

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

6 December 2022

CENTRE FOR  
LONGITUDINAL  
STUDIES



Economic  
and Social  
Research Council

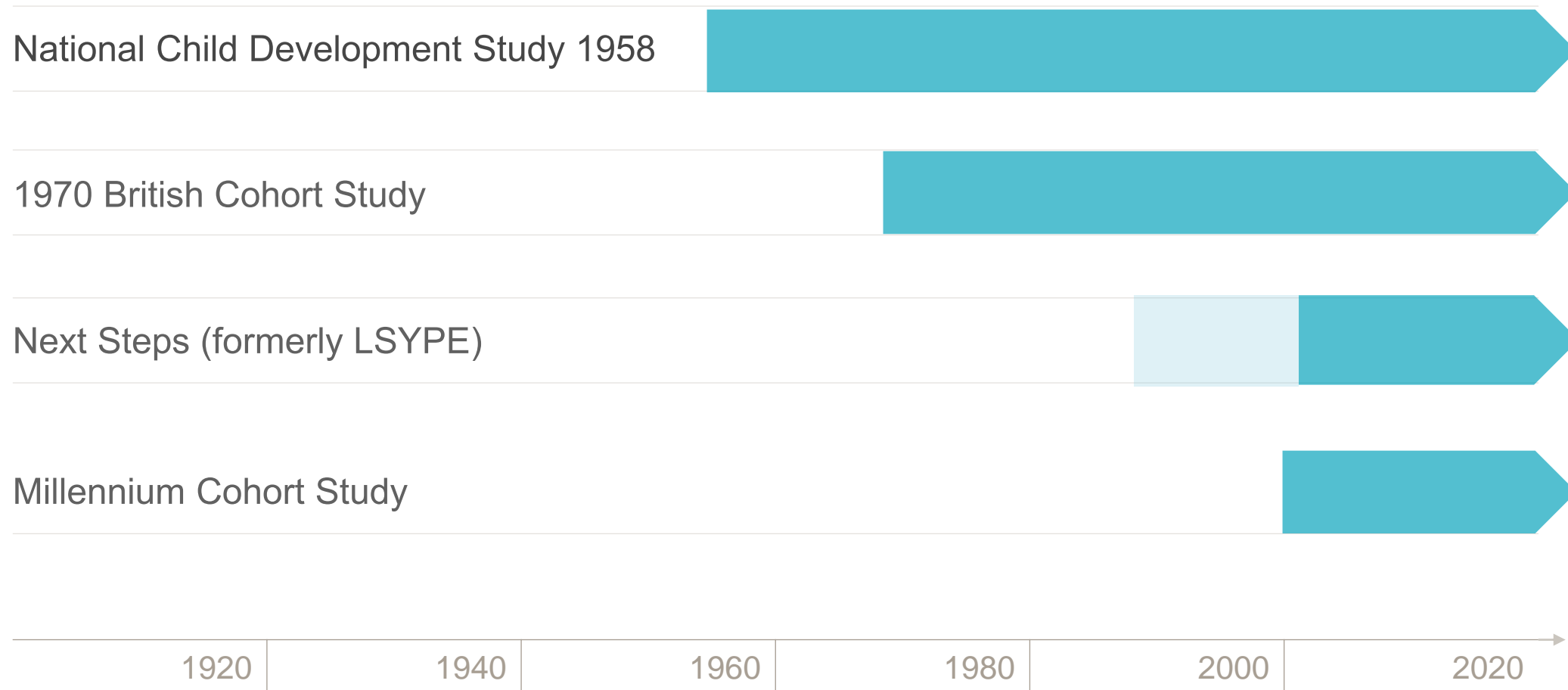
# Housekeeping

- 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](mailto: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.

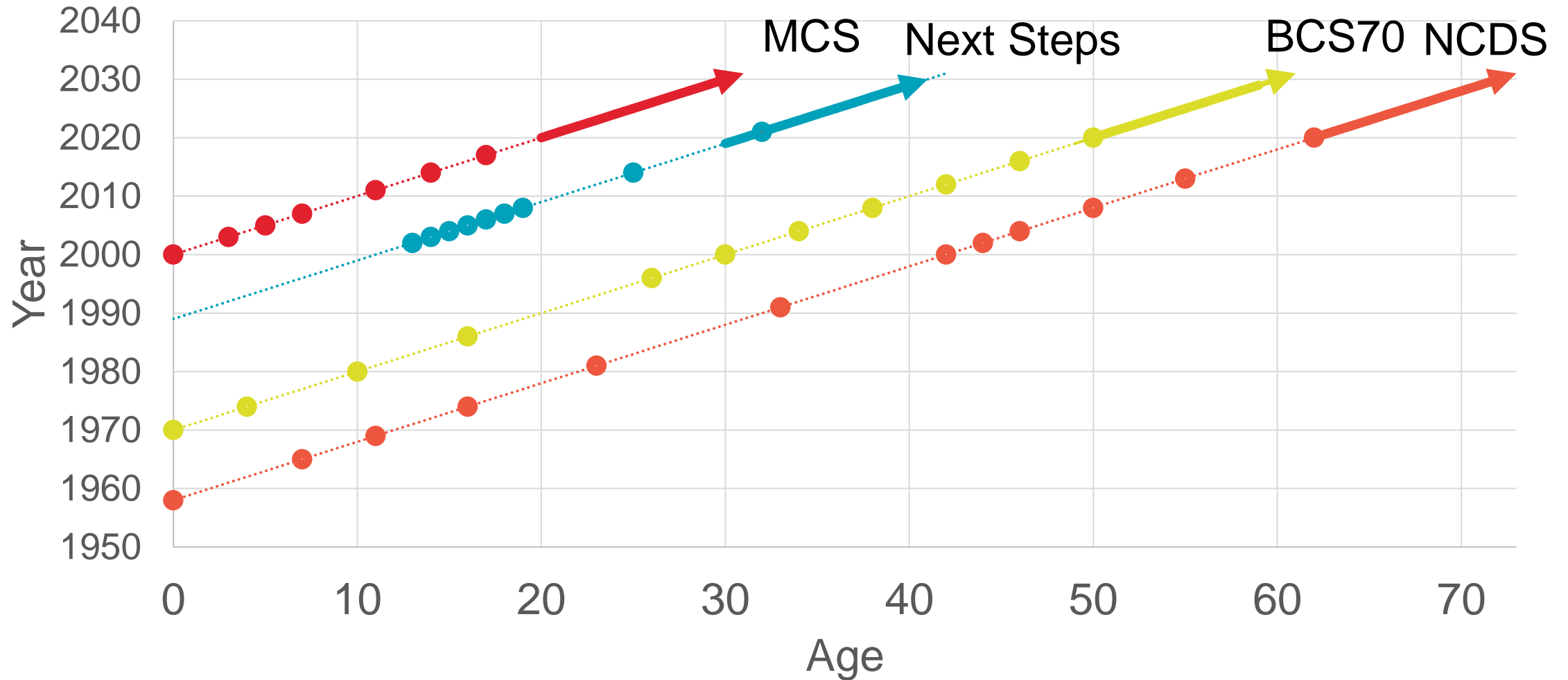
# Today's schedule

Time	Topics covered	Speaker
1.00 – 1.15pm	Brief introduction to the cohorts	Richard Silverwood Associate Professor of Statistics
1.15 – 1.30pm	HES data and linkage overview	Sarah Kerry-Barnard Research Data Manager
1.30 – 1.45pm	Accessing linked HES datasets	Danielle Gomes and Karen Dennison Record Linkage Managers
1.45 – 2.15pm	Analysing linked HES data	Richard Silverwood Associate Professor of Statistics
2.15 – 2.30 pm	Q&A	All




# CLS studies follow people across life



# Study timelines



# Typical information covered

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

# 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

# Hospital Episodes Statistics User Guide(s)

- Covers everything in this webinar and more
- Available publicly from UK Data Service (UKDS) and CLS websites for NCDS, BCS70 and Next Steps

