Getting Started: An introduction to four British cohort studies

15 November 2023

Centre for Longitudinal Studies, UCL Social Research Institute
Housekeeping

- Please keep your cameras off and mics muted at all times
- If you have a question, please use the chat function, and please note your question will be visible to all attendees
- Technical issues – please email us: ioe.clsevents@ucl.ac.uk
- We would be grateful for your feedback. Please follow the link in the chat at the end of the event for the short survey

Thank you for joining us today
<table>
<thead>
<tr>
<th>Session</th>
<th>Time</th>
<th>Topics covered</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>13.00 – 13.15</td>
<td>Introduction</td>
<td>Dr Vanessa Moulton</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Senior Research Associate</td>
</tr>
<tr>
<td>2.</td>
<td>13.15 – 13.35</td>
<td>Content by ‘subject area’</td>
<td>Prof. Morag Henderson</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Professor in Sociology</td>
</tr>
<tr>
<td>3.</td>
<td>13.35 – 13.55</td>
<td>Overview of the type of analysis</td>
<td>Dr Richard Silverwood</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Associate Professor</td>
</tr>
<tr>
<td>4.</td>
<td>13.55 – 14.20</td>
<td>Getting started with the data</td>
<td>Dr Vanessa Moulton</td>
</tr>
<tr>
<td></td>
<td></td>
<td>And where to go for more information</td>
<td>Senior Research Associate</td>
</tr>
<tr>
<td>5.</td>
<td>14.20 – 14.30</td>
<td>General Q&amp;A</td>
<td>All</td>
</tr>
</tbody>
</table>
Birth cohort studies

- Cohort studies are a type of *longitudinal study*—an approach that follows participants over a period of time (often many years).
- Participants share a common characteristic, i.e. birth cohorts follow individuals born in a particular period - a day, week, month, year.
- It follows these people throughout their lives, and collects information from them at particular ages.
- During the period of follow-up, some of the cohort will be exposed to a specific risk factor or characteristic; by measuring outcomes over a period of time, it is then possible to explore the impact of this variable.
The life course approach – life stages

- Preconception and pregnancy
- Infancy and early years (0 to 5)
- Childhood and adolescence (5 to 24)
- Working age and adults (16 to 64)
- Older people
Our new studies

**Early Life Cohort Feasibility Study** (fieldwork underway in 2023)
- ESRC funded study following a cohort of several thousand babies born in the UK in 2022.
- Will provide vital new insights into the health and development of children + test feasibility of setting up a new full-scale birth cohort study in future.
- Expected data release: late 2024.

**Children of the 2020s Study** (fieldwork began in 2022)
- Department for Education commissioned study following a cohort of babies born in England between September and November 2021 (around 8,500 families).
- Will answer important scientific and policy questions about family, early education and childcare determinants of early school success.
- Wave 1 (nine months old) completed and Wave 2 (two years old) underway.

**COVID Social Mobility and Opportunities study** (fieldwork began in 2021)
- A UKRI funded study following over 13,000 young people (in Year 11 in the academic year 2020-21).
- Is providing new information about the effects of COVID-19 pandemic and the cost of living crisis on young people’s lives and prospects.
- Wave 1 and 2 data available now.
Four national longitudinal studies

1958 National Child Development Study (NCDS)

- born in GB in one week. N = 17,415

1970 British Cohort Study (BCS70)

- born in GB in one week. N = 17,196

Next Steps (formerly LSYPE)


Millennium Cohort Study (MCS)

- born in the UK in 2000-02. N = 18,818
An example: **NCDS** A study of everyone born in one week in 1958 (GB)

<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>
</tr>
</thead>
<tbody>
<tr>
<td>Birth</td>
<td>7</td>
<td>11</td>
<td>16</td>
<td>23</td>
<td>33</td>
<td>42</td>
<td>44</td>
<td>46</td>
<td>50</td>
<td>55</td>
<td></td>
</tr>
</tbody>
</table>

- **Main respondent:**
  - 1958: mother
  - 2008, 2013: cohort member

- **Others:**
  - 1958: school
  - 2013: children (1 in 3)

- **Medical exam:**
  - 1958: medical exam
  - 2000: Ht/Wt
  - 2003: Ht/Wt
  - 2004: Ht/Wt
  - 2008, 2013: Ht/Wt

- **Cognitive mental h.**
  - 1958: cognitive mental h.
  - 2008, 2013: mental h.

- **Linked data:**
  - 1958: area of residence (census)

- **Response rate:**
  - 1958: 17,415
  - 1965: 15,425
  - 1969: 15,337
  - 1974: 14,654
  - 1981: 12,537
  - 1991: 11,469
  - 2000: 11,419
  - 2003: 9,377
  - 2004: 9,534
  - 2008: 9,790
  - 2013: 9,137
Other data enhancements in the birth cohorts

- Genetic data in the NCDS, BCS70 and MCS
  - MCS - Trios (cohort, mother, father)

- Linked administrative data
  - Health and education
  - Consent: Employment (all) and crime (Next Steps, MCS)

- Geographical data
  - e.g. electoral wards, output areas, Points of Interest etc

- Harmonised datasets across the cohorts
  - Socio-economic, BMI, mental health, child environment

- COVID-19 online surveys
  - Possible impacts of pandemic on multiple aspects of life
  - Wide range of topics including family, employment, home schooling, mental health during lockdown and an open question on affects of the pandemic
Thank you
The cohort studies by scientific theme/discipline
Subject areas

- Physical health
- Mental health & Wellbeing
- Family and relationships
- Earnings and income
- Education, ability and cognitive measures

...but there are many more
## Typical information covered

<table>
<thead>
<tr>
<th>Birth</th>
<th>School years</th>
<th>Adult</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household composition</td>
<td>Household composition</td>
<td>Household composition</td>
</tr>
<tr>
<td>Parental social class</td>
<td>Parental social class</td>
<td>Employment</td>
</tr>
<tr>
<td>Obstetric history</td>
<td>Parental employment</td>
<td>Social class</td>
</tr>
<tr>
<td>Smoking in pregnancy</td>
<td>Financial circumstances</td>
<td>Income</td>
</tr>
<tr>
<td>Pregnancy (problems,</td>
<td>Housing</td>
<td>Housing</td>
</tr>
<tr>
<td>antenatal care)</td>
<td>Health</td>
<td>Health (including biomarkers)</td>
</tr>
<tr>
<td>Labour (length, pain</td>
<td>Cognitive tests</td>
<td>Well-being and mental health</td>
</tr>
<tr>
<td>relief, problems)</td>
<td>Emotions and behaviour</td>
<td>Health-related behaviour</td>
</tr>
<tr>
<td>Birthweight, length</td>
<td>School</td>
<td>Training and qualifications</td>
</tr>
<tr>
<td></td>
<td>Views and expectations</td>
<td>Basic skills</td>
</tr>
<tr>
<td></td>
<td>Attainment</td>
<td>Cognitive tests</td>
</tr>
</tbody>
</table>

