



*The Creative Commons-Attribution-Noncommercial License 4.0 International applies to all works published by IASSIST Quarterly. Authors will retain copyright of the work and full publishing rights.*

## **The ONS Longitudinal Study – opportunities for longitudinal research on the England and Wales population**

Alison Sizer<sup>1</sup>, Andreas Mastrosavvas<sup>2</sup> and Oliver Duke-Williams<sup>3</sup>

### **Abstract**

Comprising longitudinal data on around 1.1 million individuals, the Office for National Statistics Longitudinal Study (ONS-LS) is the largest nationally representative longitudinal dataset in the United Kingdom. It follows a 1% sample of the England & Wales population drawn from the decennial census data (1971 – 2011), linked to some administrative data. Currently comprising up to 46 years of data (1971 – 2017) on sample members, the forthcoming linkage of the 2021 England and Wales Census data to the ONS-LS will extend this follow-up to 50 years. The Centre for Longitudinal Study Information and User Support (CeLSIUS) provides assistance for researchers wishing to use the ONS-LS in their research. Based at University College London (UCL) and the Office for National Statistics (ONS), it has been supporting academic and voluntary sector users of the ONS-LS since 2001. Its work includes helping researchers with their applications to use the ONS-LS, supporting research projects and advising on research outputs.

### **Keywords**

Census, longitudinal study, longitudinal research

### **Introduction**

This paper introduces the Office for National Statistics Longitudinal Study (ONS-LS) and the opportunities that it offers for longitudinal research on the England and Wales population. It details the information held in the ONS-LS, and it introduces the Centre for Longitudinal Study Information and User Support (CeLSIUS) and the work that it does to support current and prospective users of the ONS-LS. It provides details of the ONS-LS's sister studies, the Scottish Longitudinal Study and Northern Irish Longitudinal Study and the opportunities that the three studies offer for doing comparative research across the UK nations. It compares the ONS-LS to other census-based longitudinal studies in the international context, and concludes by highlighting the challenges faced by these census-based studies, and the ONS-LS specifically.

### **What is the ONS-LS?**

The ONS-LS was set up in 1974 by the then Office for Population Censuses and Surveys (OPCS), now Office for National Statistics (ONS). Initial sample members were drawn from the respondents to the 1971 Census and were selected on the basis of four undisclosed birthdays, giving an initial sample of approximately 1% of the England and Wales population (1.1 million individuals). The study is maintained through the annual addition of new births and immigrants with the same four birthdays. It includes the census forms from the 1971, 1981, 1991, 2001 and 2011 censuses, linked to some vital registrations data, including the births of new sample members, deaths of sample members, cancer registrations, live births to sample mothers, still births to sample mothers, infant deaths and widowerhoods.

The ONS-LS is representative of the population for England and Wales and in contrast to other longitudinal studies it includes individuals living in communal establishments, for example, hospitals, residential care homes and nursing homes, prisons, boarding schools, hostels and hotels. It also includes information from the census forms of other individuals living in the same household as a sample member at the time of the census. This means that in childhood other individuals in an LS member's household might be their parents and siblings, but later in their life, it might be their partner or spouse and their children. This offers opportunities to investigate inter-generational social change. The large size of the ONS-LS, currently around 1.1 million sample members, means that it is also possible to examine small population groups, for example specific groups of ethnic minorities, older age groups, and individuals in residential care homes, which may not be possible in other longitudinal studies, which have smaller sample sizes.

The ONS-LS is a dynamic sample of all individuals completing a census and usually resident in England and Wales with one of the four confidential birthdays. The sample has been maintained since its start on census day in 1971 (26<sup>th</sup> April 1971) in the following ways:

- Children subsequently born on one of the four confidential birthdays since 26th April 1971 are added to the sample.
- Immigrants registering with the NHS Central Registry (NHSCR) and being born on one of the four confidential birthdays are added to the sample.
- Any individuals completing a subsequent census form and giving their birth date as one of the four confidential birthdays is added to the sample if they are not already included.
- Former ONS-LS members who left the sample through emigration are re-entered into the sample on re-registering with the NHSCR.
- Deaths of LS members are recorded on their records but their information is retained.
- LS members who emigrate have their date of embarkation recorded but their records are retained and they can therefore re-enter the study on re-registering with the NHS (see above).
- LS members who enlist with the Armed Forces have their date of enlistment recorded. Again, their records are retained and therefore similarly to emigrants, they can re-enter the study.

