Using linked administrative data: Hospital Episode Statistics linked with the CLS cohorts

6 December 2022
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.
• Slides and recording will be made available after the session.

• Thank you for joining us today.
## Today’s schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Topics covered</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00 – 1.15pm</td>
<td>Brief introduction to the cohorts</td>
<td>Richard Silverwood Associate Professor of Statistics</td>
</tr>
<tr>
<td>1.15 – 1.30pm</td>
<td>HES data and linkage overview</td>
<td>Sarah Kerry-Barnard Research Data Manager</td>
</tr>
<tr>
<td>1.30 – 1.45pm</td>
<td>Accessing linked HES datasets</td>
<td>Danielle Gomes and Karen Dennison Record Linkage Managers</td>
</tr>
<tr>
<td>1.45 – 2.15pm</td>
<td>Analysing linked HES data</td>
<td>Richard Silverwood Associate Professor of Statistics</td>
</tr>
<tr>
<td>2.15 – 2.30 pm</td>
<td>Q&amp;A</td>
<td>All</td>
</tr>
</tbody>
</table>
CLS studies follow people across life

- National Child Development Study 1958
- 1970 British Cohort Study
- Next Steps (formerly LSYPE)
- Millennium Cohort Study
Study timelines

Year

Age

MCS  Next Steps  BCS70  NCDS
# Typical information covered

<table>
<thead>
<tr>
<th>Birth (excluding Next Steps)</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, antenatal care)</td>
<td>Housing</td>
<td>Housing</td>
</tr>
<tr>
<td>Labour (length, pain relief, problems)</td>
<td>Health</td>
<td>Health (including biomarkers)</td>
</tr>
<tr>
<td>Birthweight, length</td>
<td>Cognitive tests</td>
<td>Well-being and mental health</td>
</tr>
<tr>
<td></td>
<td>Emotions and behaviour</td>
<td>Health-related behaviour</td>
</tr>
<tr>
<td></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>
<tr>
<td></td>
<td></td>
<td>Views and expectations</td>
</tr>
</tbody>
</table>
Forthcoming training events

Introduction to NCDS
25 January 2023, 1.00–2.30pm

Introduction CLS cohorts to biomedical researchers (CLOSER)
9 February 2023, 1.00pm–2.00pm

Handling missing data in the CLS cohorts
28 February 2023, 1.00–4.30pm

Introduction to BCS70, Next Steps and MCS
Later in 2023

https://cls.ucl.ac.uk/events/
Call for papers - Children of the noughties: a conference to celebrate 21 years of the Millennium Cohort Study

• A scientific conference in London, 13-14 June 2023

• Papers may include substantive research topics and research methods. Studies involving cross-study or intergenerational comparisons, or analyses of linked data, are all welcome. It must include MCS data or another millennial longitudinal study. Deadline: Wednesday 18 January 2022
NHS England – Hospital Episode Statistics (HES)
Linked Health Administrative Data

- NHS England – Hospital Episode Statistics (HES)
  - NCDS
  - BCS70
  - Next Steps
  - MCS*

- NHS Scotland – Scottish Medical Records (SMR)
  - NCDS
  - BCS70
  - MCS

*available Wed 7th December 2022
1. Introduction

2. Consent to health data linkage

3. Health data linkage
   3.1 HES datasets
   3.2 Matching strategy
   3.3 Matching rates

4. The research datasets
   4.1 Licensing
   4.2 Data documentation provided
   4.3 Identifiers
   4.4 Data processing
   4.5 Data de-identification
   4.6 The Accident and Emergency (A&E) data
   4.7 The Admitted Patient Care (APC) data
   4.8 The Critical Care (CC) data
   4.9 The Outpatient Care (OP) data

5. Disclosure control: requirements for data users
   5.1. UKDS requirements
   5.2. NHS Digital requirements

6. Data access and variable selection
   6.1 UKDS Secure Access application
   6.2 Selection of variables
   6.3 CLS Licence Agreement

Appendices
Modifications to the Accident and Emergency Data
Modifications to the Admitted Patient Care Data
Modifications to the Outpatient Care Data
NHS England – Hospital Episode Statistics (HES)

Datasets available:

- Admitted Patient Care (APC)
- Critical Care (CC) – linked to APC records
- Accident and Emergency (A&E)
- Outpatient Care (OP)
- Emergency Care Dataset (ECDS – MCS only)
Hospital Episode Statistics over time – Next Steps, NCDS, BCS70

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accident and Emergency (A&amp;E)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Admitted Patient Care (APC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critical Care (CC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outpatient Care (OP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Hospital Episode Statistics over time – MCS

- Accident and Emergency (A&E)
- Admitted Patient Care (APC)
- Critical Care (CC)
- Outpatient Care (OP)
- Emergency Care Dataset (ECDS)
# HES - Linkage

<table>
<thead>
<tr>
<th></th>
<th>Number in sweep*</th>
<th>Gave consent (% of in sweep)</th>
<th>Has matched data (% of consented)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Next Steps</td>
<td>7,707</td>
<td>4,895 (63.5)</td>
<td>4,579 (93.5)</td>
</tr>
<tr>
<td>BCS70</td>
<td>9,841</td>
<td>7,048 (71.6)</td>
<td>5,488 (77.9)</td>
</tr>
<tr>
<td>NCDS</td>
<td>9,790</td>
<td>7,065 (72.2)</td>
<td>6,188 (87.6)</td>
</tr>
<tr>
<td>MCS</td>
<td>10.757</td>
<td>9,214 (86.7)</td>
<td>7,396 (81.5)</td>
</tr>
</tbody>
</table>

* Sweep in which consent was sought for linkage:

- Next Steps: Age 25 Survey
- BCS70: Age 42 Survey
- NCDS: Age 50 Survey
- MCS: Age 17 Survey
HES – Linkage

Number of cohort members represented in each HES dataset for each cohort

- Admitted Patient Care: 7,396
- Critical Care: 6,188
- Accident and Emergency: 5,488
- Outpatient: 4,578

Legend:
- NCDS
- BCS
- Next Steps
- MCS
The NHS organizes the data into episodes

- **Episode**
  - A continuous period of patient care administered under one consultant within health care providers.
  - Each episode is a row in each dataset.