(Centre for Longitudinal Studies)
Physical Health
<table>
<thead>
<tr>
<th>Physical health measures</th>
<th>NCDS 58</th>
<th>BCS 70</th>
<th>NS 89</th>
<th>MCS 01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self assessed general health</td>
<td>7, 11, 16, 33, 44, 46, 50, 55</td>
<td>5, 10, 16, 34, 42, 46</td>
<td>25</td>
<td>3, 5, 7, 11, 14, 17</td>
</tr>
<tr>
<td>BMI, Height, Weight</td>
<td>7, 11, 16, 23, 33, 42, 44, 45, 50, 55</td>
<td>10, 16, 26, 30, 34, 42, 46</td>
<td>25</td>
<td>3, 5, 7, 11, 14, 17</td>
</tr>
<tr>
<td>DNA / biomarkers</td>
<td>44 (available)</td>
<td>46 (genotyping)</td>
<td>Planned for age 32</td>
<td>14 (available)</td>
</tr>
<tr>
<td>Physical activity (leisure time)</td>
<td>11, 16, 23, 33, 42, 44, 50, 55</td>
<td>5, 10, 16, 34, 42, 46</td>
<td>20, 25</td>
<td>5, 7, 11, 14, 17</td>
</tr>
<tr>
<td>Diet related measures (intake, overeating)</td>
<td>7, 33, 42, 44</td>
<td>10, 16, 30, 34, 42, 46</td>
<td>25</td>
<td>9 months, 3, 7, 11, 14, 17</td>
</tr>
<tr>
<td>Anthropometry (e.g. blood pressure, body fat, grip strength, vision, motor skills)</td>
<td>7, 11, 16, 44</td>
<td>10, 16, 46 + accelerometry</td>
<td>-</td>
<td>3, 7, 11, 14, 17 (10, 14 acceler)</td>
</tr>
<tr>
<td>Medical conditions/ *long term illness</td>
<td>0, 7, 11, 26, 23, 33, 42, 44, 46, 50, 55</td>
<td>0, 5, 10, 16, 26, 30, 34, 38, 42, 46</td>
<td>14*, 15*, 16*, 17*, 18*, 19*, 20*, 25*</td>
<td>9m, 3, 5, 7, 11, 14, 17</td>
</tr>
<tr>
<td>Drugs &amp; alcohol consumption</td>
<td>16, 23, 33, 42, 44, 46, 50, 55</td>
<td>16, 26, 30, 34, 42, 46</td>
<td>14, 15, 16, 17, 18, 19, 20, 25</td>
<td>11, 14, 17</td>
</tr>
<tr>
<td>Physical health measures</td>
<td>NCDS 58</td>
<td>BCS 70</td>
<td>NS 89</td>
<td>MCS 01</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>----------------------------------</td>
<td>---------------------------------</td>
<td>--------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Self assessed general health</td>
<td>7, 11, 16, 33, 44, 46, 50, 55</td>
<td>5, 10, 16, 34, 42, 46</td>
<td>25</td>
<td>3, 5, 7, 11, 14, 17</td>
</tr>
<tr>
<td>BMI, Height, Weight</td>
<td>7, 11, 16, 23, 33, 42, 44, 50, 55</td>
<td>10, 16, 26, 30, 34, 42, 46</td>
<td>25</td>
<td>3, 5, 7, 11, 14, 17</td>
</tr>
<tr>
<td>DNA /biomarkers</td>
<td>44 (available)</td>
<td>46 (genotyping)</td>
<td>Planned for age 32</td>
<td>14 (available)</td>
</tr>
<tr>
<td>Physical activity (leisure time)</td>
<td>11, 16, 23, 33, 42, 44, 50, 55</td>
<td>5, 10, 16, 34, 42, 46</td>
<td>20, 25</td>
<td>5, 7, 11, 14, 17</td>
</tr>
<tr>
<td>Diet related measures (intake, overeating)</td>
<td>7, 33, 42, 44</td>
<td>10, 16, 30, 34, 42, 46</td>
<td>25, 32</td>
<td>9 months, 3, 7, 11, 14, 17</td>
</tr>
<tr>
<td>Anthropometry (e.g. blood pressure, body fat, grip strength, vision, motor skills)</td>
<td>7, 11, 16, 44</td>
<td>10, 16, 46 + accelerometry</td>
<td>-</td>
<td>3, 7, 11, 14, 17 (7, 14 acceler)</td>
</tr>
<tr>
<td>Medical conditions/ *long term illness</td>
<td>0, 7, 11, 26, 33, 42, 44, 50, 55</td>
<td>0, 5, 10, 16, 26, 30, 34, 38, 42, 46</td>
<td>14*, 15*, 16*, 17*, 18*, 19*, 20*, 25*</td>
<td>9m, 3, 5, 7, 11, 14, 17</td>
</tr>
<tr>
<td>Drugs &amp; alcohol consumption</td>
<td>16, 23, 33, 42, 44, 46, 50, 55</td>
<td>16, 26, 30, 34, 42, 46</td>
<td>14, 15, 16, 17, 18, 19, 20, 25</td>
<td>11, 14, 17</td>
</tr>
<tr>
<td>Health measures in COVID-19 web surveys</td>
<td>NCDS 58</td>
<td>BCS 70</td>
<td>NS 89</td>
<td>MCS 01</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>---------</td>
<td>--------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>COVID-19 antibodies N=10,442</td>
<td>n=3,222</td>
<td>n=2,547</td>
<td>n=1,267</td>
<td>CM n=1,140 Par n=2,266</td>
</tr>
<tr>
<td>Long COVID Symptoms Testing COVID presence</td>
<td>W3 W1, W2; W1, W2; W1, W2, W3</td>
<td>W3 W1, W2; W1, W2; W1, W2, W3</td>
<td>W3 W1, W2; W1, W2; W1, W2, W3</td>
<td>W3 W1, W2; W1, W2; W1, W2, W3</td>
</tr>
<tr>
<td>Self-related general health</td>
<td>W1, W2, W3</td>
<td>W1, W2, W3</td>
<td>W1, W2, W3</td>
<td>W1, W2, W3</td>
</tr>
<tr>
<td>Long-standing health conditions</td>
<td>W1, W2, W3</td>
<td>W1, W2, W3</td>
<td>W1, W2, W3</td>
<td>W1, W2, W3</td>
</tr>
<tr>
<td>Disruption to medical appointments</td>
<td>W1, W2, W3</td>
<td>W1, W2, W3</td>
<td>W1, W2, W3</td>
<td>W1, W2, W3</td>
</tr>
<tr>
<td>Difficulty obtaining medication</td>
<td>W2, W3</td>
<td>W2, W3</td>
<td>W2, W3</td>
<td>W2, W3</td>
</tr>
<tr>
<td>Defined as vulnerable</td>
<td>W1, W2, W3</td>
<td>W1, W2, W3</td>
<td>W1, W2, W3</td>
<td>W1, W2, W3</td>
</tr>
</tbody>
</table>
Mental Health & Wellbeing
Mental health & Wellbeing— all cohorts