Figure 1 illustrates the dynamic nature of the ONS-LS sample and how it is maintained. It shows the samples at each of the censuses included in the study, and the members of the sample that have been traced to the NHSCR, which facilitates the linkage of their record to the vital registrations information that is included in the study. It also shows the approximate number of entry events (births, immigrations, re-entries) and exit events (deaths, emigrations and enlistments) over the course of the study to 2017.

Figure 1: ONS Longitudinal Study: Samples at each census and the linked registry information

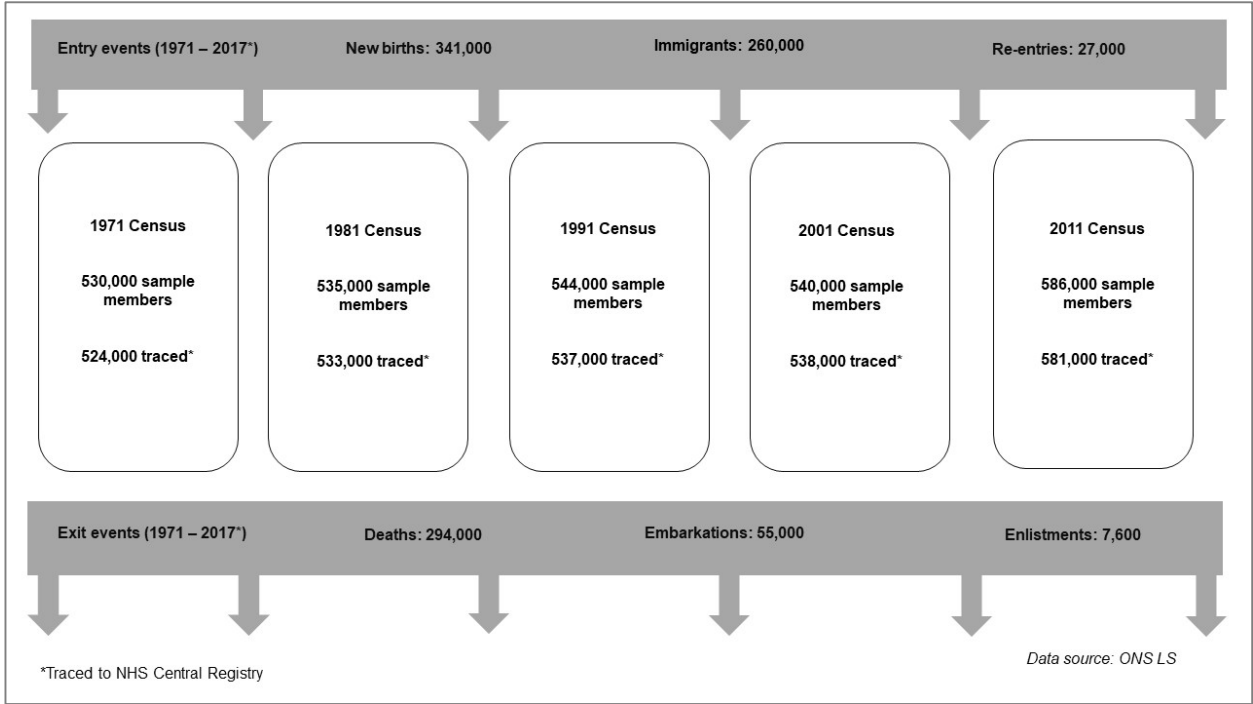
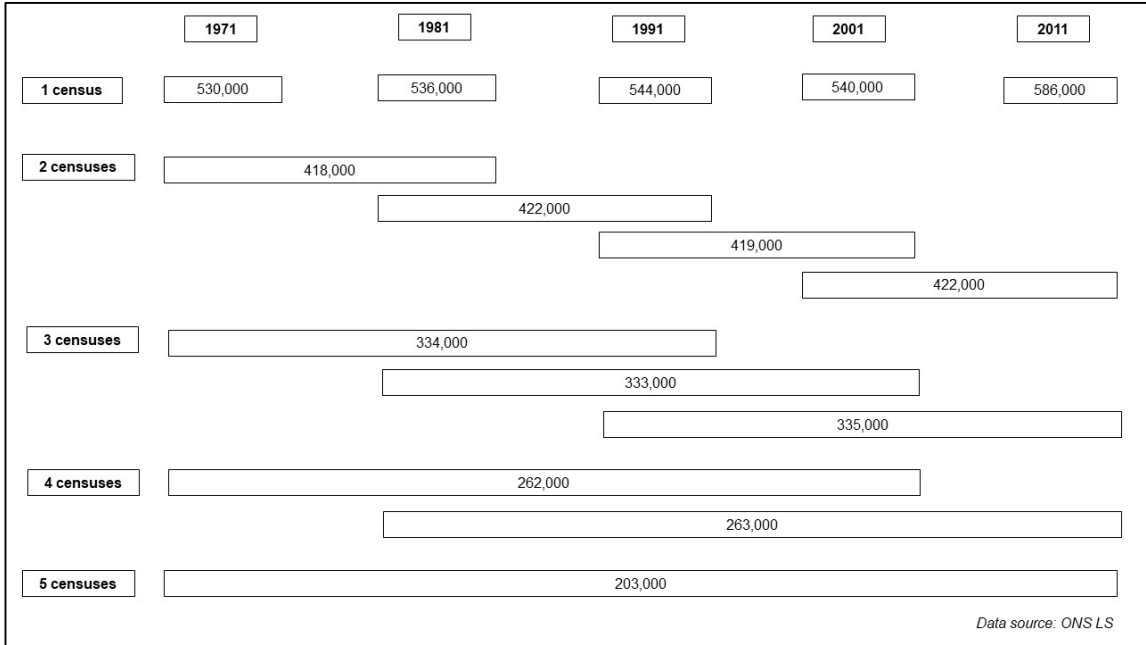