- **Hospital Spell**
  - Total time spent by a patient in the same care provided by a hospital, from date of admission to date of discharge.
  - Spells cover one or more episodes.
  - To avoid double counting, use the last episode in a spell (spelend = Y).
HES - What data are available?

Admitted Patient Care 1997-2017

- Diagnostic codes (ICD-10)
- Operation Chapter (OPCS-4)
- Maternity data
- Indices of Multiple Deprivation
- Administrative data e.g. Health Authority
- Dates of admission, discharge etc

Critical Care 2009-2017

- Types of specialist support i.e. Renal
- Number of days spent with assistance for specialist care.
- Every critical care episode links with an APC code (D_SUSID)
HES - What data are available?

**Outpatient Care 2003-2017**
- Diagnostic codes (ICD-10) (poorly completed)
- Operation Chapter (OPCS-4)
- Indices of Multiple Deprivation
- Administrative data
- Dates (DD/MM/YYYY) seen

**Accident and Emergency 2007-2017**
- Diagnostic codes (A&E)
- Treatment codes
- Administrative data
- Dates (DD/MM/YYYY) time admission and discharge
HES – what data are available?

Emergency Care Dataset

- SNOMED-CT
- Diagnostic codes
- Causes of injury & drug information
- Treatment
- Dates of arrival/being seen
HES - Dictionaries & Supplementary Documentation

- CLS User Guide
- NHS Data Dictionaries
- HES Analysis Guide
- International Classification of Diseases ICD-10 (Diagnoses, Symptoms, Causes)
- Operational codes OPCS-4
- A&E Diagnosis and Treatment codes

➢ Available on the Secure Access server at the UK Data archive
➢ Can be viewed alongside the data
➢ Links can also be found in the CLS user guide
Chapter I
Certain infectious and parasitic diseases (A00-B99)

Incl.: diseases generally recognized as communicable or transmissible

Use additional code (U80-U89), if desired, to identify the antibiotic to which a bacterial agent is resistant.

Excl.: carrier or suspected carrier of infectious disease (Z22.-)
certain localized infections - see body system-related chapters
infectious and parasitic diseases complicating pregnancy, childbirth and the puerperium [except obstetrical tetanus] (Q98.-)
infectious and parasitic diseases specific to the perinatal period [except tetanus neonatorum, congenital syphilis, perinatal gonococcal infection and perinatal human immunodeficiency virus [HIV] disease] (P35-P39)
influenza and other acute respiratory infections (J00-J22)

This chapter contains the following blocks:
A00-A09 Intestinal infectious diseases
A15-A19 Tuberculosis
A20-A28 Certain zoonotic bacterial diseases
A30-A49 Other bacterial diseases
A50-A64 Infections with a predominantly sexual mode of transmission
A65-A69 Other spirochaetal diseases
A70-A79 Other diseases caused by chlamydiae
A75-A79 Rickettsioses
A80-A89 Viral infections of the central nervous system
A90-A99 Arthropod-borne viral fevers and viral haemorrhagic fevers
B00-B09 Viral infections characterized by skin and mucous membrane lesions
B15-B19 Viral hepatitis
B20-B24 Human immunodeficiency virus [HIV] disease
B25-B34 Other viral diseases
B35-B49 Mycoses
B50-B64 Protozoal diseases
ICD-10 Version: 2010

A15 - Respiratory tuberculosis, bacteriologically and histologically confirmed

A15.0 Tuberculosis of lung, confirmed by sputum microscopy with or without culture
    Tuberculosis:
    - bronchietasis
    - fibrosis of lung
    - pneumonia
    - pneumothorax
    confirmed by sputum microscopy with or without culture

A15.1 Tuberculosis of lung, confirmed by culture only
    Conditions listed in A15.0, confirmed by culture only

A15.2 Tuberculosis of lung, confirmed histologically
    Conditions listed in A15.0, confirmed histologically

A15.3 Tuberculosis of lung, confirmed by unspecified means
    Conditions listed in A15.0, confirmed but unspecified whether bacteriologically or histologically

A15.4 Tuberculosis of intrathoracic lymph nodes, confirmed bacteriologically and histologically
    Tuberculosis of lymph nodes:
    - hilar
    - mediastinal
    - tracheobronchial
    confirmed bacteriologically and histologically

A15.5 Tuberculosis of larynx, trachea and bronchus, confirmed bacteriologically and histologically
    Tuberculosis of:
    - bronchus
    - glottis
    - larynx
    confirmed bacteriologically and histologically

A15.6 Tuberculous pleurisy, confirmed bacteriologically and histologically
    Tuberculosis of pleura
    confirmed bacteriologically and histologically