- Bristol Social Adjustment Guide (BSAG)
- Conners teachers Hyperactivity Rating Scale (Conn)
- Rutter Behavioural Scale (RUT)
- Child Development Scale (combination of Rutter and Connor) (CDS)
- Strengths and difficulty questionnaire (SDQ)
- Mood and feelings questionnaire (MFQ)
- Malaise inventory (MAL)
- Kessler Scale (4 item) (K4)
- General Health Questionnaire (12-item version) (GHQ-12)
- Short Form Health Survey (SF-36)
- The Warwick-Edinburgh Mental Wellbeing Scale (WEMWEBS)
Malaise

A nine-item Malaise Inventory: a measure of psychological distress

- Do you feel tired most of the time?
- Do you often feel miserable or depressed?
- Do you often get worried about things?
- Do you often get into a violent rage?
- Do you often suddenly become scared for no reason?
- Are you easily upset or irritated?
- Are you constantly keyed up and jittery?
- Does every little thing get on your nerves?
- Does your heart often race like mad?
<table>
<thead>
<tr>
<th>Mental Health measure</th>
<th>NCDS 58</th>
<th>BCS 70</th>
<th>NS 89</th>
<th>MCS 01</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSAG</td>
<td>7, 11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conn</td>
<td></td>
<td>10, 16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RUT</td>
<td>7, 11, 16, 16</td>
<td>5, 10, 16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDS</td>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQ</td>
<td></td>
<td></td>
<td>3, 5, 7, 7, 11, 11, 14, 17, 17</td>
<td></td>
</tr>
<tr>
<td>MFQ</td>
<td></td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAL</td>
<td>23, 33, 42, 50</td>
<td>16, 26, 30, 34, 42, 46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K4/K6</td>
<td></td>
<td>34</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>GHQ-12</td>
<td>42</td>
<td>16, 30</td>
<td>15, 17, 25</td>
<td></td>
</tr>
<tr>
<td>SF-36</td>
<td>50</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEMWEBS</td>
<td>50</td>
<td>42, 46</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>

Parent, teacher and self-report
COVID-19 sweeps: All Cohorts

- **Patient Health Questionnaire-2 (PHQ-2):**
  - Over the last 2 weeks, how often have you been bothered by the following problems?
    - Little interest or pleasure in doing things
    - Feeling down, depressed or hopeless

- **Generalised-Anxiety Disorder (GAD-2):**
  - Over the last 2 weeks, how often have you been bothered by the following problems?
    - Feeling nervous, anxious or on edge
    - Not being able to stop or control worrying
  
  Not at all – Nearly every day
Family and Relationships
Family and relationships

- Who is in the household
- Relationship to cohort member
- Age /number of siblings
- Biological, step, adoptive parents
- Age of parents when the cohort member was born
- Fertility intentions
- Pregnancy history
- Partnership formation, cohabitation, marriage, divorce, dissolution, formation
Education, ability and cognitive measures
<table>
<thead>
<tr>
<th>Education measure</th>
<th>NCDS 58</th>
<th>BCS 70</th>
<th>NS 89</th>
<th>MCS 01</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>School</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Key stage 1</td>
<td></td>
<td></td>
<td></td>
<td>NPD</td>
</tr>
<tr>
<td>Key stage 2</td>
<td></td>
<td></td>
<td>NPD</td>
<td>NPD</td>
</tr>
<tr>
<td>Key stage 3</td>
<td></td>
<td></td>
<td>NPD</td>
<td></td>
</tr>
<tr>
<td><strong>GCSE or equivalent (subject and grades) (KS4)</strong></td>
<td>23, 42</td>
<td>32</td>
<td>NPD &amp; self report</td>
<td>NPD &amp; self report</td>
</tr>
<tr>
<td><strong>A level or equivalent (KS5)</strong></td>
<td>23, 42</td>
<td>26, 32, 34, 42</td>
<td>NPD &amp; self report</td>
<td>NPD &amp; self report</td>
</tr>
<tr>
<td><strong>Study intentions</strong></td>
<td>16</td>
<td>16</td>
<td>14, 15, 16, 17</td>
<td>11, 14</td>
</tr>
<tr>
<td><strong>Further education</strong></td>
<td></td>
<td></td>
<td>Individualised Learning Record</td>
<td></td>
</tr>
<tr>
<td><strong>Higher education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree subject</td>
<td>42, 46</td>
<td>38, 42</td>
<td>20, 25</td>
<td>17</td>
</tr>
<tr>
<td>University type</td>
<td>23, 33</td>
<td>42</td>
<td>20, 25</td>
<td>17</td>
</tr>
<tr>
<td>Degree grade</td>
<td>42</td>
<td>38</td>
<td>20, 25</td>
<td>-</td>
</tr>
</tbody>
</table>
Cognitive ability in childhood

