CLOSER’s resources for teaching and research

Professor Alison Park, Director
Session 2

Focus on resources for researchers

• CLOSER Discovery

• Harmonisation and data linkage resources
About CLOSER

- **Objective**: to maximise the use, value and impact of the UK’s longitudinal studies

- **Consortium**: 8 longitudinal studies, the British Library and the UK Data Service

- **Funded** by the ESRC and the MRC
Understanding Society: UK Longitudinal Household Study

Millennium Cohort Study

Southampton Women’s Survey

Avon Longitudinal Study of Parents and Children (ALSPAC)

1970 British Cohort Study

1958 National Child Development Study

MRC National Survey of Health and Development

Hertfordshire Cohort Study
Key areas of CLOSER’s work

- Guidance and exemplar studies on data harmonisation and data linkage
- Promoting value of longitudinal research
- Training and capacity building: Learning Hub
- Data discovery: CLOSER Discovery
Data discovery tools
Oceanography
Astronomy

The SAO/NASA Astrophysics Data System

Welcome to the Digital Library for Physics and Astronomy

This site is hosted by the
High Energy Astrophysics Division at the
Harvard-Smithsonian Center for Astrophysics
Environment
New York

Providing Unprecedented Access to Data

Browse and download nearly 1,400 New York State data resources on topics ranging from farmers' markets to solar photovoltaic projects to MTA turnstile usage to Arts for Transit.
# Data discovery tools

<table>
<thead>
<tr>
<th>Data</th>
<th>Search</th>
<th>Engine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata</td>
<td>Discovery</td>
<td>Platform</td>
</tr>
<tr>
<td>Uniform</td>
<td>Documentation</td>
<td>Portal</td>
</tr>
</tbody>
</table>
## Data discovery tools

<table>
<thead>
<tr>
<th>Data</th>
<th>Search</th>
<th>Engine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata</td>
<td>Discovery</td>
<td>Platform</td>
</tr>
<tr>
<td>Uniform</td>
<td>Documentation</td>
<td>Portal</td>
</tr>
</tbody>
</table>
Metadata

• “What do you understand by the term ‘metadata’?”
  • “data about data”

• What is the difference between data and information?
  • Context

• What gives data context?
  • Metadata

• Metadata + Data = Information
Example

<table>
<thead>
<tr>
<th>Fields</th>
<th>Units</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eye colour</td>
<td></td>
<td>Brown</td>
</tr>
<tr>
<td>Shoe size</td>
<td>Paris points</td>
<td>42</td>
</tr>
<tr>
<td>Height</td>
<td>Centimetres</td>
<td>175</td>
</tr>
<tr>
<td>Favourite colour</td>
<td></td>
<td>Green</td>
</tr>
<tr>
<td>Condition</td>
<td></td>
<td>Handsome</td>
</tr>
</tbody>
</table>
Large amount of information

• 8 studies
  • > 80 sweeps
  • > 400 instruments
    • > 100,000 questions
• > 450 datasets
  • > 120,000 variables
The solution

CLOSER Discovery
CLOSER Discovery

• Search engine

• International standards (e.g. DDI-Lifecycle)
  • Comparable
  • Consistent

• Enhanced search-ability
  • Topics
  • Life stages
  • Equivalencies
# Data harmonisation

<table>
<thead>
<tr>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height, weight and BMI</td>
</tr>
<tr>
<td>Socio-economic status</td>
</tr>
<tr>
<td>Use/collection of biological samples</td>
</tr>
<tr>
<td>Measures of vision</td>
</tr>
<tr>
<td>Childhood environment and adult wellbeing</td>
</tr>
<tr>
<td>Review of methods for determining pubertal status</td>
</tr>
<tr>
<td>Exploiting CLOSER biomarker data</td>
</tr>
<tr>
<td>Overcrowding and health</td>
</tr>
</tbody>
</table>
Tracking the rise of obesity in the UK
(91st percentile)

Reference
Johnson, W., Li, L., Kuh, D., & Hardy, R. (2015). 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 Medicine*, e1001828. Contact: Rebecca Hardy rebecca.hardy@ucl.ac.uk
Harmonisation resources

- Harmonised datasets available via UKDS, with documentation
- Future outputs drawing out learning from across CLOSER’s harmonisation work
- Further harmonisation projects during 2017-2019
- Training events – eg cross-study harmonisation
Data linkage

- Economic and education data
- Geographic data
- Hospital Episode Statistics
- Data linkage in cohorts and longitudinal studies
Longitudinal analysis of schools

Note: multiple rows per school, as linked forward and backward and include school mergers. (Churchfields, the village school formed from Atworth Primary School and Monkton Farleigh & South Wraxall.)
Data linkage resources

- Outputs as part of ongoing data linkage coordination and advocacy
- Schools database
- Dataset linking P14 employment data (HMRC) to birth cohort data
- Training events (e.g., workshop on linking study data to geographic variables)
Thank you

alison.park@ucl.ac.uk
@parkali

w.poynter@ucl.ac.uk
@willpoynter

closer@ucl.ac.uk
@CLOSER_UK

closer.ac.uk
closer.ac.uk/discovery
closer.ac.uk/learning