I Certain infectious and parasitic diseases
   A00-A09 Intestinal infectious diseases
      A15-A19 Tuberculosis
         A20-A28 Certain zoonotic bacterial diseases
         A30-A49 Other bacterial diseases
         A50-A64 Infections with a predominantly sexual mode of transmission
         A65-A69 Other spirochaetal diseases
         A70-A74 Other diseases caused by chlamydiae
         A75-A79 Rickettsioses
         A80-A89 Viral infections of the central nervous system
         A90-A99 Arthropod-borne viral fevers and viral haemorrhagic fevers
         B00-B09 Viral infections characterized by skin and mucous membrane lesions
         B15-B19 Viral hepatitis
         B20-B24 Human immunodeficiency virus (HIV) disease
         B25-B34 Other viral diseases
         B35-B49 Mycoses
         B50-B64 Protozoal diseases
         B65-B83 Helminthiases
         B85-B89 Pediculosis, scabies and other infestations
         B90-B94 Sequelae of infectious and parasitic diseases
         B95-B98 Bacterial, viral and other infectious agents
         B99-B99 Other infectious diseases
      II Neoplasms
      III Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
      IV Endocrine, nutritional and metabolic diseases
      V Mental and behavioural disorders
      VI Diseases of the nervous system
      VII Diseases of the eye and adnexa
      VIII Diseases of the ear and mastoid process
NHS Scotland – Scottish Medical Records (SMR)
NHS Scotland – Scottish Medical Data (SMR)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>NCDS</th>
<th>BCS70</th>
<th>MCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prescribing Information System (PIS)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Outpatient Records (SMR00)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Inpatient Records (SMR01)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Maternity Records (SMR02)</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth and Neonatal Records (SMR11)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Scottish Immunisation &amp; Recall System (SIRS)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Child Health Reviews (CHR)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>1st visit / 6-8 week visit / 8-9 month visit / 22-24 month visit / 37-42 month visit / Preschool visit / School reviews</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Accessing Health Data

UK Data Service (UKDS)
UK Longitudinal Linkage Collaboration (UKLLC)
Accessing the data

Due to the potential disclosivity and sensitivity of HES and SMR data, these datasets are only accessible under Secure Access conditions through the UK Data Service Secure Lab.

In order to have access to the data and as part of the application process:

- you must be based in a UK organisation, and be affiliated with a UK HE or FE institution.
- you must be an accredited researcher to use ONS data or apply to be one as part of the project application or you must fill in the accredited researcher form each time you apply to access ESRC data
- your User Agreement must be authorised and signed by someone who can take legal responsibility on behalf of your institution
- you must have attended and passed a Safe Researcher training course (a place on this course can be offered as part of the application process)
- you must be able to access the data from a suitable safe environment that meets the UK Data Service’s security requirements
To apply to use the data:

**Step 1: Register**

**Step 2: Create a project**
To apply to use the data:

**Step 3: Select your data**

   - Access data

   The Data Collection is available to users registered with the UK Data Service.
   - Commercial use is not permitted.
   - Use of the data requires approval from the data owner or their nominee. Users must apply for access via a Secure Access application.
   - The Data Collection is to be accessed remotely by registered, approved (by the data owner or his/her nominee), and trained users through a secure virtual private network via the researcher’s own institutional desktop PC.
   - The Data Collection must be accessed via a secure virtual private network in a safe environment approved by the UK Data Service.
   - Users should indicate on their Research Project Application form all Safeguarded dataset(s) that they wish to access alongside the study (selected from the 1970 British Cohort Study series elsewhere).

**Step 4: Assign dataset to a project**

1. Add dataset to your account (Only available via Secure Lab remote access)

   - Add to account

   - Add to project
To apply to use the data

Step 5: Access the appropriate forms

My Project to request data

Datasets in project

<table>
<thead>
<tr>
<th>SN</th>
<th>Dataset</th>
<th>Status</th>
<th>Actions</th>
</tr>
</thead>
</table>

- Accept standard EUL
- Complete Research Project Application documentation
- Complete ESRC Accredited Researcher Application
- Complete Secure Access User Agreement
- Depositor approval
- Training
- Secure Lab access

[Complete actions]
To apply to use the data

Step 5: Download Application forms

Research Project Application documentation

1. UKDS Project Application form
2. ESRC Accredited Researcher Application
3. Secure Access User Agreement
4. CLS Licence Agreement (HES only)
5. HES Variables list (HES only)
6. Complete the Secure access section
To apply to use the data

Step 5: Application forms (HES and SMR data)

- Completed by: Project Lead

- Completed by: the Principal Investigator/Project Lead

- Any project member who requires access to statistical analysis before release from the UK Data Service

- PhD supervisors and co-authors (even if they will not be analysing the data)
To apply to use the data

Step 5: Application forms- Agreements (HES and SMR data)

Project lead and their organisation signatory need to read and agree to terms.

Signed by:
- Project Lead &
- Institution Contracts Office or person authorised to sign on behalf of the organisation.
To apply to use the data

Step 5: Application – HES Variable selection (e.g. NCDS_HES_variables_List_1.xlsx)