Figure 2 shows the approximate numbers of LS members who are enumerated in each individual census (the 1971, 1981, 1991, 2001 and 2011 censuses); two consecutive censuses (the 1971-1981, 1981-1991, 1991-2001, and 2001-2011 censuses); three consecutive censuses (the 1971-1991, 1981-2001, and 1991-2011 censuses); four consecutive censuses (the 1971-2001, and 1981-2011 censuses); and all five censuses (1971-2011 censuses).

Figure 2: ONS Longitudinal Study: samples for single census years and multiple consecutive censuses



### The information included in the ONS-LS?

Household and sociodemographic information collected through the decennial census (1971 – 2021) is included in the ONS-LS, along with variables that are derived from this information. The variables derived from the census information include social class, derived from the employment and occupation related information in the census, and 10-year migration indicators, derived by comparing the current address in successive censuses. The household information included in the ONS-LS covers type of accommodation and whether it is self-contained, tenure, household amenities (e.g. type of central heating), number of rooms, and number of cars/ vans. The individual related topics included in the ONS LS include age, sex, country of birth and marital/ civil partnership status; educational qualifications; economic position or activity (employed/ unemployed/ retired/ inactive); employment status (employed/ self-employed, full-time/ part-time), occupation, work address, number of hours worked and travel to work mode; and geographical location (current and historical). In addition to these variables, which are included across the censuses, topics were added to the censuses in 1981, 1991, 2001 and 2011. In 1991 (Office of Population Censuses and Surveys, 1991), ethnicity and long-term limiting illness<sup>4</sup> were added to the census. In 2001 (Office for National Statistics, 2001), a voluntary question on religion<sup>5</sup>, and questions on informal caregiving<sup>6</sup> and self-rated health<sup>7</sup> were added. In 2011 (Office for National Statistics, 2011), questions were added related to national identity<sup>8</sup>, passports held<sup>9</sup>, time since arrival in the United Kingdom<sup>10</sup>, and intended length of stay<sup>11</sup>.