## **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

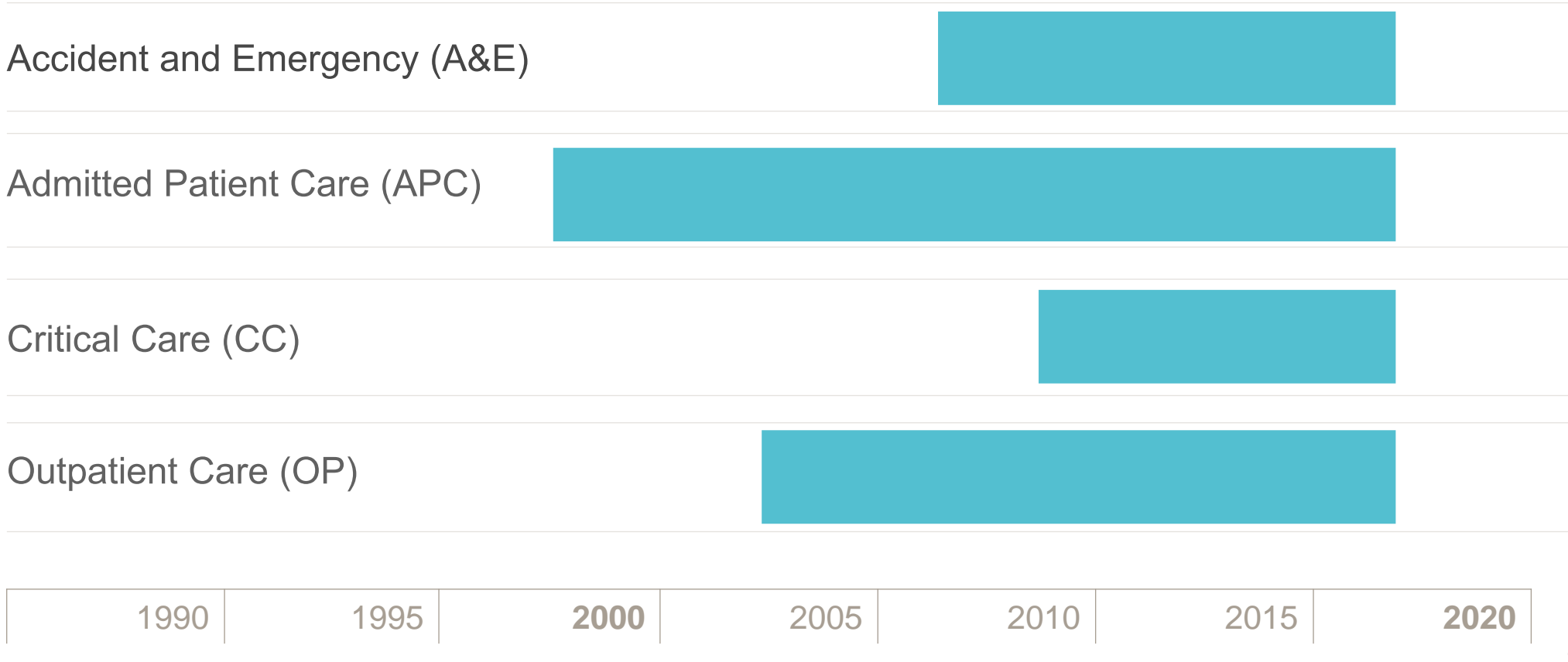
Modifications to the Outpatient Care Data

# NHS England – Hospital Episode Statistics (HES)

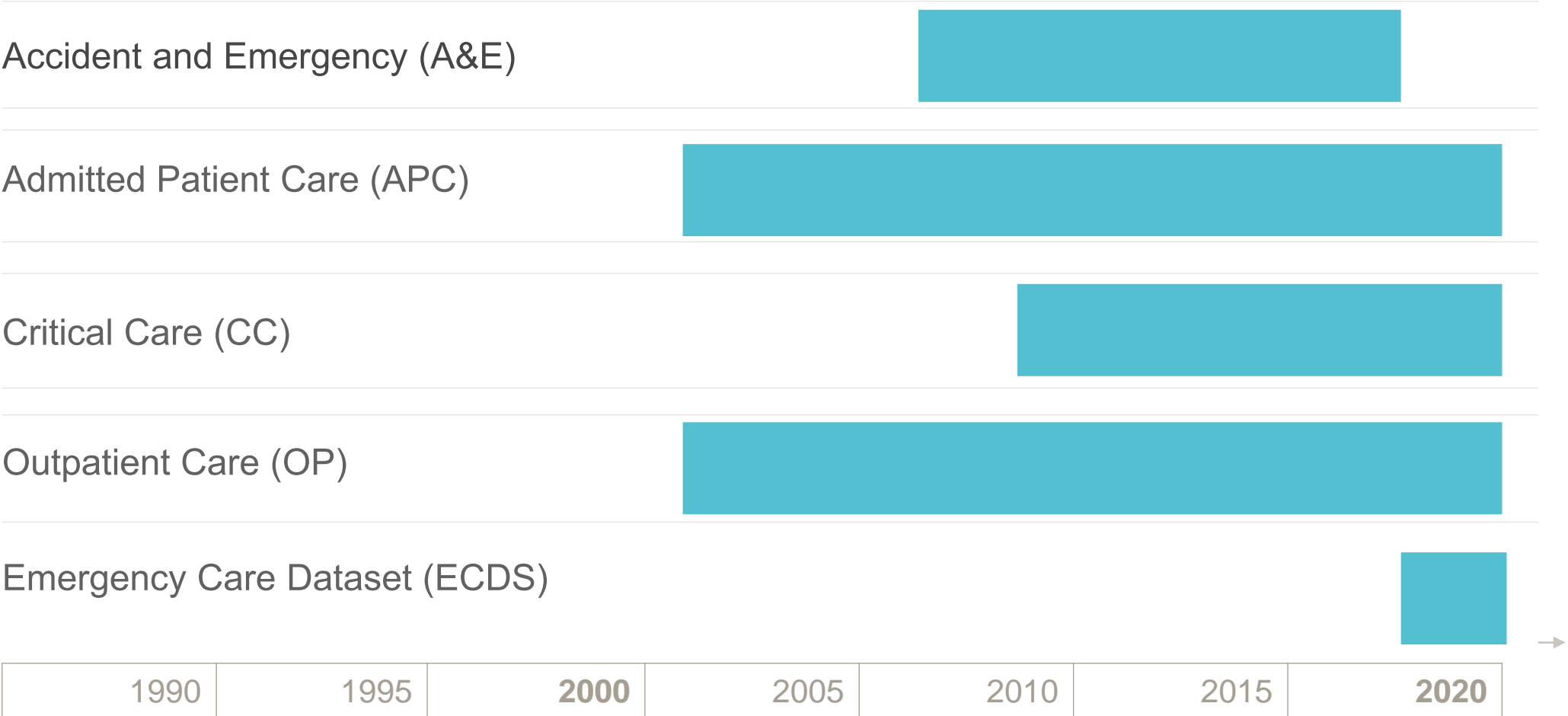
## Datasets available:

- 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



# Hospital Episode Statistics over time – MCS



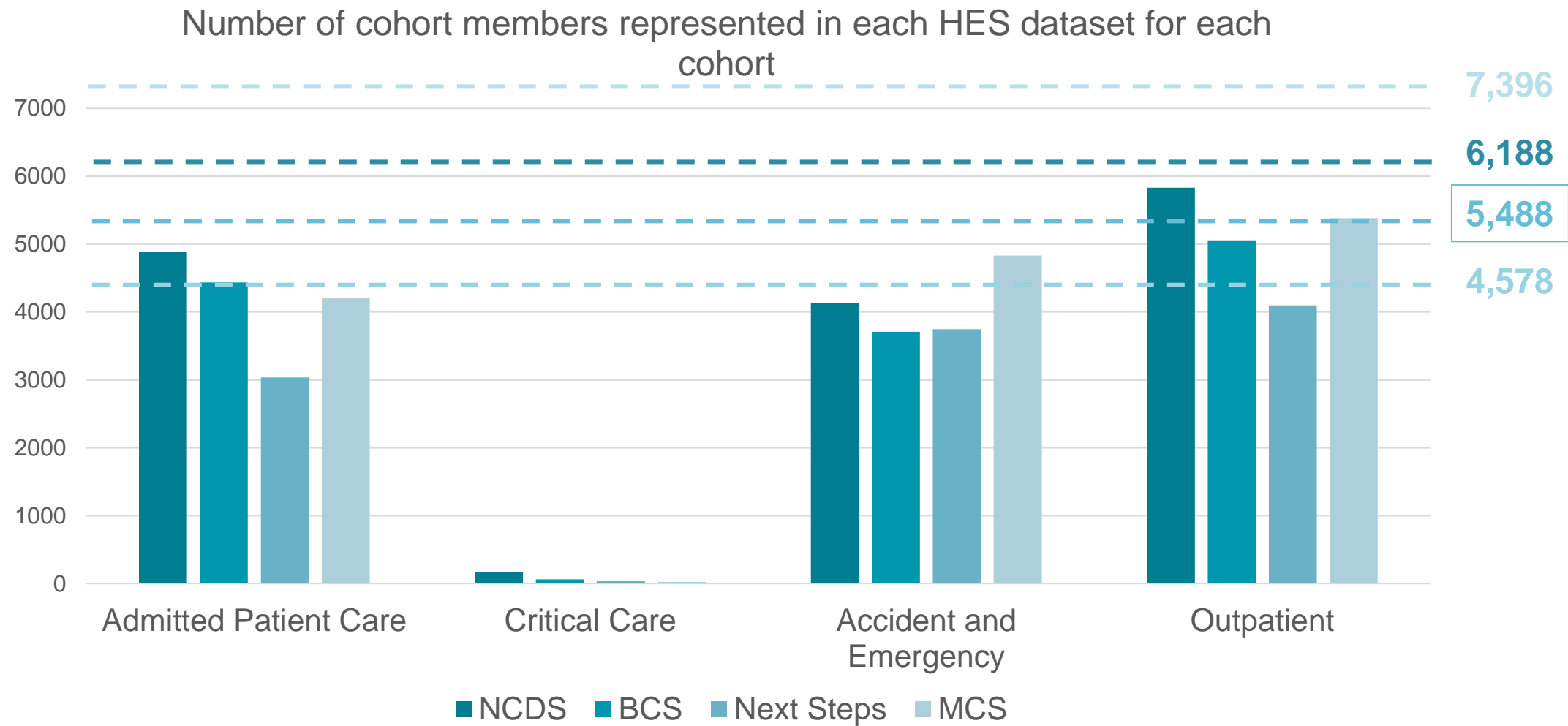
# HES - Linkage

	Number in sweep*	Gave consent (% of in sweep)	Has matched data (% of consented)
Next Steps	7,707	4,895 (63.5)	4,579 (93.5)
BCS70	9,841	7,048 (71.6)	5,488 (77.9)
NCDS	9,790	7,065 (72.2)	6,188 (87.6)
MCS	10,757	9,214 (86.7)	7,396 (81.5)

\* 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





# HES – Structure of the data

## 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

### ▼ ICD-10 Version:2010

- ▶ I Certain infectious and parasitic 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
- ▶ IX Diseases of the circulatory system
- ▶ X Diseases of the respiratory system
- ▶ XI Diseases of the digestive system
- ▶ XII Diseases of the skin and subcutaneous tissue
- ▶ XIII Diseases of the musculoskeletal system and connective tissue
- ▶ XIV Diseases of the genitourinary system
- ▶ XV Pregnancy, childbirth and the puerperium
- ▶ XVI Certain conditions originating in the perinatal period
- ▶ XVII Congenital malformations, deformations and chromosomal abnormalities
- ▶ XVIII Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified
- ▶ XIX Injury, poisoning and certain other consequences of external causes
- ▶ XX External causes of morbidity and mortality
- ▶ XXI Factors influencing health status and contact with health services
- ▶ XXII Codes for special purposes

## International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) Version for 2010

### 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] ([O98.-](#))  
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-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

ICD-10 Version:2010

Search

[ Advanced Search ]

ICD-10

Versions - Languages

Info

ICD-10 Version:2010

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 Helminthiasis

B85-B89 Pediculosis, acariasis 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

Excl.: congenital tuberculosis ([A97.0](#))

human immunodeficiency [HIV] disease resulting in tuberculosis ([B20.0](#))

pneumoconiosis associated with tuberculosis ([J65](#))

sequelae of tuberculosis ([B90.-](#))

silicotuberculosis ([J65](#))

A15 Respiratory tuberculosis, bacteriologically and histologically confirmed

A15.0 Tuberculosis of lung, confirmed by sputum microscopy with or without culture

Tuberculous:

• bronchiectasis

• 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

Excl.: specified as primary ([A15.7](#))

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

Tuberculosis of:

• bronchus

• glottis

• larynx

• trachea

confirmed bacteriologically and histologically

A15.6 Tuberculous pleurisy, confirmed bacteriologically and histologically

Tuberculosis of pleura

Tuberculous empyema

confirmed bacteriologically and histologically

# NHS Scotland – Scottish Medical Records (**SMR**)



# NHS Scotland – Scottish Medical Data (SMR)

Dataset	NCDS	BCS70	MCS
Prescribing Information System (PIS)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Outpatient Records (SMR00)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Inpatient Records (SMR01)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Maternity Records (SMR02)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Birth and Neonatal Records (SMR11)			<input checked="" type="checkbox"/>
Scottish Immunisation & Recall System (SIRS)			<input checked="" type="checkbox"/>
Child Health Reviews (CHR) 1 <sup>st</sup> visit / 6-8 week visit / 8-9 month visit / 22-24 month visit / 37-42 month visit / Preschool visit / School reviews			<input checked="" type="checkbox"/>

# Accessing Health Data

UK Data Service (UKDS)

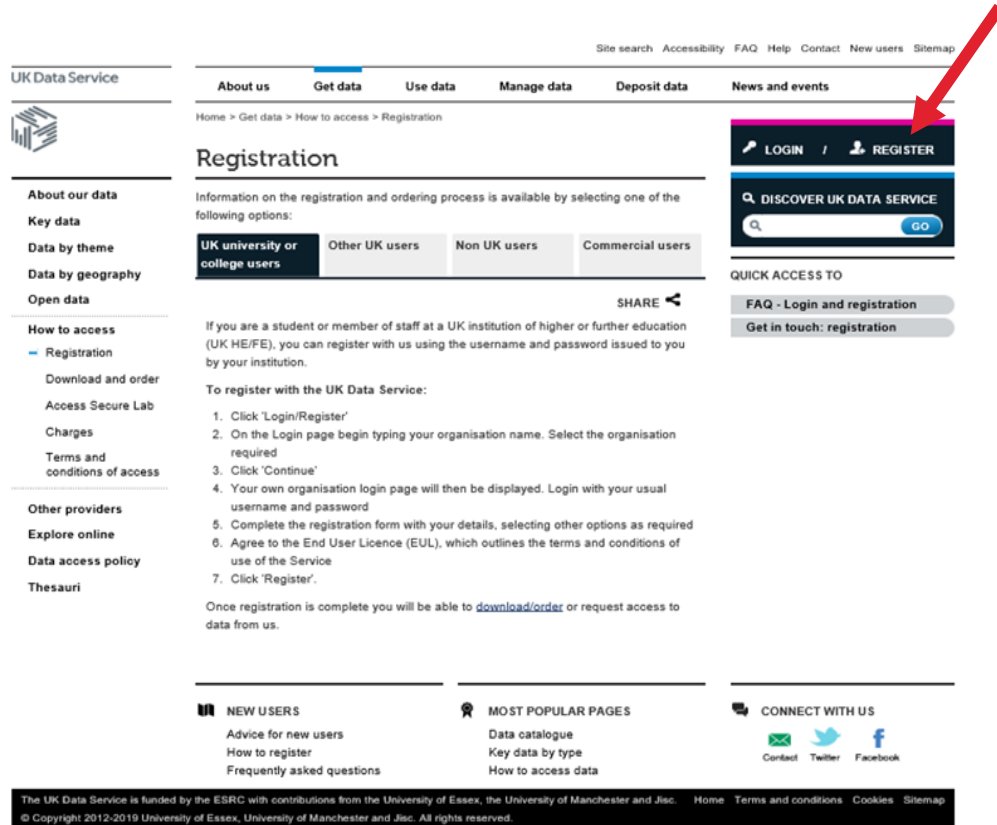
UK Longitudinal Linkage Collaboration (UKLLC)

# 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



The screenshot shows the UK Data Service website. The navigation bar includes links for About us, Get data, Use data, Manage data, Deposit data, and News and events. The 'Get data' link is highlighted. Below the navigation bar, the 'Registration' page is displayed. A red arrow points to the 'REGISTER' button in the top right corner of the page. The page content includes information about the registration process and a list of steps to follow.

UK Data Service

Site search Accessibility FAQ Help Contact New users Sitemap

About us Get data Use data Manage data Deposit data News and events

Home > Get data > How to access > Registration

### Registration

Information on the registration and ordering process is available by selecting one of the following options:

UK university or college users Other UK users Non UK users Commercial users

SHARE

If you are a student or member of staff at a UK institution of higher or further education (UK HE/FE), you can register with us using the username and password issued to you by your institution.