<table>
<thead>
<tr>
<th>var_order</th>
<th>variable_name</th>
<th>variable_label</th>
<th>value_labels</th>
<th>Requested?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NSID</td>
<td>NSID - cohort member identifier</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot; (2) &quot;Brought in by ambulance (including helicopter)&quot; (3) &quot;Other&quot;</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>ARRIVALMODE</td>
<td>Arrival mode</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot; (2) &quot;First Accident and Emergency attendance&quot; (3) &quot;Follow-up Accident and Emergency attendance - planned&quot; (4) &quot;Follow-up Accident and Emergency attendance - unplanned&quot; (5) &quot;Other&quot;</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>ARRATTEND_EX_PLANED</td>
<td>Attendances excluding planned</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot; (2) &quot;First Accident and Emergency attendance&quot; (3) &quot;Follow-up Accident and Emergency attendance - planned&quot; (4) &quot;Follow-up Accident and Emergency attendance - unplanned&quot; (5) &quot;Other&quot;</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>ARRATTENDCAT</td>
<td>Attendance category</td>
<td>(1) &quot;Admitted to hospital bed at same MCP&quot; (2) &quot;Discharged follow-up treatment to be provided by RP&quot; (3) &quot;Discharged did not require any follow-up treatment&quot; (4) &quot;Referred to A&amp;E clinic&quot; (5) &quot;Referred to fracture clinic&quot; (6) &quot;Referred to other outpatient clinic&quot; (7) &quot;Transferred to other healthcare provider&quot; (8) &quot;Died in department&quot; (9) &quot;Referred to other healthcare professional&quot; (10) &quot;Left department before being treated&quot; (11) &quot;Left department having refused treatment&quot; (12) &quot;Other&quot;</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ARRATTENDDISP</td>
<td>Attendance disposal</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot; (2) &quot;Consultant-led minor specialty accident and emergency service&quot; (3) &quot;Other type of A&amp;E minor Injury activity with designated accommodation for reception of patients&quot; (4) &quot;NHS walk-in centres&quot; (5) &quot;Other&quot;</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>ADEPTTYPE</td>
<td>Department type</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot; (2) &quot;Consultant-led minor specialty accident and emergency service&quot; (3) &quot;Other type of A&amp;E minor Injury activity with designated accommodation for reception of patients&quot; (4) &quot;NHS walk-in centres&quot; (5) &quot;Other&quot;</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>AINCLOCTYPE</td>
<td>Incident location type</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot; (2) &quot;Home&quot; (3) &quot;Work&quot; (4) &quot;Educational establishment&quot; (5) &quot;Public place&quot; (6) &quot;Other&quot;</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>APATGROUP</td>
<td>Patient group</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot; (2) &quot;Road traffic accident&quot; (3) &quot;Assault&quot; (4) &quot;Deliberate self-harm&quot; (5) &quot;Sports injury&quot; (6) &quot;Firework injury&quot; (7) &quot;Other accident&quot; (8) &quot;Brought in dead&quot; (9) &quot;Other than above&quot;</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>AERESOURCE</td>
<td>Source of referral for A&amp;E</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot; (2) &quot;General medical practitioner&quot; (3) &quot;Self referral&quot; (4) &quot;Local authority social services&quot; (5) &quot;Emergency services&quot; (6) &quot;Work&quot; (7) &quot;Educational establishment&quot; (8) &quot;Police&quot; (9) &quot;Other&quot; (10) &quot;General dental practitioner&quot;</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>ARRIVALDATE</td>
<td>Arrival start date</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot; (2) &quot;Yes&quot; (3) &quot;No&quot;</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>ARRIVALTIME</td>
<td>Arrival Time HH:MM (00:00 to 23:59)</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot; (2) &quot;Yes&quot; (3) &quot;No&quot;</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>CARERSI</td>
<td>Caretaker support indicator</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot; (2) &quot;Yes&quot; (3) &quot;No&quot;</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>CONCLUDUR</td>
<td>Duration to conclusion (in minutes)</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot;</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>CONCLTIME</td>
<td>Conclusion time HH:MM (00:00 to 23:59)</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot;</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>DEPDur</td>
<td>Duration to departure (in minutes)</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot;</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>DEPTime</td>
<td>Departure time HH:MM (00:00 to 23:59)</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot;</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>DIAG_01</td>
<td>A&amp;E diagnosis</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot;</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>DIAG_02</td>
<td>A&amp;E diagnosis</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot;</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>DIAG_03</td>
<td>A&amp;E diagnosis</td>
<td>(1) &quot;Missing: not applicable/not available/not given&quot;</td>
<td></td>
</tr>
</tbody>
</table>
To apply to use the data

Step 5: Application forms - Additional agreement required (HES data only)
To apply to use the data
Step 5: Application forms - Additional agreement required (HES data only)

Important sections in the document: These are NHS Digital requirements.

✓ Legal basis
✓ Organisational Security
✓ Project proposal
✓ Provide project’s expected measurable benefits to Health and/or Social Care
✓ Select variables (this is different to Scottish health data where only datasets are specified)
✓ Completed by Project Lead and each member applying to use the data.
✓ Signed by Institution Contracts Offices or authorised signatory.
Access to secure data

Step 6: Depositor approval

- The application will be presented to the CLS Data Access Committee (DAC) for approval.
- For applications requesting access to HES data, DAC will assess whether the researcher has provided satisfactory information to the NHS Digital requirements sections highlighted previously.
Step 7: Training

- New Secure Lab researchers who haven’t previously completed the Safe User of Research Environment (SURE) or Safe Researcher Training will be invited by the UKDS to attend training.
- Researchers who completed training since January 2016 will be required to complete a short online training that covers the specifics of the Secure Lab,
- Current Secure Lab researchers who have not accessed the Secure lab for 30 months will be asked to complete a short online refresher course.
Access to secure data: Secure Lab Access

Complete actions for access

- Introduction
- Accept standard EUL
- Secure Access User Agreement form
- Depositor approval required
- Training required
- SecureLab access required

SecureLab access required

This Controlled dataset requires approval from the depositor (Research Application), training (Accredited Researcher) and a Secure Lab account to be set up before it can be accessed.