On 21<sup>st</sup> March 2021, the most recent census was held in England and Wales<sup>12</sup>. Further questions were added to the census, voluntary response questions related to gender identity<sup>13</sup> and sexuality<sup>14</sup>, a revised question related to marital or civil partnership status<sup>15</sup>, and one related to being a veteran of the Armed Forces<sup>16</sup>. The linkage of the 2021 Census to the ONS-LS, and the testing of the linkage, is currently on-going and it is hoped that the linkage will be completed in the latter half of 2025. This linkage will extend the follow-up of the ONS-LS to 2021 (50 years), enabling researchers to examine changes that have taken place in 2011 – 2021 period, a period which saw Brexit and the Covid 19 pandemic.

Table 1 lists the topics included in the census forms and the ONS-LS, and the census years that they were addressed.

Table 1: Topics included in the ONS LS by census year (1971 – 2021) (Data source: ONS LS)

Topic	Census year					
	1971	1981	1991	2001	2011	2021
<b><i>Household related</i></b>						
Location (various geographical levels)	√	√	√	√	√	√
Tenure	√	√	√	√	√	√
Whether accommodation is self-contained	√	√	√	√	√	√
Type of accommodation (caravan, flat, temporary structure, block of flats, whole house)		√	√	√	√	√
Number of rooms	√	√	√	√	√	√
Household amenities (e.g. heating, hygiene)	√	√	√	√	√	√
Number of cars/ vans	√	√	√	√	√	√
<b><i>Individual related</i></b>						
Sex	√	√	√	√	√	√
Date of birth	√	√	√	√	√	√
Relationship to head of household/ household reference person	√	√	√	√	√	√
Usual resident or visitor at enumeration address	√					
Geographical location one-year ago	√	√	√	√	√	√

Geographical location five-years ago	√					
Geographical location of second address					√	√
Marital status (including civil partnership)	√	√	√	√	√	√
Sex of marital or civil partner						√
Marital history (women under-60 who are married/ widowed/ divorced)	√					
Live born children (women under-60 who are married/ widowed/ divorced)	√					
Country of birth	√	√	√	√	√	√
Mother and father country of birth	√					
National identity					√	√
Passports held					√	√
Year of arrival in UK (if born outside UK)					√	√
Intended length of stay in UK					√	√
Ethnic group			√	√	√	√
Religion (voluntary answer)				√	√	√
Welsh language (Wales only) (speak, read, write)	√	√	√	√	√	√
Main language, English language ability					√	√
Sexuality (voluntary answer)						√
Gender identity (voluntary answer)						√
Education qualifications (NB: question varies over time)	√	√	√	√	√	√
Working/ unemployed/ retired/ inactive (previous week)	√	√	√	√	√	√
Whether a student (previous week)	√	√	√	√	√	√
Employment status (employed/ self-employed, full/ part-time)	√	√	√	√	√	√
Occupation	√	√	√	√	√	√
Location of main place of work						√
Work address	√	√	√	√	√	√
Occupation one year ago	√					
Year last worked				√	√	√
Hours worked	√	√	√	√	√	√
Transport to work mode	√	√	√	√	√	√
Long-term limiting illness			√	√	√	√
Self-rated health				√	√	√
Informal caregiving				√	√	√
Veteran status						√

Life events data have also been linked to the ONS-LS, which means that the LS also includes information that is not addressed through the decennial census, including births, deaths and cancer registrations. The life events data in the ONS-LS are updated annually with a two-year delay. Linkage happens through two processes:

- The birth and death register and cancer registration files that are sent to ONS annually are searched for individuals with one of the four ONS-LS birthdates. These are then linked to actual sample members.