To register with the UK Data Service:

1. Click 'Login/Register'
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 from us.

NEW USERS

- Advice for new users
- How to register
- Frequently asked questions

MOST POPULAR PAGES

- Data catalogue
- Key data by type
- How to access data

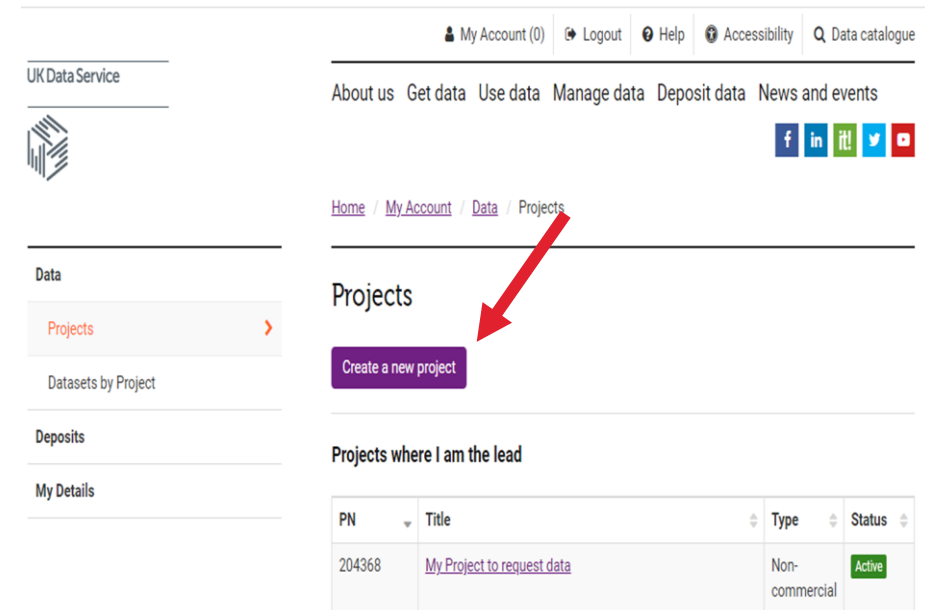
CONNECT WITH US

Contact Twitter Facebook

The UK Data Service is funded by the ESRC with contributions from the University of Essex, the University of Manchester and Jisc. Home Terms and conditions Cookies Sitemap

© Copyright 2012-2019 University of Essex, University of Manchester and Jisc. All rights reserved.

## Step 2: Create a project



The screenshot shows the UK Data Service website. The navigation bar includes links for About us, Get data, Use data, Manage data, Deposit data, and News and events. The 'Use data' link is highlighted. Below the navigation bar, the 'Projects' page is displayed. A red arrow points to the 'Create a new project' button. The page content includes a table of projects where the user is the lead.

UK Data Service

My Account (0) Logout Help Accessibility Data catalogue

About us Get data Use data Manage data Deposit data News and events

Home / My Account / Data / Projects

### Projects

Create a new project

Projects where I am the lead

PN	Title	Type	Status
204368	<a href="#">My Project to request data</a>	Non-commercial	Active

# To apply to use the data:

## Step 3: Select your data

Studies

Series

http://doi.org/10.5255/UKDA-SN-8733-1

Copy study DOI

### 1970 British Cohort Study: Linked Health Administrative Datasets (Hospital Episode Statistics), England, 1997-2017: Secure Access

Details

Documentation

Resources

Access 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.

Use of the data requires the Lead Researcher to request a one-time bespoke subset of variables per project, for which approval is requested from the data owner or their nominee. If approved, the researcher(s) will be able to access the bespoke dataset via their Secure Lab project area.

Approved users must complete specialist training.

Users must have a UK HE or FE affiliation and must be based in the UK when accessing data

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 and

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 webpage](#)).

Add these data to your account. (Only available via Secure Lab remote access)

Add to account

## Step 4: Assign dataset to a project

UK Data Service

Data

Projects

Datasets by Project

Deposits

My Details

My Account (1) Logout Help Accessibility Data catalogue

About us Get data Use data Manage data Deposit data News and events

f in itl t y

Home / My Account / Data / Assign To Project

### Assign datasets to project

SN	Dataset	Status	
8733	1970 British Cohort Study: Linked Health Administrative Datasets (Hospital Episode Statistics), England, 1997-2017: Secure Access	Ready to assign	<input checked="" type="checkbox"/>

Create a new project

Add to an existing project

Project: \*

My Project to request data

Only checked datasets will be added to this project.

Add to project

Cancel

Existing datasets in project

# To apply to use the data

## Step 5: Access the appropriate forms

My Project to request data

Project Datasets Members Notes Log

Datasets in project

Select all datasets

SN	Dataset	Status		
8733	1970 British Cohort Study: Linked Health Administrative Datasets (Hospital Episode Statistics), England, 1997-2017: Secure Access	<div>⚠ Request access</div>	Actions ▼	<input type="checkbox"/>
<div><div><input checked="" type="checkbox"/></div>Accept standard EUL</div> <div><div>⚠</div>Complete Research Project Application documentation</div> <div><div>⚠</div>Complete ESRC Accredited Researcher Application</div> <div><div>⚠</div>Complete Secure Access User Agreement</div> <div><div>🕒</div>Depositor approval</div> <div><div>🕒</div>Training</div> <div><div>🕒</div>Secure Lab access</div>				
<div>Complete actions</div>				

# To apply to use the data

## Step 5: Download Application forms

Complete actions for access

Introduction

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 Research Project Application documentation

The Project Lead must download the zip file below and complete all three forms within the zip file. Please read the cover sheet of the Project Application form included in the zip first.

Project members will need to enter into an additional access agreement with CLS as part of the application, on a per organisation/institution basis. Any queries regarding this additional agreement should be sent to [clsfeedback@ucl.ac.uk](mailto:clsfeedback@ucl.ac.uk).

The completed set of forms should be returned by email to [secure.applications@ukdataservice.ac.uk](mailto:secure.applications@ukdataservice.ac.uk).

[BCS70 HES Linked Application](#)

First downloaded on 18/01/2021.

Your changes will be saved automatically.

Close

Back

Next

## 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 )



**Secure Access Research Project  
Application (NHS-D)**

Public  
16 November 2020  
Version: 1.00


T +44 (0)1206 872143  
E [secure.applications@ukdataservice.ac.uk](mailto:secure.applications@ukdataservice.ac.uk)  
W [ukdataservice.ac.uk](http://ukdataservice.ac.uk)

- **Completed by :** Project Lead

**ESRC ACCREDITED RESEARCHER FORM**  
Accredited Researcher application form (Secure Lab access)

**PUBLIC**  
3 DECEMBER 2014  
Version: 07.00

T +44 (0)1206 872001  
E [kdenn@data-archive.ac.uk](mailto:kdenn@data-archive.ac.uk)



**UK DATA ARCHIVE**  
UNIVERSITY OF ESSEX  
WIVENHOE PARK  
COLCHESTER  
ESSEX, CO4 3SQ

**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)

UK Data Service



### Secure Access User Agreement

External  
13 October 2016  
Version: 09.00

T +44 (0)1206 873546  
E collections@ukdataservice.ac.uk  
ukdataservice.ac.uk

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)