- Bracken school readiness (BSRA-R)
- British Ability Scales: verbal similarities, word definitions, matrices, recall of digits, pattern construction, picture similarities, naming vocabulary, word reading
- General Ability Test (GAT)
- Cambridge Neuropsychological Test Automated Battery (CANTAB): Decision making, Working memory
- National Foundation for Education Research (NFER): maths tests, reading comprehension
- Applied Psychology Unit (APU): Vocab test, Maths test
- Number Analogies (GL Assessment)
- Schonell Reading Test; Southgate Group Reading Test; Edinburgh Reading Test; English Picture Vocabulary Test
- Copying Designs Test; Human Figure Drawing; Complete a Profile Test
<table>
<thead>
<tr>
<th>(Main) cognitive ability/skill</th>
<th>NCDS 58</th>
<th>BCS 70</th>
<th>NS 89</th>
<th>MCS 01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developmental milestones</td>
<td></td>
<td>22 months*, 42 months*</td>
<td></td>
<td>9 months</td>
</tr>
<tr>
<td>School readiness (BSRA-R)</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Verbal reasoning</td>
<td>11</td>
<td>10</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>Non-verbal reasoning</td>
<td>11</td>
<td>10,16</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Verbal skills (i.e. reading, comprehension, vocabulary, literacy)</td>
<td>7,11,16,37*</td>
<td>5,10,16,21*,34,42</td>
<td>3,5,7,14</td>
<td></td>
</tr>
<tr>
<td>Mathematics and numeracy</td>
<td>7,11,16,37*</td>
<td>10,16,21*,34</td>
<td></td>
<td>7,17</td>
</tr>
<tr>
<td>Visual/spatial processing</td>
<td>7</td>
<td>5</td>
<td></td>
<td>5,7</td>
</tr>
<tr>
<td>Decision making</td>
<td></td>
<td></td>
<td>11,14</td>
<td></td>
</tr>
<tr>
<td>Memory (short-term, long-term, spatial working)</td>
<td>50</td>
<td>10,46</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Processing speed</td>
<td>50</td>
<td>46</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* sub-sample
Earnings and income
<table>
<thead>
<tr>
<th>Earnings and Income</th>
<th>NCDS 58</th>
<th>BCS 70</th>
<th>NS 89</th>
<th>MCS 01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings from work (CM and parents)</td>
<td>7, 11, 16, 23, 33, 42, 46, 50, 55</td>
<td>5, 10, 16, 26, 30, 34, 38, 42, 46</td>
<td>14, 15, 16, 20, 25</td>
<td>3, 7, 11, 14, 17</td>
</tr>
<tr>
<td>Income (investments, income support, benefits, etc.)</td>
<td>16, 33, 42, 46, 50, 55</td>
<td>10, 16, 30, 34, 38, 42</td>
<td>25</td>
<td>3, 7, 11, 14, 17</td>
</tr>
<tr>
<td>Occupation</td>
<td>11, 33, 42, 46, 50, 55</td>
<td>10, 30, 34, 38, 42, 46</td>
<td>25</td>
<td>3, 7, 11, 14, 17</td>
</tr>
<tr>
<td>Social mobility (generational analysis)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Wealth (actual):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Housing</td>
<td>55</td>
<td>42</td>
<td>-</td>
<td>11,14</td>
</tr>
<tr>
<td>Financial</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Savings</td>
<td>23,33,50</td>
<td>34, 42, 46</td>
<td>-</td>
<td>7,11,14</td>
</tr>
<tr>
<td>Debt</td>
<td>-</td>
<td>42, 46</td>
<td>25</td>
<td>7,11,14</td>
</tr>
</tbody>
</table>
How to search these resources
SEARCH Closer Discovery

**Item types:** All  
**Query:** life satisfaction  
**Search within:** 1970 British Cohort Study

Results 1 to 8 of 8 (0.02 seconds)

1. **COVID, Social, and Mental health — Life satisfaction**

2. B960667
   - **Satisfaction** about how life has turned out (10)
     - Please tick the box with the number above it which shows how dissatisfied or satisfied you are about the way your life has turned out so far.
Alternative methods for searching

- Or questionnaires (UK Data Service or CLS website: Our Studies; Sweeps; Documentation; Questionnaires)
- Or download the actual datasets and search the variables (UK Data Service)
- Descriptions of variables in published papers
More detailed picture of the contents by scientific theme & events:

https://cls.ucl.ac.uk/data-access-training/training-and-support-2/

Mental health in four British cohort studies: measurement, research and access

(2023, 44 minutes)

This short webinar explores the wide-ranging opportunities for mental health research using British cohort studies.

Events

Register for upcoming events, discover previous events and download materials.

On this page you can find out what events we have coming up and book your place.

• To hear about future training, please sign up to our events mailing list.

Past webinars and training events

• Selected webinars including our introductions to all four cohort studies can be found on the Training and support page.
• You can also find our COVID-19 online training sessions on the COVID-19 survey page.
Additional sources of information

Physical health


Mental Health and Wellbeing


Cognitive ability

Examples of the types of analyses that can be undertaken using CLS cohort data

Richard Silverwood
Associate Professor of Statistics &
CLS Chief Statistician
Outline

1. Simple analyses
2. Confounder control
3. Repeated measures
4. Cross-cohort analysis
Simple analyses
Simple analyses

• The cohorts provide rich data collected on cohort members over many years/decades, so complex analyses possible.
• But let’s start with some simple examples…
Prevalence of COVID-19, symptoms and testing in the U.K.

Initial findings from the COVID-19 Survey in Five National Longitudinal Studies

By Dylan M. Williams, Gabriella Conti, Nishi Chaturvedi, Alun Hughes, George B. Ploubidis and Richard J. Silverwood
Simple analyses: Examples

What is the prevalence of test-confirmed COVID-19 in each cohort?

COVIDTEST Have you been tested for Coronavirus?
- Yes
- No

COVIDRESULT What was the result of your coronavirus test? If you had more than one test please report the findings of the latest test.
- Positive - it showed I had coronavirus
- Negative - it showed I did not have coronavirus
- Inconclusive
- Waiting for results

Table 1 – Percentages of respondents that had tested positive for SARS-CoV-2 by cohort

<table>
<thead>
<tr>
<th>Cohort</th>
<th>n</th>
<th>N</th>
<th>Percent</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS</td>
<td>8</td>
<td>2609</td>
<td>0.3</td>
<td>0.1, 0.8</td>
</tr>
<tr>
<td>NS</td>
<td>10</td>
<td>1876</td>
<td>0.6</td>
<td>0.3, 1.1</td>
</tr>
<tr>
<td>BCS70</td>
<td>28</td>
<td>4132</td>
<td>0.7</td>
<td>0.5, 1.0</td>
</tr>
<tr>
<td>NCDS</td>
<td>15</td>
<td>5119</td>
<td>0.3</td>
<td>0.2, 0.5</td>
</tr>
<tr>
<td>NSHD</td>
<td>1</td>
<td>1170</td>
<td>0.1</td>
<td>0.0, 0.5</td>
</tr>
</tbody>
</table>

n – number reporting a positive test; N – total sample size
Simple analyses: Examples

What is the prevalence of self-reported COVID-19 in each cohort?

COVID19 Do you think that you have or have had Coronavirus?

- Yes, confirmed by a positive test
- Yes, based on strong personal suspicion or medical advice
- Unsure
- No

Figure 1 – Prevalences of confirmed or self-reported COVID-19 by cohort

Ranges show point estimates with 95% CI
Simple analyses: Examples

Does the prevalence of self-reported COVID-19 differ between males and females in each cohort?