- Sample members who are marked as ONS-LS sample members in the NHSCR are sent to the Longitudinal Study team at ONS when a birth or death is recorded in the Register. These are then added to the ONS-LS database.

Currently the following life events data from 1971 to 2017 have been linked to ONS-LS study members records (Centre for Longitudinal Study Information and User Support, 2018):

- New births of sample members.
- Deaths of sample members.
- Live births and still births registered to sample mothers.
- Infant deaths of births registered to sample mothers.
- Cancer registrations.
- Widowerhoods (these are drawn from death registrations when an LS member's legal spouse/ civil partner dies, and the LS members are identified through the date of birth of surviving spouse information in the death registration form).

The information available in the ONS-LS from the linkage comprises information that is recorded in the birth, death and cancer registration forms. Table 2 shows the information available in the ONS-LS according to the different life events that are linked to it.

Table 2: Topics included in the life events information that is linked to ONS-LS members (data source: ONS-LS)

Life events data	Number of events (1971 – 2017)	Information available
<i>New births of sample members</i>	340,542	Day, month and year of birth*
		Place and geographical location of birth
		Sex
		Birthweight
		Whether multiple birth; type of multiple birth
		Number of previous live or still births to mother
		Parent's age
		Occupation of working parent
<i>Deaths</i>	293,931	Day, month and year of death
		Place and geographical location of death
		Underlying and contributory causes of death (ICD8, ICD9, ICD10)**
		Whether the death was in a communal establishment and establishment type
<i>Live births to sample mothers</i>	331,304	Date of birth
		Sex
		Place and geographical location of birth
		Birthweight
		Whether multiple birth; type of multiple birth
		Number of previous live or still births to mother

<i>Still births to sample mothers</i>	1,864	Age of sample mother at birth
		Date of birth
		Sex
		Place and geographical location of birth
		Birthweight
		Whether multiple birth; type of multiple birth
		Number of previous live or still births to mother
		Age of sample mother at birth
		Whether the sample mother was married at the time of the birth (and year of current marriage)
<i>Infant deaths to sample mothers***</i>	2,418	Social class, employment status and occupation of sample mother and father
		Date of birth of infant/ child
		Date of death of infant/ child
		Sex of child
		Usual address of sample mother at time of infant/ child birth and death
		Place of birth and place of death
		Birthweight
		Whether multiple birth; type of multiple birth
		Cause of infant death (according to ICD8, 9, 10)
		Number of previous live and still births to mother
		Age of sample mother at birth
<i>Cancer registrations (to 2016)</i>	160,727	Whether the sample mother was married at the time of the birth (and year of current marriage)
		Social class, employment status and occupation of sample mother and father
		Date of cancer registration
		Age at time of registration
		Site and sub-site of cancer
<i>Widow(er)hoods</i>	99,892	Treatment type
		Date of death (if dead)
		Age and date of death of LS member's spouse
		Cause of death of LS member's (according to ICD8, ICD9, ICD10)

