sampling, instrument development and fieldwork procedures
- Data Dictionaries
- Cohort Studies Data Notes
- Working Papers
Data can be accessed via the UK data service:

http://ukdataservice.ac.uk

Datasets come under three different licence regimes:

1. **End User Licence (EUL)** - standard access

2. **Special Access Licence** - data collections are anonymised but contain more detailed information than End User Licence (EUL) data

3. **Secure Access Licence** - deemed too confidential/sensitive to be released via download; not possible to access for non-UK based researchers
Datasets with End User Licence (EUL) conditions attached are accessible and downloadable once users are registered.

Registration for non-UK based researchers has the following steps:

1. Apply for a username and password. This can be done on this page: http://www.data-archive.ac.uk/sign-up/credentials-application

2. Once researchers receive their username and password, will need to complete an online registration form after logging in

3. In process of downloading the data, they will also be asked to register their project online (30 words).

→ All v quick and straight forward!
How to access data

Some data, such as low level geography, hospital of birth are only available via special licence, which imposes certain restrictions on the handling and usage of the data.

To access these data, you need to complete a Special Licence Access form.

Process to apply is explained here:

http://ukdataservice.ac.uk/get-data/how-to-access/downloadorder/ordersl.aspx

→ Process is not difficult, but more time-consuming than end user data.

Once downloaded, the researcher can use the data in the same way as the EUL data.
Some data, such as administrative data on educational outcomes and geocodes, only available via Secure Access - restricted to UK researchers.

Data are not downloadable and can only be accessed via remote desktop.