	A	B	C	D	E	F
1	<b>Next Steps - Accident &amp; Emergency - Data Dictionary</b>			Please select the variables you need in the 'Requested' drop down cells		
2						
3	<b>var_order</b>	<b>variable_name</b>	<b>variable_label</b>	<b>value_labels</b>	<b>Requested ? (YES/blank)</b>	
4	1	NSID	NSID - cohort member identifier		Yes	
5	2	AEARRIVALMODE	Arrival mode	(-1) "Missing: not applicable/not available/not given" (1) "Brought in by ambulance (including helicopter)" (2) "Other"		
6	3	AEATTEND_EXC_PLANNED	Attendances excluding planned	(0.0) "Planned AE Attendance " (1.0) "Unplanned AE Attendance"		
7	4	AEATTENDCAT	Attendance category	(-1) "Missing: not applicable/not available/not given" (1) "First Accident and Emergency attendance" (2) "Follow-up Accident and Emergency attendance - planned" (3) "Follow-up Accident and Emergency attendance - unplanned" (9) "Not known".		
8	5	AEATTENDDISP	Attendance disposal	(1)"Admitted to hospital bed at same HCP" (2)"Discharged follow-up treatment to be provided by GP" (3)"Discharged did not require any follow-up treatment" (4)"Referred to A&E clinic" (5)"Referred to fracture clinic" (6)"Referred to other outpatient clinic" (7)"Transferred to other healthcare provider" (10)"Died in department" (11)"Referred to other healthcare professional" (12)"Left department before being treated" (13)"Left department having refused treatment" (14)"Other" (99)"Not known"		
9	6	AEDEPTTYPE	Department type	(-1) "Missing: not applicable/not available/not given" (1) "Emergency departments" (2) "Consultant-led mono specialty accident and emergency service" (3) "Other type of A&E/ minor injury activity with designated accommodation for reception of patients" (4) "NHS walk-in centres" (99) "Not known"		
10	7	AEINCLOCTYPE	Incident location type	(-1) "Missing: not applicable/not available/not given" (10) "Home" (40) "Work" (50) "Educational establishment" (60) "Public Place" (91) "Other" (99) "Not known"		
11	8	AEPATGROUP	Patient group	(-1) "Missing: not applicable/not available/not given" (10) "Road traffic accident" (20) "Assault" (30) "Deliberate self-harm" (40) "Sports injury" (50) "Firework injury" (60) "Other accident" (70) "Brought in dead" (80) "Other than above" (99) "Not known"		
12	9	AEREFSOURCE	Source of referral for A&E	(-1) "Missing: not applicable/not available/not given" (0) "General medical practitioner" (1) "Self referral" (2) "Local authority social services" (3) "Emergency services" (4) "Work" (5.0) "Educational establishment" (6) "Police " (7.0) "Health care provider: same or other" (8.0) "Other " (92.0) "General dental practitioner" (93.0) "Community dental service"		
13	10	ARRIVALDATE	Arrival start date			
14	11	ARRIVALTIME	Arrival Time HH:MM (00:00 to 25:59)			
15	12	CARERSI	Carer support indicator	(-1) "Missing: not applicable/not available/not given" (1) "Yes" (2) "No"		
16	13	CONCLDUR	Duration to conclusion (in minutes)	(-1) "Missing: not applicable/not available/not given"		
17	14	CONCLTIME	Conclusion time HH:MM (00:00 to 25:59)			
18	15	DEPDUR	Duration to departure (in minutes)	(-1) "Missing: not applicable/not available/not given"		
19	16	DEPTIME	Departure time HH:MM (00:00 to 25:59)			
20	17	DIAG_01	A&E diagnosis (1)	he A&E diagnosis is a six character code made up of, diagnosis condition (n2), sub-analysis (n1), anatomical area (n2) and anatomical side (an1). (-1) "Missing: not applicable/not available/not given"		
21	18	DIAG_02	A&E diagnosis (2)	he A&E diagnosis is a six character code made up of, diagnosis condition (n2), sub-analysis (n1), anatomical area (n2) and anatomical side (an1). (-1) "Missing: not applicable/not available/not given"		
22	19	DIAG_03	A&E diagnosis (3)	he A&E diagnosis is a six character code made up of, diagnosis condition (n2), sub-analysis (n1), anatomical area (n2) and anatomical side (an1). (-1) "Missing: not applicable/not available/not given"		

# To apply to use the data

## Step 5: Application forms- Additional agreement required (HES data only)

CENTRE FOR LONGITUDINAL STUDIES



---

**CLS Licence Agreement for Linked NHS-Digital Data**

30<sup>th</sup> May 2022  
Version: 17.00-  
Public

Email: [clsrecordslinkage@ucl.ac.uk](mailto:clsrecordslinkage@ucl.ac.uk)  
Website: <https://cls.ucl.ac.uk/>

CENTRE FOR LONGITUDINAL STUDIES



---

**CLS Licence Agreement for Linked NHS-Digital Data**  
**UCL users only**

30<sup>th</sup> May 2022  
Version 04  
Public

Email: [clsrecordslinkage@ucl.ac.uk](mailto:clsrecordslinkage@ucl.ac.uk)  
Website: <https://cls.ucl.ac.uk/>

# To apply to use the data

Step 5: Application forms- Additional agreement required (HES data only)

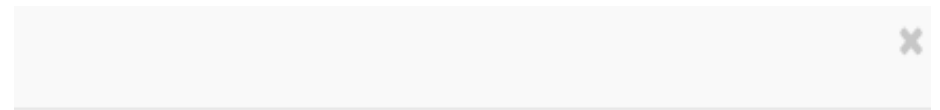
**Important sections in the document: These are NHS Digital requirements.**

- ✓ 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.



### Depositor approval required

This dataset requires approval from the depositor before it can be accessed. Please wait while we service this request. You will be notified by email when approval has been granted.



In progress with the UK Data Service.

# Access to secure data

## Training 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.


## 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


 ESRC Accredited Researcher Proposal form

 ESRC Accredited Researcher Application form

 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.

Your changes will be saved automatically.

Close

Back

Next

# 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](mailto: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?

CENTRE FOR  
LONGITUDINAL  
STUDIES



Economic  
and Social  
Research Council



# The UK Longitudinal Linkage Collaboration:

a research resource for the longitudinal research community



UK LLC led by the Universities of Bristol & Edinburgh



THE UNIVERSITY  
of EDINBURGH

Funded by



UK Research  
and Innovation

In collaboration with



UNIVERSITY OF  
LEICESTER



DATA MIND  
The Health Data Research Hub for Mental Health

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



BRADFORD INSTITUTE  
FOR HEALTH RESEARCH  
| MAKING RESEARCH REAL



University of the  
West of England



THE UNIVERSITY  
of EDINBURGH



University  
of Glasgow



NICE National Institute for  
Health and Care Excellence

# COVID-19 Longitudinal Health and Wellbeing National Core Study



## A globally unique cross-cutting resource

### 20+ Longitudinal Population Studies

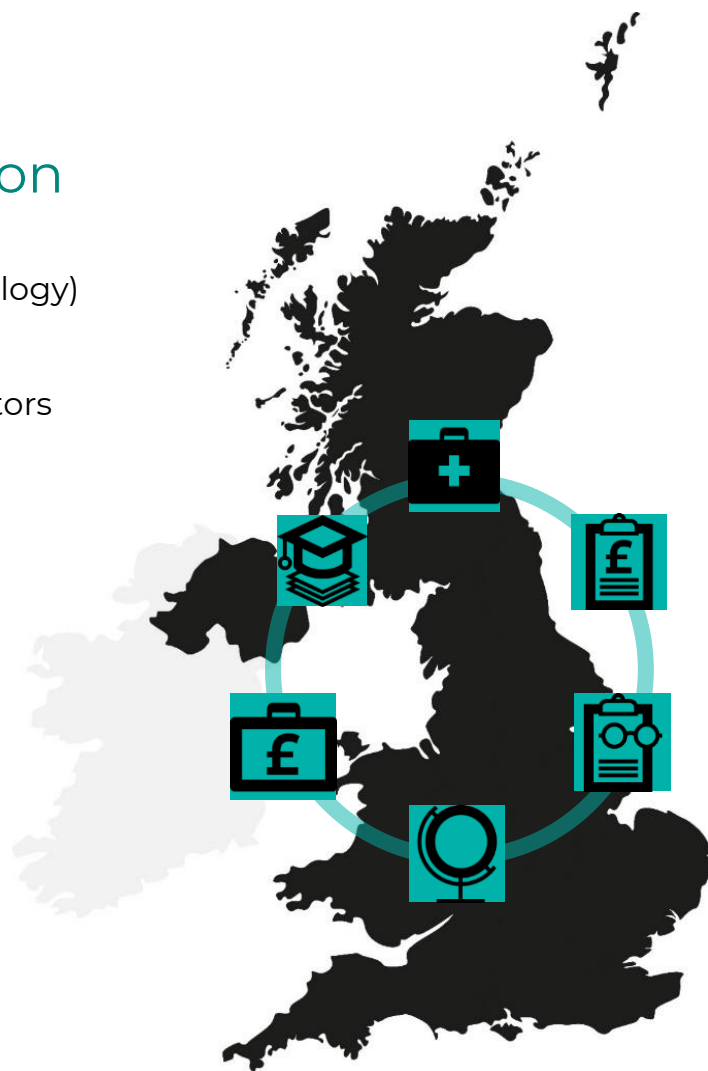
- COVID-19 collections (Questionnaires, Serology)
- SES & demographic
- Baseline physical & mental health
- Baseline family, SES and life-course indicators

### NHS COVID-19 datasets

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

### 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

**NIHR BioResource\_COPING**: National Institute of Health Research BioResource, COVID-19 Psychiatry and Neurological Genetics Study