Please wait while we set up your account. You will be notified when this is completed.

- In progress.
Accessing the data

- For more information on how to apply for access, please visit https://www.ukdataservice.ac.uk/get-data/how-to-access/accesssecurelab

- If you have any enquiries regarding the application process, please contact the UK Data Service via their ‘Accessing data’ web form: https://beta.ukdataservice.ac.uk/help?id=accessingData

- Alternatively, the UK Data Service can be contacted via email at help@ukdataservice.ac.uk, or by telephone on 01206 872143.

Timescales
All applications must be validated and processed by the UK Data Service and then approved by the data owners. The UK Data Service works through a high volume of applications throughout the year. The UK Data Service endeavour to process applications as soon as possible but each application may take a number of months from start to finish.
Welsh health data in the MCS via the SAIL Data Bank

Datasets available:

• Emergency Department Dataset

• Primary Care GP dataset

• Outpatient Dataset

• Patient Episode Database (episodes of inpatient and day case care)
  ▪ MCS1-6 currently available
  ▪ Data for CM’s (up to age 14) and parents

Access

Secure Anonymised Information Linkage (SAIL) Databank – a Wales-wide research resource offering a range of health related datasets.

• Can apply for any health data in the SAIL Databank via https://data.ukserp.ac.uk/Organisation/Category? nodeId=1&orgId=0

• Application process can be found at https://saildatabank.com/application-process/two-stage-process/

• There is a charge determined on a case by case basis
Linked health administrative data available at the UKDS in the near future

- HES refresh for Next Steps, BCS70 and NCDS.
- CLS will refresh the Welsh health dataset linked to MCS (2001-2012) currently available at the UKDS.
Thank you
Any questions?
The UK Longitudinal Linkage Collaboration: a research resource for the longitudinal research community

UK LLC led by the Universities of Bristol & Edinburgh

Funded by

In collaboration with

A work package in the COVID-19 Longitudinal Health & Wellbeing National Core Study

Slides provided by UK LLC and adapted/presented by CLS
COVID-19 Longitudinal Health and Wellbeing National Core Study

A globally unique cross-cutting resource

20+ Longitudinal Population Studies
- COVID-19 collections (Questionnaires, Serology)
- SES & demographic
- Baseline physical & mental health
- Baseline family, SES and life-course indicators

NHS COVID-19 datasets
- GP Data
- Vaccinations
- COVID-19 test data (Pillar 1-3)
- Accident & Emergency
- CHESS

Wider NHS datasets
- Hospital Inpatient data
- Cancer & mortality registers
- Community Mental Health (MHSDS, IAPT)
- Prescribing data

Administrative records
- HMRC
- DWP
- DfE

Approvals in place

Environmental & neighbourhood
- Air pollution
- Noise
- Greenspace
- SES & demographic indicators
- Service provision
- Property data
UK Longitudinal Linkage Collaboration

20+ longitudinal studies with >280,000 participants
A new Trusted Research Environment
Novel NHS Digital data pipeline
Approvals for hitherto inaccessible data
Public contributors & application panel
23 approved research projects

AIRWAVE Health Monitoring Study Tissue Bank
ALSPAC: Avon Longitudinal Study of Parents and Children
BCS70 British Cohort Study
Born in Bradford
ELSA: English Longitudinal Study of Ageing
EPIC-Norfolk: The European Prospective Investigation into Cancer (EPIC) Norfolk Study
EXCEED: Extended Cohort for E-health, Environment & DNA
The Fenland Study
Generation Scotland
GLAD: Genetic Links to Anxiety and Depression Study
MCS: Millennium Cohort Study

NCDS 58: 1958 National Child Development Study
Next Steps
NICOLA: Northern Ireland Cohort for the Longitudinal Study of Ageing
NSHD46: MRC National Survey of Health and Development Cohort/1946 Birth Cohort
SABRE: Southall & Brent Revisited
TRACK-COVID Study
TwinsUK
Understanding Society - the UK Household Longitudinal Study
UK LLC: Data Access Overview

- **Stage 1**: Submit expression of interest
  - 1 week

- **Stage 2**: Submit Full Application
  - 2 weeks

- **Stage 3**: Application Review Panel
  - 4 weeks

- **Agreements signed by institution**: 1 week

- **Automated data provision**: 1 week

Submit expression of interest via Health Data Research UK gateway [https://www.healthdatagateway.org/](https://www.healthdatagateway.org/) - search for UK LLC - in the UK LLC record, click on ‘How to access’

- Free to ONS accredited UK Researchers
- Public benefit with public review
- Requirements for reproducible research
- UK LLC distributes applications to contributing studies
- UK LLC has delegated review authority from NHS Digital

- **8 - week Service Level Agreement**
Access to the data

All processing *within* the ISO 27001 accredited Trusted Research Environment

- Access via Windows 10 remote desktop environment using two-factor authentication
- Project-specific database views
- Access to code lists and code mappings
- Standard packages for analyses include - Python, R(RStudio), Stata and SPSS
- Outputs checked for statistical disclosure control before released from the Trusted Research Environment
Looking ahead

- Currently funded until May 2023 - application for new 5-year funding submitted
- Q3 and Q4 2023
  - Opening up to research beyond Covid-19
  - Addition of Welsh NHS data
  - Addition of HMRC data
- Q1 and Q2 2024
  - Addition of Scottish NHS data
  - DWP then DfE to follow