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Male/female</th>
<th>N (total)</th>
<th>N (C-19) (%)</th>
<th>Risk ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS</td>
<td>Male</td>
<td>770</td>
<td>38 (4.9)</td>
<td>1.00</td>
<td>(ref)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>1,839</td>
<td>120 (6.5)</td>
<td>1.45</td>
<td>0.92, 2.27</td>
</tr>
<tr>
<td>Next Steps</td>
<td>Male</td>
<td>643</td>
<td>82 (12.8)</td>
<td>1.00</td>
<td>(ref)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>1,233</td>
<td>115 (9.3)</td>
<td>0.68</td>
<td>0.44, 1.06</td>
</tr>
<tr>
<td>BCS70</td>
<td>Male</td>
<td>1,711</td>
<td>160 (9.4)</td>
<td>1.00</td>
<td>(ref)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>2,420</td>
<td>219 (9.0)</td>
<td>1.00</td>
<td>0.71, 1.41</td>
</tr>
<tr>
<td>NCDS</td>
<td>Male</td>
<td>2,432</td>
<td>137 (5.6)</td>
<td>1.00</td>
<td>(ref)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>2,686</td>
<td>159 (5.9)</td>
<td>1.17</td>
<td>0.83, 1.66</td>
</tr>
</tbody>
</table>
Confounder control
Confounder control

• If we want an estimated association between an independent variable and a dependent variable to have any causal interpretation, we need to consider confounder control.

• **Confounder**: A variable that causes non-causal (spurious) association between an independent variable and a dependent variable.
Confounder control

- If we want an estimated association between an independent variable and a dependent variable to have any causal interpretation, we need to consider confounder control.

- **Confounder**: A variable that causes non-causal (spurious) association between an independent variable and a dependent variable.
Confounding control

• If we want an estimated association between an independent variable and a dependent variable to have any causal interpretation, we need to consider confounder control.

• **Confounder**: A variable that causes non-causal (spurious) association between an independent variable and a dependent variable.
Confounder control

• If we want an estimated association between an independent variable and a dependent variable to have any causal interpretation, we need to consider confounder control.

• **Confounder**: A variable that causes non-causal (spurious) association between an independent variable and a dependent variable.
Confounder control

- If we want an estimated association between an independent variable and a dependent variable to have any causal interpretation, we need to consider confounder control.

- **Confounder**: A variable that causes non-causal (spurious) association between an independent variable and a dependent variable.
Confounder control

• If we want an estimated association between an independent variable and a dependent variable to have any causal interpretation, we need to consider confounder control.

• **Confounder**: A variable that causes non-causal (spurious) association between an independent variable and a dependent variable.

• Thankfully, the rich data collected on cohort members over many years/decades provide great opportunity for confounder control.
Confounder control: Example

JAMA Psychiatry | Original Investigation

Association of Early-Life Mental Health With Biomarkers in Midlife and Premature Mortality
Evidence From the 1958 British Birth Cohort

George B. Ploubidis, PhD; G. David Batty, PhD, DSc; Praveetha Patalay, PhD; David Bann, PhD; Alissa Goodman, MSc

**IMPORTANCE** Early-life mental health is known to be associated with socioeconomic adversity and psychological distress in adulthood, but less is known about potential associations with biomarkers and mortality.

**OBJECTIVE** To investigate the association between early-life mental health trajectories with biomarkers in midlife and premature mortality.

**DESIGN, SETTING, AND PARTICIPANTS** This study used data from the British National Child Development Study, a population-based birth cohort. The initial sample of 17,415 individuals consisted of all infants born in Great Britain in a single week in 1958. Analysis began February 2017 and ended May 2020.
Confounder control: Example

Early-life mental health (age 7-16) ⟷ Biomarkers in midlife (age 44-45)
Confounder control: Example

**Early-life mental health (age 7-16)**
Rutter Child Scale A at ages 7 and 11 (mothers) and at age 16 (teachers):
- Conduct problems
- Affective symptoms

Early-life mental health (age 7-16) \( ? \) Biomarkers in midlife (age 44-45)
Confounder control: Example

Biomarkers in midlife (age 44-45)
- Fibrinogen
- C-reactive protein
- Glycated haemoglobin
- High-density lipoprotein
- Low-density lipoprotein
- High blood pressure

Early-life mental health (age 7-16)  ❯  Biomarkers in midlife (age 44-45)
Confounder control: Example

Birth characteristics

Early-life mental health (age 7-16) → Biomarkers in midlife (age 44-45)

?
Confounder control: Example

- Birth characteristics
- Parental characteristics
- Early-life mental health (age 7-16)
- Biomarkers in midlife (age 44-45)
Confounder control: Example

Birth characteristics → Early-life mental health (age 7-16)

Parental characteristics → Early-life mental health (age 7-16) → Biomarkers in midlife (age 44-45)

Early life socio-economic position → Biomarkers in midlife (age 44-45)
Confounder control: Example

Birth characteristics → Early-life mental health (age 7-16)
Parental characteristics → Early-life mental health (age 7-16)
Early life socio-economic position → Early-life mental health (age 7-16)
Early life cohort member characteristics → Early-life mental health (age 7-16)

Early-life mental health (age 7-16) → Biomarkers in midlife (age 44-45)
Biomarkers in midlife (age 44-45) → ?

Biomarkers in midlife (age 44-45) → Early life cohort member characteristics
Early life cohort member characteristics → Early-life mental health (age 7-16)
Early life cohort member characteristics → Parental characteristics
Parental characteristics → Early-life mental health (age 7-16)
Parental characteristics → Early life socio-economic position
Early life socio-economic position → Early-life mental health (age 7-16)
Early life socio-economic position → Parental characteristics
Parental characteristics → Early life socio-economic position
Confounder control: Example

**Birth characteristics**
- Birthweight
- Maternal smoking during pregnancy
- Maternal age
- Maternal partnerships status
- Whether cohort member was ever breastfed
- Region of residence at birth

Early-life mental health (age 7-16) → Biomarkers in midlife (age 44-45)
Confounder control: Example

**Parental characteristics**
- Maternal employment up to age 5
- Parents reading to child
- Parental interest in school
- Divorce by age 7
- Separation from child from more than one month

**Birth characteristics**
- Early-life mental health (age 7-16)

**Daily life cohort characteristics**
- Biomarkers in midlife (age 44-45)
Confounder control: Example

Early life socio-economic position
- Paternal social class at birth
- Financial difficulties at age 7
- Age mother left school
- Housing tenure at age 7
- Access to household amenities
- Housing difficulties at age 7
Confounder control: Example

Early life cohort member characteristics
- Cognitive ability
- Nocturnal enuresis
- Health conditions
- Experience of bullying
- Body mass index

Birth characteristics

Early life mental health (age 7-16)

Midlife health (age 44-45)

Biomarkers in midlife (age 44-45)
Confounder control: Example

Birth characteristics

Parental characteristics

Early life socio-economic position

Early life cohort member characteristics

Early-life mental health (age 7-16)

Biomarkers in midlife (age 44-45)
Confounder control: Key message

- The rich data collected on cohort members over many years/decades provide great opportunity for confounder control.
Repeated measures
Repeated measures

- Long-running cohorts measuring consistent topics over time provide repeated measures of the same measurement/construct.