**NSHD46**: MRC National Survey of Health and Development Cohort/1946 Birth Cohort

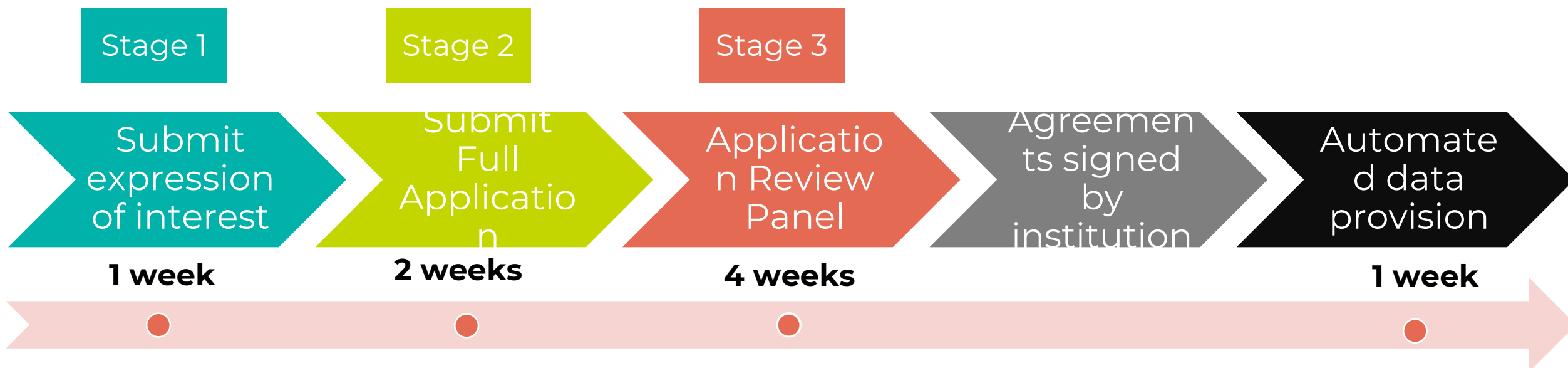
**SABRE**: Southall & Brent Revisited

**TRACK-COVID** Study

**TwinsUK**

**Understanding Society** - the UK Household Longitudinal Study

## UK LLC: Data Access Overview



Submit expression of interest via Health Data Research UK gateway  
<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**

# COVID-19 Longitudinal Health and Wellbeing National Core Study



## 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](mailto:access@ukllc.ac.uk)

# Analysing linked HES data

Richard Silverwood  
Associate Professor of Statistics &  
CLS Chief Statistician

CENTRE FOR  
LONGITUDINAL  
STUDIES



6 December 2022



Economic  
and Social  
Research Council

# 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

NCDS identifier	Generalised hospital provider	Spells	Order	Start date	End date	Discharge	Spells
-----------------	-------------------------------	--------	-------	------------	----------	-----------	--------

ncdsid	provspnops	admidate	epiorder	epistart	epiend	disdate	spelend
N3674M	HS6Q9B4HU27D96MF61P5	16jan2010	1	16jan2010	18jan2010	18jan2010	Y
N8429G	JS8G36DJV8538JSBDNH8	07aug2003	1	07aug2003	07aug2003	.	N
N8429G	JS8G36DJV8538JSBDNH8	07aug2003	2	07aug2003	13aug2003	13aug2003	Y
N4275R	H836FTK2SN9JTEB533N2	21feb1999	1	21feb1999	22feb1999	.	N
N4275R	H836FTK2SN9JTEB533N2	21feb1999	2	22feb1999	22feb1999	.	N
N4275R	H836FTK2SN9JTEB533N2	21feb1999	3	22feb1999	25feb1999	25feb1999	Y
N4275R	64NQF48J6D2UH76WN382	09nov2015	1	09nov2015	09nov2015	.	N
N4275R	64NQF48J6D2UH76WN382	09nov2015	2	09nov2015	14nov2015	14nov2015	Y

# HES data structure: Example (made up) data

ncdsid	provspnops	Admidate	epiorder	epistart	epiend	disdate	spelend
N3674M	HS6Q9B4HU27D96MF61P5	16jan2010	1	16jan2010	18jan2010	18jan2010	Y
N8429G	JS8G36DJV8538JSBDNH8	07aug2003	1	07aug2003	07aug2003	.	N
N8429G	JS8G36DJV8538JSBDNH8	07aug2003	2	07aug2003	13aug2003	13aug2003	Y
N4275R	H836FTK2SN9JTEB533N2	21feb1999	1	21feb1999	22feb1999	.	N
N4275R	H836FTK2SN9JTEB533N2	21feb1999	2	22feb1999	22feb1999	.	N
N4275R	H836FTK2SN9JTEB533N2	21feb1999	3	22feb1999	25feb1999	25feb1999	Y
N4275R	64NQF48J6D2UH76WN382	09nov2015	1	09nov2015	09nov2015	.	N
N4275R	64NQF48J6D2UH76WN382	09nov2015	2	09nov2015	14nov2015	14nov2015	Y


## 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.

Search  [ Advanced Search ]

ICD-10

Versions - Languages

Info



## ▼ ICD-10 Version:2010

- ▶ I Certain infectious and parasitic 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
- ▼ IX Diseases of the circulatory system
  - ▶ I00-I02 Acute rheumatic fever
  - ▶ I05-I09 Chronic rheumatic heart diseases
  - ▶ **I10-I15 Hypertensive diseases**
  - ▶ I20-I25 Ischaemic heart diseases
  - ▶ I26-I28 Pulmonary heart disease and diseases of pulmonary circulation
  - ▶ I30-I52 Other forms of heart disease
  - ▶ I60-I69 Cerebrovascular diseases
  - ▶ I70-I79 Diseases of arteries, arterioles and capillaries
  - ▶ I80-I89 Diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified
  - ▶ I95-I99 Other and unspecified disorders of the circulatory system
- ▶ X Diseases of the respiratory system
- ▶ XI Diseases of the digestive system
- ▶ XII Diseases of the skin and subcutaneous tissue
- ▶ XIII Diseases of the musculoskeletal system and connective tissue
- ▶ XIV Diseases of the genitourinary system

**(I00-I99)**
**Hypertensive diseases**  
**(I10-I15)**

**Excl.:** complicating pregnancy, childbirth and the puerperium ([O10-O11](#), [O13-O16](#))  
 involving coronary vessels ([I20-I25](#))  
 neonatal hypertension ([P29.2](#))  
 pulmonary hypertension ([I27.0](#))

**I10 Essential (primary) hypertension**

**Incl.:** High blood pressure  
 Hypertension (arterial)(benign)(essential)(malignant)(primary)(systemic)

**Excl.:** involving vessels of:

- brain ([I60-I69](#))
- eye ([H35.0](#))

**I11 Hypertensive heart disease**

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

**I11.0 Hypertensive heart disease with (congestive) heart failure**

Hypertensive heart failure

**I11.9 Hypertensive heart disease without (congestive) heart failure**

Hypertensive heart disease NOS

**I12 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

# ICD-10 coding: Chapters

Chapter	Block	Title
I	A00–B99	Certain infectious and parasitic diseases
II	C00–D48	Neoplasms
III	D50–D89	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
IV	E00–E90	Endocrine, nutritional and metabolic diseases
V	F00–F99	Mental and behavioural disorders
VI	G00–G99	Diseases of the nervous system
...	...	...
XXII	U00–U99	Codes for special purposes

# ICD-10 coding: Example (made up) data

Diagnosis	Diagnosis	Diagnosis	Diagnosis	Diagnosis	Diagnosis	Diagnosis	Diagnosis
Diagnosis	Diagnosis	Diagnosis	Diagnosis	Diagnosis	Diagnosis	Diagnosis	Diagnosis

ncdsid	provspnops	D_DIAG_01	D_DIAG_02	D_DIAG_03	D_DIAG_04	...	D_DIAG_20
N3674M	HS6Q9B4HU27D96MF61P5	I77	.	.	.	...	.
N8429G	JS8G36DJV8538JSBDNH8	Z85	G35	Z51	.	...	.
N8429G	JS8G36DJV8538JSBDNH8	Z85	G35	Z51	.	...	.
N4275R	H836FTK2SN9JTEB533N2	I42	K29	.	.	...	.
N4275R	H836FTK2SN9JTEB533N2	I42	K29	.	.	...	.
N4275R	H836FTK2SN9JTEB533N2	I42	K29	.	.	...	.
N4275R	64NQF48J6D2UH76WN382	R06	Z87	K12	N23	...	.
N4275R	64NQF48J6D2UH76WN382	R06	Z87	K12	N23	...	.