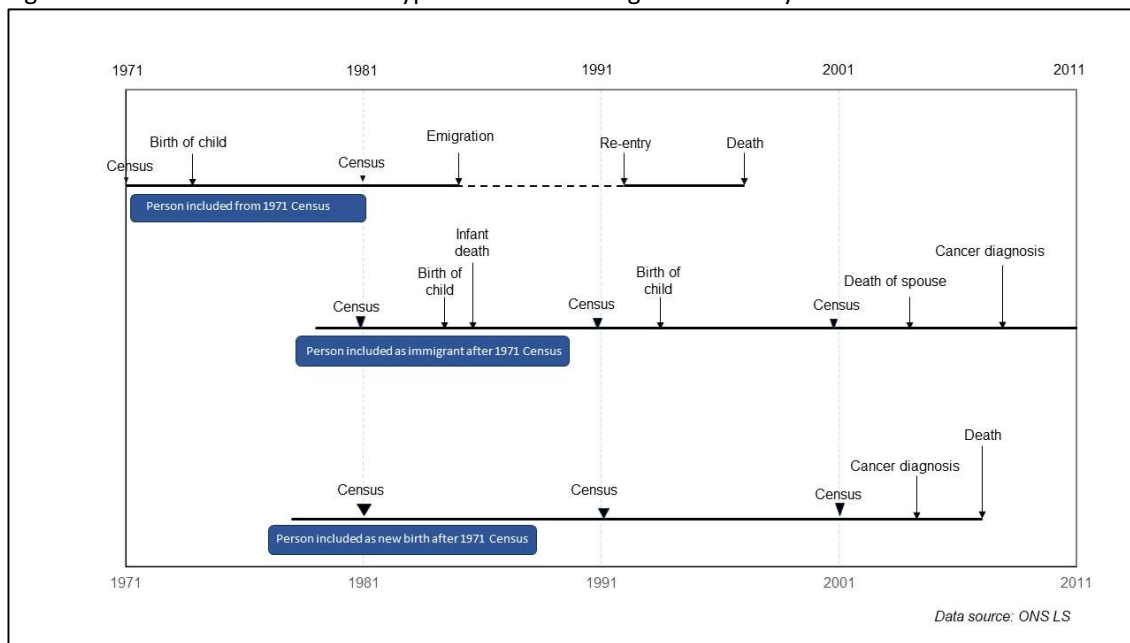
Notes to Table 2: \* Access to the day and month of birth requires special permission from ONS, because they will disclose the confidential birthdays used to select ONS-LS sample members and hence disclose the identity of LS members.

\*\*Three different versions of the International Classification of Diseases (ICD) have been used over the course of the ONS-LS: ICD8 (from census day 1971 to 4th April 1981); ICD9 (from 5th April 1981 to 31st December 2000); and ICD10 (from 1st January 2001 onwards).

\*\*\* Until January 1993 this only included infants up to one-year old. Following a change in the national database, from January 1993, the death of any child born in 1993 was included in the LS, which means that currently this includes deaths of children under 16-years old.

The life courses of three hypothetical ONS-LS members based on their enumeration through the census and the life events information that is linked to their ONS-LS record through the NHSCR are shown in Figure 3. The first ONS-LS member is included from the 1971 Census on account of being born on one of the four LS birthdays. They have a child between the 1971 and 1981 censuses and then emigrate in 1985. They return to the UK and re-enter the sample in 1992 through registration with a GP. The ONS-LS member then dies in 1997. The second ONS-LS member migrates into England and Wales in 1979 and enters the sample through registering with the NHSCR and being born on one of the four confidential birthdays. They then complete the 1981 Census. Between the 1981 and 1991 censuses they have a live birth and subsequently that child dies (infant death). They complete the 1991 Census and a further child is born between the 1991 and 2001 censuses. They then complete the 2001, but prior to the 2011 Census their spouse dies and they are subsequently diagnosed with cancer in 2009. They go onto complete the 2011 Census. The third LS member is born in 1978 and is then enumerated in the 1981, 1991 and 2001 censuses. After the 2001 Census they are diagnosed with cancer and then they die in 2007.

Figure 3: The life courses of three hypothetical ONS Longitudinal Study members



### Geographical information included in the ONS-LS

The ONS-LS includes the geographical location of LS members at the time of census based on the address on the front of the census form. The location is recorded at a range of different geographical levels including Government Office Region, county and county district, local authority district and ward. Location at lower geographical levels is also recorded including postcode area, Output Area



(OA)<sup>17</sup>, Lower Super Output Area (LSOA)<sup>18</sup> and Middle Super Output Area (MSOA)<sup>19</sup> (Office for National Statistics, 2022). However, these lower-level geographical variables are subject to additional access permissions and restrictions owing to the potential for the identification (disclosure) of individual LS members. In addition to geographical location of the LS member at the time of the census, the census forms also ask for the address one year ago if it is different (the 1971 Census also asked about the address five years ago). This information can be used to look at one-year moves for LS members, and comparing the geographical location of LS members in subsequent censuses enables 10-year moves to be examined.