- Examples:
  - Physical measurements
  - General physical health, mental health, specific diseases/conditions, health behaviours
  - Relationships, marital status, household composition
  - Employment status, occupation, earnings and income

- Allows you to characterise changes or trajectories over time.
Patterns of Atopic Eczema Disease Activity From Birth Through Midlife in 2 British Birth Cohorts

Katrina Abuabara, MD, MA, MSCE; Morgan Ye, MPH; David J. Margolis, MD, PhD; Charles E. McCulloch, PhD; Amy R. Mulick, MSc; Richard J. Silverwood, PhD; Alice Sullivan, PhD; Hywel C. Williams, DSc; Sinéad M. Langan, PhD

**IMPORTANCE** Atopic eczema is characterized by a heterogenous waxing and waning course, with variable age of onset and persistence of symptoms. Distinct patterns of disease activity such as early-onset/resolving and persistent disease have been identified throughout childhood; little is known about patterns into adulthood.

**OBJECTIVE** This study aimed to identify subtypes of atopic eczema based on patterns of disease activity through mid-adulthood, to examine whether early life risk factors and participant characteristics are associated with these subtypes, and to determine whether subtypes are associated with other atopic diseases and general health in mid-adulthood.

**DESIGN, SETTING, AND PARTICIPANTS** This study evaluated members of 2 population-based birth cohorts, the 1958 National Childhood Development Study (NCDS) and the 1970 British Cohort Study (BCS70). Participant data were collected over the period between 1958 and 2016. Data were analyzed over the period between 2018 and 2020.
Repeated measures: Example

- Aimed to identify subtypes of eczema based on patterns of disease activity in NCDS and BCS70.

- Parent-reported or self-reported eczema period prevalence available from standardised questions at ages 7, 11, 16, 23, 42 and 50 in NCDS and ages 5, 10, 16, 26, 30, 34, 38, 42 and 46 in BCS70.

- Then examined whether:
  - early life risk factors associated with eczema subtypes
  - eczema subtypes associated with other atopic diseases and general health in mid-adulthood
Repeated measures: Example

Figure. Estimated Probabilities of Atopic Eczema Symptoms at Each Age for Each Subtype in 4-Class Models From 2 British Birth Cohorts

A 1958 Cohort

B 1970 Cohort

Predicted probabilities at each age generated from generalized linear and latent mixed models.
Repeated measures: Key message

- British cohort studies provide repeated observations of the same measurement/construct.
- Allows you to characterise changes or trajectories over time.
Cross-cohort analysis
Cross-cohort analysis

- Conducting analyses across multiple cohorts allows us to extend our hypotheses: how do things change over time or between cohorts?
- Ideally want to analyse *identical* measures across cohorts.
- In absence of this, need consider how measures can best be *harmonised*.
- COVID-19 surveys offer great opportunity for cross-cohort analysis as most questions identical.
Socioeconomic inequalities in childhood and adolescent body-mass index, weight, and height from 1953 to 2015: an analysis of four longitudinal, observational, British birth cohort studies

David Barr, William Johnson, Leah Li, Diana Kuh, Rebecca Hardy

Summary
Background Socioeconomic inequalities in childhood body-mass index (BMI) have been documented in high-income countries; however, uncertainty exists with regard to how they have changed over time, how inequalities in the composite parts (i.e., weight and height) of BMI have changed, and whether inequalities differ in magnitude across the outcome distribution. Therefore, we aimed to investigate how socioeconomic inequalities in childhood and adolescent weight, height, and BMI have changed over time in Britain.

Methods We used data from four British longitudinal, observational, birth cohort studies: the 1946 Medical Research Council National Survey of Health and Development (1946 NSHD), 1958 National Child Development Study (1958 NCDS), 1970 British Cohort Study (1970 BCS), and 2001 Millennium Cohort Study (2001 MCS). BMI (kg/m²) was derived in each study from measured weight and height. Childhood socioeconomic position was indicated by the
Cross-cohort analysis: Example

- Investigated how socioeconomic inequalities in childhood and adolescent weight, height, and BMI have changed over time.
- Used data from NSHD (BMI at ages 7, 11 and 15), NCDS (7, 11 and 16), BCS70 (10 and 16) and MCS (7, 11 and 14).
- Childhood socioeconomic position indicated by father's occupational social class reported at age 10-11.
- Examined associations between childhood socioeconomic position and BMI to assess socioeconomic inequalities.
- Examined whether inequalities widened or narrowed from childhood to adolescence.
Cross-cohort analysis: Example
Cross-cohort analysis: Example

“There was little inequality in childhood BMI in the 1946–70 cohorts, whereas inequalities were present in the 2001 cohort and widened from childhood to adolescence in the 1958–2001 cohorts…”
Cross-cohort analysis: Key message

• Conducting analyses across multiple cohorts allows us to extend our hypotheses: how do things change over time or between cohorts?

• Bann D, et al. *Investigating change across time in prevalence or association using observational data: guidance on utility, methodology, and interpretation.* Discover Social Science and Health. 2022; 2: 18.
CLLS bibliography

Bibliography

What is the bibliography?
The CLS bibliography is a searchable database of over 6000 publications based on data from the 1958, 1970 and Millennium cohort studies, and more recently the Health and Retirement Survey data in the USA. It is intended for anyone who wishes to obtain background information on CLS sources, or to cite them in their research.

CLS relays on researchers to inform us when they have published research using the cohort data. If you have a publication to contribute to the bibliography please contact us.

Tips for searching
The database is searchable by year (1958, 1960, 1970, etc.) author and journal title. You can search by keywords or phrases in the filter or abstract. If your search contains a hyphen or a dash, try a shorter version of it that includes all that character.

Filter By
Choose a filter from the drop downs and fill in to narrow your search:

Search Year: From [ ] To [ ]
Study:

Author [ ] Journal [ ] Title/Abstract [ ]

Search Clear

https://www.bibliography.cls.ucl.ac.uk
The CLS bibliography is a searchable database of over 5,000 publications based on data from the 1958, 1970 and Millennium birth cohort studies, and more recently the Next Steps cohort study. It's a useful resource for finding out what's already been published on certain subjects, and for building reading lists for literature reviews and courses.
Thank you.
Getting started with the data
This section

- Available resources
- Accessing the data
- Key ID’s and other data protocols
- Merging data within and across sweeps
- Study design and sample weights
- Non-response and attrition
- Where to go for more information
Available resources  https://cls.ucl.ac.uk/

- User guides
  - Overview of measures
  - Response and weights
- Questionnaires
  - Exact question wording
  - Questionnaire routing
  - Variable names
- Data documentation
  - Data notes
  - Coding frames
  - Variables lists, including derived variables
- Technical reports
  - Sample and questionnaire design, development
  - Fieldwork, response, ethics
  - Coding, editing
- Data dictionaries
- Cohort profiles e.g.
- Previous journal publications
Resources available: UK Data Service
Data freely available to researchers, government analysts and third sector
And most research data is accessed via the UK Data Service
Access to different types of data at the UKDS

Access to data held by the UK Data Service varies depending on how the data is classified:

**Safeguarded data** available under **End User Licence** (EUL): data with a low level of sensitivity and disclosivity.
- Most of our data are available under this licence.
- Your application is authorised directly by the UK Data Service, and you can download the data directly from there.