# ICD-10 coding: Example (made up) data

Personal history of family mental disorders	Multiple sclerosis	Optic neuritis	Medical care
---	--------------------	----------------	--------------

ncdsid	provspnops	D_DIAG_01	D_DIAG_02	D_DIAG_03	D_DIAG_04	...	D_DIAG_20
N3674M	HS6Q9B4HU27D96MF61P5	I77	.	.	.	...	.
N8429G	JS8G36DJV8538JSBDNH8	Z85	G35	Z51	.	...	.
N8429G	JS8G36DJV8538JSBDNH8	Z85	G35	Z51	.	...	.
N4275R	H836FTK2SN9JTEB533N2	I42	K29	.	.	...	.
N4275R	H836FTK2SN9JTEB533N2	I42	K29	.	.	...	.
N4275R	H836FTK2SN9JTEB533N2	I42	K29	.	.	...	.
N4275R	64NQF48J6D2UH76WN382	R06	Z87	K12	N23	...	.
N4275R	64NQF48J6D2UH76WN382	R06	Z87	K12	N23	...	.

# ICD-10 coding: Example (made up) data

ncdsid	provspnops	D_DIAG_01	D_DIAG_02	D_DIAG_03	D_DIAG_04	...	D_DIAG_20
N3674M	HS6Q9B4HU27D96MF61P5	I77	.	.	.	...	.
N8429G	JS8G36DJV8538JSBDNH8	Z85	G35	Z51	.	...	.
N8429G	JS8G36DJV8538JSBDNH8	Z85	G35	Z51	.	...	.
N4275R	H836FTK2SN9JTEB533N2	I42	K29	.	.	...	.
N4275R	H836FTK2SN9JTEB533N2	I42	K29	.	.	...	.
N4275R	H836FTK2SN9JTEB533N2	I42	K29	.	.	...	.
N4275R	64NQF48J6D2UH76WN382	R06	Z87	K12	N23	...	.
N4275R	64NQF48J6D2UH76WN382	R06	Z87	K12	N23	...	.

# 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.

## ICD-10 coding: Example code

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_ch_1_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.

appendix

OPCS-4.8 ▶ Volume I - Tabular List

- OPCS-4.8
  - Volume I - Tabular List
  - Volume II - Alphabetical Index

- Found: 16 matches
- Volume I - Tabular List
    - Part 2 - Tabular List of Four Digit Subcategories
      - H01 Emergency excision of appendix
        - H01.1 Emergency excision of abnormal appendix and drainage HFQ
        - H01.2 Emergency excision of abnormal appendix NEC
        - H01.3 Emergency excision of normal appendix
      - H02 Other excision of appendix
      - H03 Other operations on appendix
        - H03.1 Drainage of abscess of appendix
        - H03.2 Drainage of appendix NEC
        - H03.3 Exteriorisation of appendix
      - Z28.1 Appendix
  - Volume II - Alphabetical Index
    - I. Alphabetical Index of Interventions and Surgical Procedures
      - H03.- Appendix Operations NEC
      - Z28.1 Appendix site
      - H03.- Drainage Appendix
      - H01.- Excision Appendix Emergency
      - H02.- Excision Appendix NEC
      - H03.3 Exteriorisation Appendix

# OPCS coding: Chapters

Chapter/first character*	Title
A	Nervous System
B	Endocrine System and Breast
C	Eye
D	Ear
E	Respiratory Tract
F	Mouth
...	...
Z	Subsidiary Classification of Sites of Operation

\*Codes beginning with an “O” are overflow codes and should be handled differently.

**CENTRE FOR  
LONGITUDINAL  
STUDIES**

**CENTRE FOR  
LONGITUDINAL  
STUDIES**

**CENTRE FOR  
LONGITUDINAL  
STUDIES**

# OPCS coding: Example (made up) data

Comp	Printed	Following	signifying	NEC	Chroplasty of joint
------	---------	-----------	------------	-----	---------------------

ncdsid	provspnops	opertn_01	opertn_02	opertn_03	opertn_04	...	opertn_23
N3674M	HS6Q9B4HU27D96MF61P5	Z815	.	.	.	...	.
N8429G	JS8G36DJV8538JSBDNH8	U212	W581	.	.	...	.
N8429G	JS8G36DJV8538JSBDNH8	.	.	.	.	...	.
N4275R	H836FTK2SN9JTEB533N2	.	.	.	.	...	.
N4275R	H836FTK2SN9JTEB533N2	Y534	Z664	G451	.	...	.
N4275R	H836FTK2SN9JTEB533N2	T202	.	.	.	...	.
N4275R	64NQF48J6D2UH76WN382	.	.	.	.	...	.
N4275R	64NQF48J6D2UH76WN382	.	.	.	.	...	.

# OPCS coding: Example (made up) data

ncdsid	provspnops	opertn_01	opertn_02	opertn_03	opertn_04	...	opertn_23
N3674M	HS6Q9B4HU27D96MF61P5	Z815	.	.	.	...	.
N8429G	JS8G36DJV8538JSBDNH8	U212	W581	.	.	...	.
N8429G	JS8G36DJV8538JSBDNH8	.	.	.	.	...	.
N4275R	H836FTK2SN9JTEB533N2	.	.	.	.	...	.
N4275R	H836FTK2SN9JTEB533N2	Y534	Z664	G451	.	...	.
N4275R	H836FTK2SN9JTEB533N2	T202	.	.	.	...	.
N4275R	64NQF48J6D2UH76WN382	.	.	.	.	...	.
N4275R	64NQF48J6D2UH76WN382	.	.	.	.	...	.

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.

Silverwood RJ, et al. Examining the linkage quality and sample representativeness of linked National Child Development Study and Hospital Episode Statistics data. CLS Working Paper 2022/5. London: UCL Centre for Longitudinal Studies; 2022.