Help/questions
- access@ukllc.ac.uk
Analysing linked HES data

Richard Silverwood
Associate Professor of Statistics &
CLS Chief Statistician

6 December 2022
Outline

1. HES data structure and variables
2. Defining your analysis sample
3. Recent analyses
4. Resources
HES data structure and variables
HES data structure: Episodes and spells

• Each record in HES is a hospital episode.
• Episodes relate to a period of care for a patient under a single consultant within one hospital provider.
• A stay in hospital from admission to discharge is called a spell and can be made up of one or more episodes of care.
# HES data structure: Example (made up) data

<table>
<thead>
<tr>
<th>ncdsid</th>
<th>provspnops</th>
<th>admidate</th>
<th>epiorder</th>
<th>epistart</th>
<th>epiend</th>
<th>disdate</th>
<th>spelend</th>
</tr>
</thead>
<tbody>
<tr>
<td>N3674M</td>
<td>HS6Q9B4HU27D96MF61P5</td>
<td>16jan2010</td>
<td>1</td>
<td>16jan2010</td>
<td>18jan2010</td>
<td>18jan2010</td>
<td>Y</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>07aug2003</td>
<td>1</td>
<td>07aug2003</td>
<td>07aug2003</td>
<td>.</td>
<td>N</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>09nov2015</td>
<td>1</td>
<td>09nov2015</td>
<td>09nov2015</td>
<td>.</td>
<td>N</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>09nov2015</td>
<td>2</td>
<td>09nov2015</td>
<td>14nov2015</td>
<td>14nov2015</td>
<td>Y</td>
</tr>
</tbody>
</table>
HES data structure: Example (made up) data

<table>
<thead>
<tr>
<th>ncdsid</th>
<th>provspnops</th>
<th>Admidate</th>
<th>epiorder</th>
<th>epistart</th>
<th>epiend</th>
<th>disdate</th>
<th>spelend</th>
</tr>
</thead>
<tbody>
<tr>
<td>N3674M</td>
<td>HS6Q9B4HU27D96MF61P5</td>
<td>16jan2010</td>
<td>1</td>
<td>16jan2010</td>
<td>18jan2010</td>
<td>18jan2010</td>
<td>Y</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>07aug2003</td>
<td>1</td>
<td>07aug2003</td>
<td>07aug2003</td>
<td>.</td>
<td>N</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>09nov2015</td>
<td>1</td>
<td>09nov2015</td>
<td>09nov2015</td>
<td>.</td>
<td>N</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>09nov2015</td>
<td>2</td>
<td>09nov2015</td>
<td>14nov2015</td>
<td>14nov2015</td>
<td>Y</td>
</tr>
</tbody>
</table>
Basic data checks

• Variable completeness and quality differs by variable and dataset.

• Important to conduct some basic data checks, e.g.:
  • Dates correctly recorded/formatted (e.g. within plausible range).
  • Ordering of dates is plausible (e.g. within spells).
  • Observed episodes within spell consistent with epiorder.
ICD-10 coding

• International Classification of Diseases version 10 (ICD-10).

• Codifies diseases, signs and symptoms, abnormal findings, complaints, social circumstances, and external causes of injury or diseases.
Hypertensive diseases
(I00-I15)

Excl.: complicating pregnancy, childbirth and the puerperium (O10-O11, O13-O16)

Incl.: 
- involving coronary vessels (I20-I22)
- neonatal hypertension (P29.2)
- pulmonary hypertension (I27.0)

110 Essential (primary) hypertension

Incl.: 
- High blood pressure

Excl.: involving vessels of:
- brain (I60-I69)
- eye (H35.0)

111 Hypertensive heart disease

Incl.: any condition in I50.-, I51.4-I51.9 due to hypertension

111.0 Hypertensive heart disease with (congestive) heart failure

111.9 Hypertensive heart disease without (congestive) heart failure

112 Hypertensive renal disease

Incl.: any condition in N00-N07, N18.-, N19 or N26 due to hypertension
- arteriosclerosis of kidney
- arteriosclerotic nephritis (chronic)(interstitial)
- hypertensive nephropathy
- nephrosclerosis
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Block</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>A00–B99</td>
<td>Certain infectious and parasitic diseases</td>
</tr>
<tr>
<td>II</td>
<td>C00–D48</td>
<td>Neoplasms</td>
</tr>
<tr>
<td>III</td>
<td>D50–D89</td>
<td>Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism</td>
</tr>
<tr>
<td>IV</td>
<td>E00–E90</td>
<td>Endocrine, nutritional and metabolic diseases</td>
</tr>
<tr>
<td>V</td>
<td>F00–F99</td>
<td>Mental and behavioural disorders</td>
</tr>
<tr>
<td>VI</td>
<td>G00–G99</td>
<td>Diseases of the nervous system</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>XXII</td>
<td>U00–U99</td>
<td>Codes for special purposes</td>
</tr>
</tbody>
</table>
ICD-10 coding: Example (made up) data

<table>
<thead>
<tr>
<th>ncdsid</th>
<th>provspnops</th>
<th>D_DIAG_01</th>
<th>D_DIAG_02</th>
<th>D_DIAG_03</th>
<th>D_DIAG_04</th>
<th>...</th>
<th>D_DIAG_20</th>
</tr>
</thead>
<tbody>
<tr>
<td>N3674M</td>
<td>HS6Q9B4HU27D96MF61P5</td>
<td>I77</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>Z85</td>
<td>G35</td>
<td>Z51</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>Z85</td>
<td>G35</td>
<td>Z51</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>I42</td>
<td>K29</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>I42</td>
<td>K29</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>I42</td>
<td>K29</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>R06</td>
<td>Z87</td>
<td>K12</td>
<td>N23</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>R06</td>
<td>Z87</td>
<td>K12</td>
<td>N23</td>
<td>...</td>
<td>.</td>
</tr>
</tbody>
</table>
# ICD-10 coding: Example (made up) data