**Special safeguarded data** available under **Special Licence** (SL): access to moderately sensitive or disclosive data. Access through the UK Data Service and application approved by CLS before you can download the data.

**Controlled data** available under **Secure Access Licence** (SA) for access to the most sensitive and/or potentially disclosive data. Access through the UK Data Service and attend a specialised training course. CLS approval and access via UK Data Service SecureLab

For details on specialist linked administrative data, genetic data and more information on data access please visit: [https://cls.ucl.ac.uk/data-access-training/data-access/](https://cls.ucl.ac.uk/data-access-training/data-access/)
Registering with the UK Data Service

1. Click ‘Login’
2. On the Login page begin typing your organisation name. Select the organisation required
3. Click ‘Continue’
4. Your own organisation login page will then be displayed. Login with your usual username and password
5. Complete the registration form with your details, selecting other options as required
6. Agree to the End User Licence (EUL), which outlines the terms and conditions of use of the Service
7. Click ‘Register’.

Once registration is complete you will be able to download/order or request access to data
Accessing CLS cohort data

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCDS</td>
<td><a href="http://discover.ukdataservice.ac.uk/series/?sn=2000032">http://discover.ukdataservice.ac.uk/series/?sn=2000032</a></td>
</tr>
<tr>
<td>BCS70</td>
<td><a href="http://discover.ukdataservice.ac.uk/series/?sn=200001">http://discover.ukdataservice.ac.uk/series/?sn=200001</a></td>
</tr>
<tr>
<td>Next Steps</td>
<td><a href="http://discover.ukdataservice.ac.uk/series/?sn=2000030">http://discover.ukdataservice.ac.uk/series/?sn=2000030</a></td>
</tr>
<tr>
<td>MCS</td>
<td><a href="http://discover.ukdataservice.ac.uk/series/?sn=2000031">http://discover.ukdataservice.ac.uk/series/?sn=2000031</a></td>
</tr>
</tbody>
</table>

**Millennium Cohort Study**

The Millennium Cohort Study (MCS) began in 2000, conducted by the Centre for Longitudinal Studies (CLS). It aims to study the social, economic, and health advantages and disadvantages facing children born at the start of the 21st century. The study has been tracking the 'Millennium children' through their early childhood years and plans to follow them into adulthood. It provides a basis for comparing patterns of development with the preceding cohort studies (the National Child Development Study (NCDS) and the 1970 Birth Cohort Study (BCS70)).
Accessing CLS cohort data II

Before downloading the data:
- Click on Request Access
- Click on Complete actions

- Agree to standard ‘End User Licence’
- Read and agree extra conditions

Choose data format and download zip file
- SPSS
- STATA
- TAB (tab-delimited)
# Files: Datasets

<table>
<thead>
<tr>
<th>Name</th>
<th>Contents</th>
<th>Structure</th>
<th>Identifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS8_2015_Main_Interview</td>
<td>Modules 1 to 7</td>
<td>Flat</td>
<td>NSID</td>
</tr>
<tr>
<td>NS8_2015_Self_Completion</td>
<td>Module 8</td>
<td>Flat</td>
<td>NSID</td>
</tr>
<tr>
<td>NS8_2015_Partnerships</td>
<td>Relationship histories</td>
<td>Hierarchical</td>
<td>NSID, W8RELID</td>
</tr>
<tr>
<td>NS8_2015_Children</td>
<td>Details of children of CM</td>
<td>Hierarchical</td>
<td>NSID, W8CHID</td>
</tr>
<tr>
<td>NS8_2015_Household_Members</td>
<td>Details of members living in same household as CM</td>
<td>Hierarchical</td>
<td>NSID, W8HHMID</td>
</tr>
<tr>
<td>NS8_2015_Activity_History</td>
<td>Activities and Employment histories</td>
<td>Hierarchical</td>
<td>NSID, W8HISTID</td>
</tr>
<tr>
<td>NS8_2015_Benefits</td>
<td>Details of individual benefits received</td>
<td>Hierarchical</td>
<td>NSID, W8BENID</td>
</tr>
<tr>
<td>NS8_2015_Income_Unfolding_brackets</td>
<td>Unfolding brackets questions for payments and income</td>
<td>Flat</td>
<td>NSID</td>
</tr>
<tr>
<td>NS8_2015_Benefits_Unfolding_brackets</td>
<td>Unfolding brackets questions for benefits</td>
<td>Hierarchical</td>
<td>NSID, W8BENID</td>
</tr>
<tr>
<td>NS8_2015_Derived_variables</td>
<td>Derived variables</td>
<td>Flat</td>
<td>NSID</td>
</tr>
</tbody>
</table>
## Key identifiers (ID’s)

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Key cohort identifier</th>
<th>Key cohort member/family identifier format</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NCDS 1958</td>
<td>NCDSID</td>
<td>7 characters:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>N followed by 5 digits, and a single character</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. N10016V</td>
<td></td>
</tr>
<tr>
<td>BCS70 1970</td>
<td>BCSID</td>
<td>7 characters:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>B followed by 5 digits, and a single character</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. B25819Z</td>
<td></td>
</tr>
<tr>
<td>Next Steps 1989/90</td>
<td>NSID</td>
<td>8 characters:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NS followed by 5 digits and a single character</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. NS21140C</td>
<td></td>
</tr>
<tr>
<td>MCS 2000/02</td>
<td>MCSID</td>
<td>7 characters:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M followed by 5 digits, and a single character</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. M10029A</td>
<td></td>
</tr>
</tbody>
</table>

**Cohort member/family identifier:**

Every cohort member (or for the MCS family) has the same ID across sweeps.

Use these ID’s to link datasets.