In addition to using the geographical location of LS members to examine migration between geographical areas in England and Wales, the geographical location of LS members can be used to link ecological variables to the ONS-LS, including neighbourhood deprivation (Norman & Boyle, 2014), population density (Norman & Boyle, 2014) and pollution data.

Travel to work areas (TTWAs) and workplace zones of LS members are also included in ONS-LS, based on their workplace address. TTWAs are defined to approximate self-contained labour-market areas within which commuting to and from work occur. They are areas where the number of people living and working in the area should be at least 75% of the total number of workers living in the area and the total number of people working in the area (Hattersley & Creeser, 1995). In contrast to TTWAs, workplace zones are areas where people work, and they are designed to have similar numbers of workers within them (200 - 625 workers). There were 60,709 workplace zones in 2011 (Office for National Statistics, 2012).

#### Previous research using the ONS-LS

The ONS-LS currently has 46 years of follow-up data, 1971 – 2017, and it has been used by researchers to examine issues with social policy implications, including social inequalities in health, employment and education; social exclusion including the long-term health implications of non-employment and low educational status; housing and geographical mobility; and family policy. The research summarised below describes four recently completed and on-going research projects using the ONS-LS.

Murray et al. (2020) and Sacker et al (2024) used the ONS-LS to examine the health of looked after children when they are grown up. Their research found that adults who had been in care in any of the 1971 – 2001 censuses faced higher risk of mortality by 2013 than those LS members who had never been in care. They also found that the excess mortality was predominantly attributable to three causes, self-harm, accidents or mental and behavioural disorders (Murray, Lacey, Maughan, & Sacker, 2020; Sacker, Murray, Maughan, & Lacey, 2024).

Sturley et al. (2023) used the ONS-LS to study variations in colorectal cancer incidence and survival over a 15-year period by individual characteristics and area type. The researchers found that compared to LS members without a degree, those with a degree had a lower incidence of colorectal cancer. It was also higher among individuals who were employed in manual occupations compared to LS members employed in non-manual occupations. Furthermore, these disparities were even more pronounced in terms of colorectal cancer survival. Although there was no clear relationship between area deprivation and colorectal cancer incidence among LS members, individuals living in the most deprived areas had higher probability of death than those living in the least deprived areas (Sturley, Norman, Morris, & Downing, 2023).

Researchers at the Institute for Fiscal Studies (IFS) have used the ONS-LS to examine intergenerational social mobility in England and Wales (Bourquin, Joyce, & Sturrock, 2021). They analysed data on LS members born in the 1960s and 1980s comparing their average inheritances to their lifetime projected income. The study found that for LS members born in the 1960s, average inheritances were worth 9%

of their projected household lifetime income (non-inheritance), whereas for those born in the 1980s this figure rose to 16%, suggesting that inheritances were set to increase inequalities in lifetime income between individuals with richer and poorer parents.

Finally, the ONS-LS is being used to examine the labour market impact of trade shocks (Irastorza-Fadrique, Levell, & Parey, 2024). This research is looking at the LS member's own and their partner's responses to the decline of manufacturing employment due to import substitution and has found that responses vary significantly by gender. Men in households exposed to import competition respond to shocks by increasing their labour force participation into older ages, and by moving into solo self-employment. Furthermore, this response is observed both when men themselves lose employment and when their partners are affected. In contrast however, women did not increase their labour supply if their male partners were initially employed in exposed industries.