<table>
<thead>
<tr>
<th>ncdsid</th>
<th>provspnops</th>
<th>D_DIAG_01</th>
<th>D_DIAG_02</th>
<th>D_DIAG_03</th>
<th>D_DIAG_04</th>
<th>…</th>
<th>D_DIAG_20</th>
</tr>
</thead>
<tbody>
<tr>
<td>N3674M</td>
<td>HS6Q9B4HU27D96MF61P5</td>
<td>I77</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>Z85</td>
<td>G35</td>
<td>Z51</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>Z85</td>
<td>G35</td>
<td>Z51</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>I42</td>
<td>K29</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>I42</td>
<td>K29</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>I42</td>
<td>K29</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>R06</td>
<td>Z87</td>
<td>K12</td>
<td>N23</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>R06</td>
<td>Z87</td>
<td>K12</td>
<td>N23</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

**Notes:**
- Other disorders of arteries and arterioles
- Personal history of malignant neoplasm
- Multiple sclerosis
- Other medical care
ICD-10 coding: Example (made up) data

<table>
<thead>
<tr>
<th>ncdsid</th>
<th>provspnops</th>
<th>D_DIAG_01</th>
<th>D_DIAG_02</th>
<th>D_DIAG_03</th>
<th>D_DIAG_04</th>
<th>...</th>
<th>D_DIAG_20</th>
</tr>
</thead>
<tbody>
<tr>
<td>N3674M</td>
<td>HS6Q9B4HU27D96MF61P5</td>
<td>I77</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>Z85</td>
<td>G35</td>
<td>Z51</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>Z85</td>
<td>G35</td>
<td>Z51</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>I42</td>
<td>K29</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>I42</td>
<td>K29</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>I42</td>
<td>K29</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>R06</td>
<td>Z87</td>
<td>K12</td>
<td>N23</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>R06</td>
<td>Z87</td>
<td>K12</td>
<td>N23</td>
<td>...</td>
<td>.</td>
</tr>
</tbody>
</table>
ICD-10 coding: Example code

May want to group codes, e.g. into chapters:

```
. gen D_DIAG_01_ch = .

. replace D_DIAG_01_ch = 1 if substr(D_DIAG_01, 1, 1)=="A" | substr(D_DIAG_01, 1, 1)=="B"

. replace D_DIAG_01_ch = 2 if substr(D_DIAG_01, 1, 1)=="C" | substr(D_DIAG_01, 1, 2)=="D0" | substr(D_DIAG_01, 1, 2)=="D1" | substr(D_DIAG_01, 1, 2)=="D2" | substr(D_DIAG_01, 1, 2)=="D3" | substr(D_DIAG_01, 1, 2)=="D4"

...  
```

Expand to all diagnosis positions by looping over D_DIAG_01, ..., D_DIAG_20.
Often interested in whether an individual has ever had a certain diagnosis. E.g. for Chapter I:

. gen D_DIAG_01_ch1 = (D_DIAG_01_ch==1)
. gen D_DIAG_01_ch1_ever = D_DIAG_01_ch1
. bysort ncdsid (D_DIAG_01_ch1_ever): replace D_DIAG_01_ch1_ever = D_DIAG_01_ch1_ever[_N]

Extend to all chapters by looping over Chapters 1-22
OPCS coding

- Office of Population Censuses and Surveys (OPCS) Classification of Interventions and Procedures.
- Codifies operations, procedures and interventions performed during in-patient stays, day case surgery and some out-patient treatments in NHS hospitals.
### OPCS coding: Chapters

<table>
<thead>
<tr>
<th>Chapter/first character*</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Nervous System</td>
</tr>
<tr>
<td>B</td>
<td>Endocrine System and Breast</td>
</tr>
<tr>
<td>C</td>
<td>Eye</td>
</tr>
<tr>
<td>D</td>
<td>Ear</td>
</tr>
<tr>
<td>E</td>
<td>Respiratory Tract</td>
</tr>
<tr>
<td>F</td>
<td>Mouth</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td>Z</td>
<td>Subsidiary Classification of Sites of Operation</td>
</tr>
</tbody>
</table>

*Codes beginning with an “O” are overflow codes and should be handled differently.*
## OPCS coding: Example (made up) data

<table>
<thead>
<tr>
<th>ncdsid</th>
<th>provspnops</th>
<th>opertn_01</th>
<th>opertn_02</th>
<th>opertn_03</th>
<th>opertn_04</th>
<th>...</th>
<th>opertn_23</th>
</tr>
</thead>
<tbody>
<tr>
<td>N3674M</td>
<td>HS6Q9B4HU27D96MF61P5</td>
<td>Z815</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>U212</td>
<td>W581</td>
<td>.</td>
<td>.</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>Y534</td>
<td>Z664</td>
<td>G451</td>
<td>.</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>T202</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
<td>.</td>
</tr>
</tbody>
</table>
### OPCS coding: Example (made up) data