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

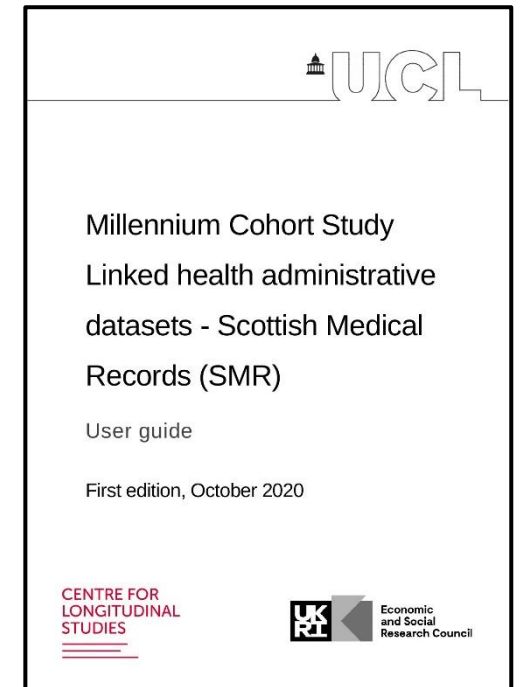
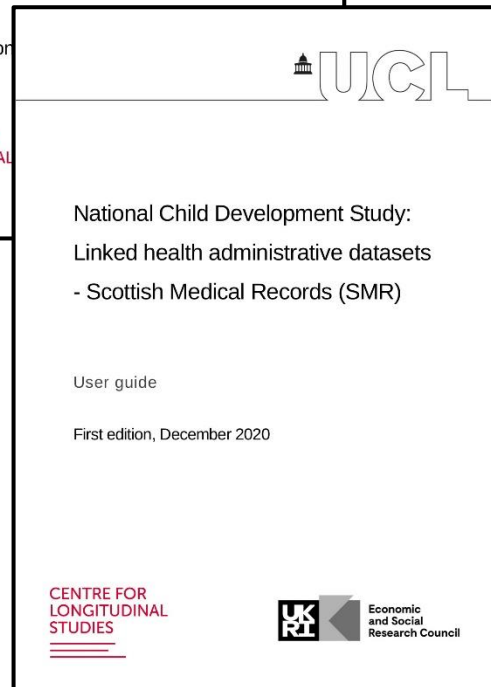
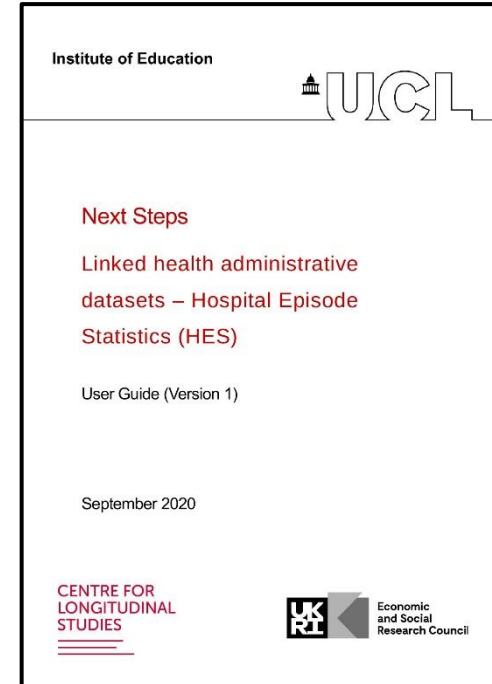
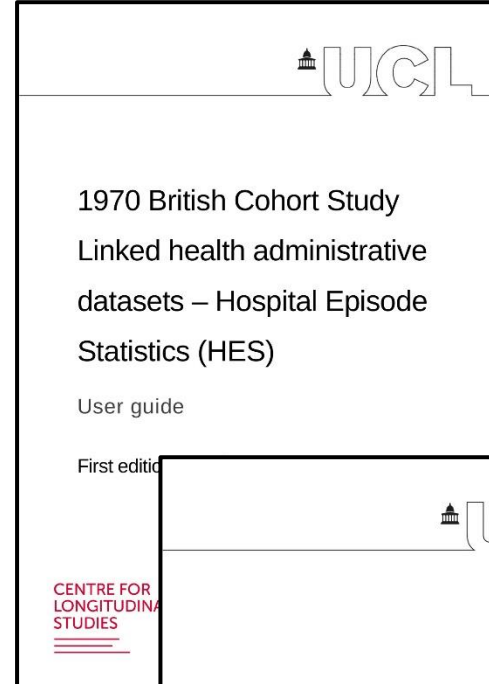
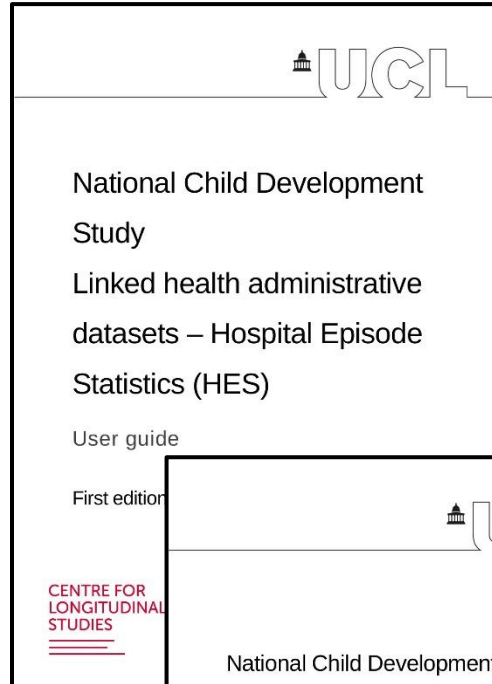
- Identified 10 such variables, e.g. cohort members treated for adult mental illness >70% more likely to be non-respondents (risk ratio 1.73; 95% CI 1.17, 2.51).
- Included these variables as auxiliary variables in multiple imputation analyses to explore the extent to which they helped restore sample 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](https://cls.ucl.ac.uk/data_documentation)

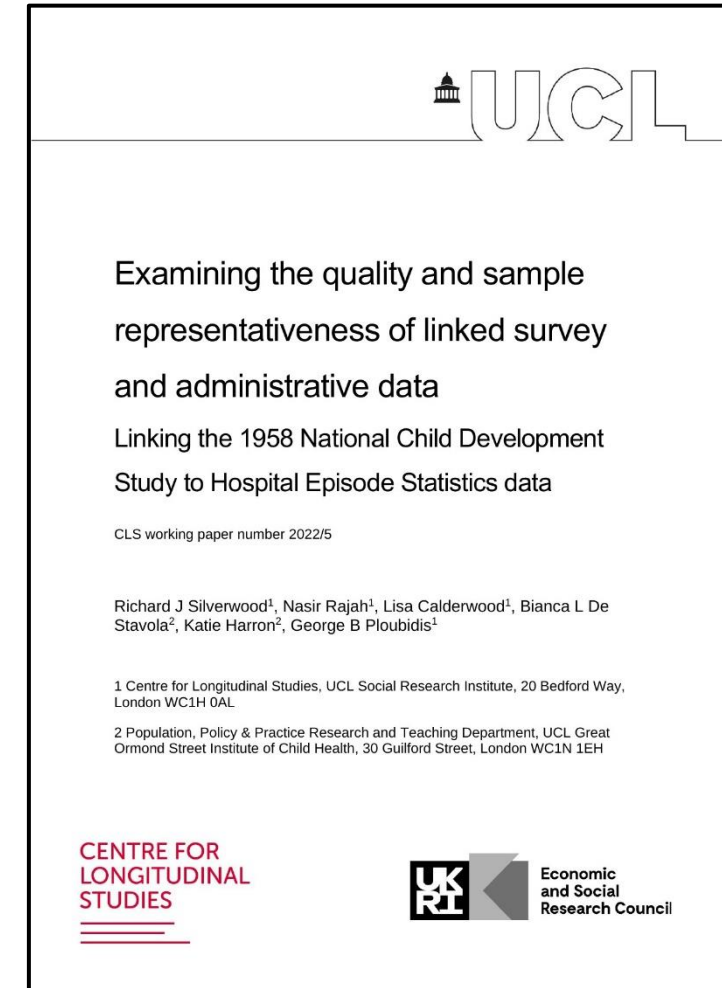
CENTRE FOR  
LONGITUDINAL  
STUDIES





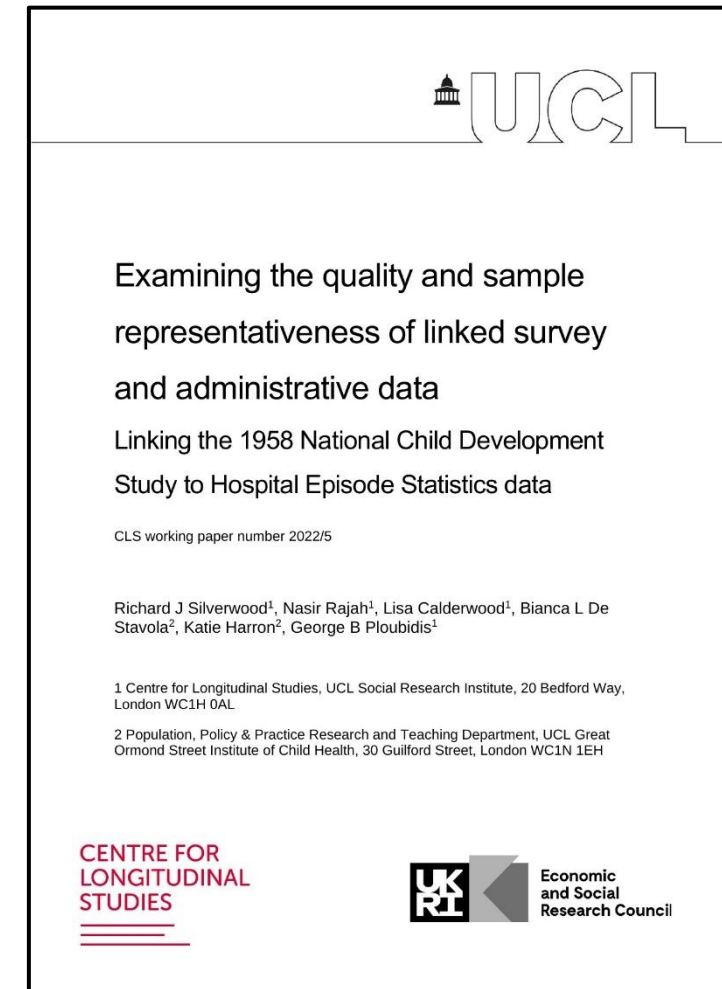
# Resources: Quality and representativeness of linked NCDS-HES data

- 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.



# 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.



## Other resources

- Boyd A, Cornish R, Johnson L, Simmonds S, Syddall H, Westbury L, et al. Understanding Hospital Episode Statistics (HES). London, UK: CLOSER; 2017.
- Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). International Journal of Epidemiology. 2017;46(4):1093-i.
- NHS Digital. Hospital Episode Statistics (HES). <https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics>.

Thank you.

CENTRE FOR  
LONGITUDINAL  
STUDIES



Economic  
and Social  
Research Council