### W8xxID

- Used in particular files to denote relationship, child, HH member etc

### CNUM

- Cohort members, 1, 2 (twins) or 3 (triplets)

### PNUM

- Person number, for everyone else in the family apart from cohort members: parents, siblings, grandparents, etc
File structures: Flat v hierarchical

<table>
<thead>
<tr>
<th>Type of file structure</th>
<th>Format</th>
<th>Examples in the cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>1 record per case</td>
<td>NCDS, BCS70, Next Steps – main data files</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MCS – family files</td>
</tr>
<tr>
<td>Hierarchical</td>
<td>1 or more records per case</td>
<td>Household files</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Activities e.g. employment histories</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Relationship histories</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Time use diaries</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MCS - Person within family</td>
</tr>
</tbody>
</table>

### Examples of data entry

<table>
<thead>
<tr>
<th>BCSID</th>
<th>Sex</th>
<th>Country</th>
<th>Emp1</th>
<th>Emp2</th>
<th>Emp3</th>
</tr>
</thead>
<tbody>
<tr>
<td>B567689A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>B467921B</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B879255C</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>B297614D</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>B349725E</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Notes
- Flat files contain 1 record per case.
- Hierarchical files contain 1 or more records per case.
- NCDS, BCS70, and Next Steps are main data files.
- MCS family files are also included.
- Household files include details of family activities, employment histories, relationship histories, and time use diaries.
- MCS records can be further divided into Person within family and other family members.
- BCSID, Sex, Country, and Employment records are examples of data entry.
Merging data within and across sweeps

1. Identify appropriate files
   - Establish number of cases in target population

2. Check file structure: flat v hierarchical
   - Transform if necessary

3. Identify merging variables:
   1. Unique ‘key’ cohort ID (member or family)
   2. Other ID’s depending on merge
      • Check the same variable name (case sensitive, changed across sweep etc)
      • Create identical variable name if necessary

4. Check merged correctly
Study design (sampling and sample weights)

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Population</th>
<th>Sample and study design</th>
<th>Design weight(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCDS 1958</td>
<td>All born in GB in one week</td>
<td><strong>Total sample:</strong> Captured 98% of the total births in GB in the target week</td>
<td>None</td>
</tr>
<tr>
<td>BCS70 1970</td>
<td>All born in GB in one week</td>
<td><strong>Total sample:</strong> Captured 95-98% of the total births in GB in the target week</td>
<td>None</td>
</tr>
<tr>
<td>Next Steps 1989/90</td>
<td>Young people in England in 2004 born between 01/09/89 and 31/8/90</td>
<td><strong>Complex sample design:</strong> Maintained schools* 2-stage sampling procedure</td>
<td>psu (SampPSU) strata (SampStratum)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stage 1: Schools</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stage 2: Pupils within schools.</td>
<td></td>
</tr>
<tr>
<td>MCS 2000/02</td>
<td>Children born in 2000 /2002 and living in the UK at age nine months,</td>
<td><strong>Complex sample design:</strong> Clusters: Areas of residence (electoral wards)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Disproportionately stratified by area disadvantage, and UK country, and ethnicity in England</td>
<td></td>
</tr>
</tbody>
</table>

*Independent and referral units sampled differently
Non-response and attrition

- Distinction between unit (respondents’) non-response and item non-response
  - Unit non-response (not responding to a particular sweep)
    - Non-response is common in longitudinal surveys
  - Item non-response i.e. not answering some questions
    - tends to be less of an issue in the cohorts

- Missing data may be a risk to representativeness
  - Potential for bias since respondents are often systematically different from nonrespondents
Dealing with unit non-response

- **Case-wise deletion**
  - Any individual in a data set is deleted from an analysis if they're missing data on any variable in the analysis.
  - Straightforward, but doesn't deal with any non-response bias.

- **Non-response weights**
  - Adjust the sample composition to take account of the loss of particular type of respondents.
  - Provided in MCS (govwt2 = overall in MCS7 for whole of UK analysis) and Next Steps (W8FINWT = final weight for age 25 survey).

- **Other more advanced methods**
  - MI involves the generation of multiple copies of the dataset in each of which missing values are replaced by imputed values sampled from their posterior predictive distribution given the observed.
Where to go for more information
CLOSER Learning Hub: [https://learning.closer.ac.uk/](https://learning.closer.ac.uk/)

CLOSER provides training and resources for students and early-career researchers to “maximise the use, value and impact of longitudinal research”

Learning Hub – demonstration video: [https://youtu.be/Z_bFCClq2Dc](https://youtu.be/Z_bFCClq2Dc)
NCRM has an extensive library of resources on research methods. These include those in our EPrints publications database, which has more than 3,000 items, and our collection of online tutorials.

- **Online tutorials**
- **NCRM EPrints - search our publications database**
- **Resources for trainers**
- **Videos**
- **Podcasts**
- **Courses and events**

[https://www.ncrm.ac.uk/resources/](https://www.ncrm.ac.uk/resources/)
Enhance your data skills and teaching

**New to using data**
Best practice and training for researchers new to accessing and using data in our collection. Includes advice and tools to correct cite data; student-specific information on our Dissertations Award for undergraduates; and more.

**Data skills modules**
There is a wealth of data available for reuse in research and reports. These free, interactive tutorials are designed for anyone who wants to start using secondary data. They show you how to get started with finding good quality data, understanding it and starting your analyses.

**Students**
Students can access most of the UK Data Service's collection of social, economic and population data. Find resources to help you find and use our data during your studies including the UK Data Service dissertation resources.

**Survey data**
Survey data, including data from long-running surveys, series and longitudinal studies, are a major part of social science research. Learn how to use survey and longitudinal data through training resources including videos, on-demand webinars and written guides.

**International data**
Our international microdata contains socio-economic time series data aggregated to a country or regional level for a range of countries over a substantial time period.

**Qualitative data**
Qualitative research gives a voice to the lived experience, offering researchers a deeper insight into a topic or individuals' experiences. Qualitative data can be combined with quantitative to enhance understanding around a policy or topic in a way that quantitative data by itself often cannot.
Resources available: CLS website

https://cls.ucl.ac.uk/
Resources available on each cohort and sweeps:
User Guide, technical resources and questionnaires
### CLS training and support

#### Upcoming training events

<table>
<thead>
<tr>
<th>Event Description</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction to Longitudinal Data: Structure and Visualisation</td>
<td>21 November 2023</td>
</tr>
<tr>
<td>Ageing in the British cohort studies: measurement, research and access</td>
<td>Early Feb 2024</td>
</tr>
<tr>
<td>Genetic data: An overview of genetic data in the British cohort studies</td>
<td>Feb/March 2024</td>
</tr>
<tr>
<td>Methods: Cross-cohort analyses</td>
<td>May/June 2024</td>
</tr>
</tbody>
</table>

[https://cls.ucl.ac.uk/data-access-training/training-and-support-2/](https://cls.ucl.ac.uk/data-access-training/training-and-support-2/)

[https://cls.ucl.ac.uk/events/](https://cls.ucl.ac.uk/events/)
So we’ve covered

- An introduction to birth cohorts
- Some of the content in the CLS cohorts by subject areas
- Examples of the types of analysis
- Getting started with the data
- Where to go for more information
Thank you
Any questions?