<table>
<thead>
<tr>
<th>ncdsid</th>
<th>provspnops</th>
<th>opern_01</th>
<th>opern_02</th>
<th>opern_03</th>
<th>opern_04</th>
<th>...</th>
<th>opern_23</th>
</tr>
</thead>
<tbody>
<tr>
<td>N3674M</td>
<td>HS6Q9B4HU27D96MF61P5</td>
<td>Z815</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>U212</td>
<td>W581</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>Y534</td>
<td>Z664</td>
<td>G451</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>T202</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>...</td>
<td>.</td>
</tr>
</tbody>
</table>

**Computed tomography NEC**

**Primary resurfacing arthroplasty of joint**
OPCS coding: Example (made up) data

<table>
<thead>
<tr>
<th>ncdsid</th>
<th>provspnops</th>
<th>opertn_01</th>
<th>opertn_02</th>
<th>opertn_03</th>
<th>opertn_04</th>
<th>…</th>
<th>opertn_23</th>
</tr>
</thead>
<tbody>
<tr>
<td>N3674M</td>
<td>HS6Q9B4HU27D96MF61P5</td>
<td>Z815</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>…</td>
<td>.</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>U212</td>
<td>W581</td>
<td>.</td>
<td>.</td>
<td>…</td>
<td>.</td>
</tr>
<tr>
<td>N8429G</td>
<td>JS8G36DJV8538JSBDNH8</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>…</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>…</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>Y534</td>
<td>Z664</td>
<td>G451</td>
<td>.</td>
<td>…</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>H836FTK2SN9JTEB533N2</td>
<td>T202</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>…</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>…</td>
<td>.</td>
</tr>
<tr>
<td>N4275R</td>
<td>64NQF48J6D2UH76WN382</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>…</td>
<td>.</td>
</tr>
</tbody>
</table>
Defining your analysis sample
Defining your analysis sample

- Cohort members may legitimately not have a HES record over the period of HES data availability, so don’t usually want to analyses only those with linked HES records.
- If we believe that any cohort member with a HES record would have been successfully matched to it, we can conclude that any eligible cohort members without a matched HES record truly had no HES record.
- In order to do this, we need to consider who is eligible for linkage.
- Linkage consent certainly a condition for eligibility.
Defining your analysis sample

- HES data only relate to English (NHS) hospitals, so restrict eligibility to cohort members living in England over the relevant period?
- E.g. In analyses of linked NCDS-HES data we have considered cohort members to be eligible for linkage if they reported living in England at any one or more waves between waves 6 (2000, age 42) and 9 (2013, age 55).
- But eligibility definition may differ depending on context.

Recent analyses
Recent analyses: Mental health and hospital interactions in young people

• Nasir Rajah, Emla Fitzsimons.

• Aimed to investigate association between poor mental health in young people and subsequent hospital service use.

• Used linked Next Steps-HES data.

• Next Steps: self-reported mental health and self-harm.

• HES: any/number of A&E visits (A&E), diagnosis of mental or behavioural disorder (APC).
Recent analyses: Mental health and hospital interactions in young people

- Poor self-reported mental health at age 17 associated with:
  
  i. 6 percentage point (95% CI 2, 10) greater probability of an A&E appointment in the following decade

  ii. 6 percentage point (95% CI 2, 9) greater probability of a diagnosis of mental or behavioural disorder in hospital.

- Conclude that there may be associations between young people’s poorer self-reported mental health and increased hospital interactions.

- Important implications for future research on demand and allocation of healthcare resources in England.
Recent analyses: Using linked HES data to aid the handling of cohort study non-response

- Nasir Rajah, Lisa Calderwood, Bianca L De Stavola, Katie Harron, George Ploubidis, Richard Silverwood.
- Growing interest in whether linked administrative data can aid analyses of cohort data subject to missingness.
- Used linked NCDS-HES data.
- Derived 58 variables from HES data (numbers of admissions and appointments, missed appointments, investigations undertaken, diagnoses and treatments received).
- Applied LASSO variable selection approach to identify HES variables predictive of non-response at age 55 sweep of NCDS.
Recent analyses: Using linked HES data to aid the handling of cohort study non-response

- Identified 10 such variables, e.g. cohort members treated for adult mental illness >70% more likely to be non-respondents (risk ratio 1.73; 95% CI 1.17, 2.51).

- Included these variables as auxiliary variables in multiple imputation analyses to explore the extent to which they helped restore sample representativeness.

- However, only helped to a limited extent. Essentially no additional gain relative to using only previously identified survey predictors of non-response (i.e. NCDS rather than HES variables).
Resources
Resources: User Guides  [https://cls.ucl.ac.uk/data_documentation]

National Child Development Study
Linked health administrative datasets – Hospital Episode Statistics (HES)
User guide
First edition, December 2020

1970 British Cohort Study
Linked health administrative datasets – Hospital Episode Statistics (HES)
User guide
First edition

Institute of Education
Next Steps
Linked health administrative datasets – Hospital Episode Statistics (HES)
User Guide (Version 1)
September 2020

Millennium Cohort Study
Linked health administrative datasets - Scottish Medical Records (SMR)
User guide
First edition, October 2020
Resources: Quality and representativeness of linked NCDS-HES data

• Examined associations between key cohort member sociodemographic characteristics and successful linkage.

• Compared the levels of successful linkage within strata of NCDS variables which may be expected to be associated with hospital attendance, and hence with successful HES linkage.

https://cls.ucl.ac.uk/working_papers/
Resources: Quality and representativeness of linked NCDS-HES data

- Evaluated the population representativeness of the linked sample using external data (hospital admission rates in the general population).

- Findings suggest that the linkage quality of the NCDS-HES data is high and that the linked sample maintains an excellent level of population representativeness.

https://cls.ucl.ac.uk/working_papers/
Other resources


Thank you.