### **Applying to use the ONS Longitudinal Study data**

Although researchers are free to use the ONS-LS data, access is strictly controlled. Researchers wanting to access the ONS-LS must be accredited researchers under the Digital Economy Act (2017). This allows them to access secure ONS data, such as the ONS-LS, through the Secure Research Service (SRS) or another Trusted Research Environment (TRE). Researchers apply to be an accredited researcher through the People and Project Services (PPS) platform (see: <https://integrateddataservice.gov.uk/apply-for-researcher-accreditation>).

Accredited researchers must have an undergraduate degree or higher, with a significant proportion of maths or statistics, or be able to demonstrate at least three years of experience of quantitative research (Office for National Statistics, 2024b). Once a researcher has applied to be an accredited researcher through the PPS, they are invited to complete compulsory online Safe Researcher training. The training covers the “five safes”<sup>20</sup> (Ritchie, 2017), the safe handling of data, and the disclosure control rules for the research outputs. The disclosure control rules are applied to research outputs in order to minimise the risk of the disclosure of personal information about the data subjects. Accredited researchers are required to renew their accreditation every five years, involving attending refresher training (UK Statistics Authority, 2020).

Accredited researchers wanting to use the ONS-LS must apply for access to the ONS-LS data. The application for the ONS-LS data comprises two parts:

- Research accreditation application form.
- ONS Longitudinal Study Supplementary form.

Both are available through the CeLSIUS website (Centre for Longitudinal Study Information and User Support, 2024b).

Information required for the research accreditation application form focuses on the research aim and objectives, as well as the research questions and hypotheses. Researchers are required to set out their research methodology and identify any research biases. In addition, all research using the ONS-LS must demonstrate public benefit. The application form also asks if external data will be linked to the ONS-LS, for example pollution data or neighbourhood deprivation levels. If such linkage is intended, researchers must explain the rationale for the linkage, provide the source and owner of the external data, confirm that permission to use the data has been granted, and describe the method of linkage to the ONS-LS. In the application form all the members of the research team must be identified, and those requiring access to the data, rather than just reviewing outputs from the data, must be accredited researchers. (Centre for Longitudinal Study Information and User Support, 2022)

In the ONS Longitudinal Study supplementary form the researcher describes their study sample and the variables that they will require for their research project. Researchers will need to refer to the

ONS-LS Data Dictionary for a list of the variables available for researchers to use in their ONS-LS research project (see: <https://iweb.dis.ucl.ac.uk/celsius/standalone/searchResults.php>). The Data Dictionary groups the variables into different tables based on their source. For example, there are separate tables for each census and separate tables for ONS-LS members (ME71, ME81, ME91, ME01 and ME11) and non-members<sup>21</sup> (NM71, NM81, NM91, NM01, NM11). There are also tables for each of the types of linked events data (for example the table called “DETH” includes the variables drawn from the death registration, and the table called “NBIR” includes the variables based on the birth registrations of new ONS-LS study members). This grouping of the variables into a series of tables means that researchers can browse through the variables in each table. It is also possible to do a dictionary search of the Data Dictionary in order to search for specific variables or keywords. (Centre for Longitudinal Study Information and User Support & Office for National Statistics, 2014)

Once a researcher has completed both parts of the research accreditation application (application form and ONS LS supplementary form), they should send them to the Centre for Longitudinal Study Information and User Support (CeLSIUS) to submit on their behalf. The application then undergoes a series of feasibility checks by ONS, the LS Data owners and ONS-LS Research Board, and finally ONS’s Research Accreditation Panel. If none of these raise any serious concerns about the research project being applied for, the application is approved.

Once an application to use the ONS-LS has been approved, all the accredited researchers named in the research team must sign an Accredited Researchers Assurance Registration (ARAR) form. The form sets out the commitments that Accredited Researchers must make before they will be able to access the Secure Research Service (SRS) environment through which they will access their data. It consists of sections which address, i) the responsibilities of the accredited researcher; ii) security of the data; iii) lawful access and data confidentiality; iv) data clearance; v) non-compliance and breach sanctions; vi) dispute procedures; and vii) incident management and reporting.

### **Accessing the ONS LS data**

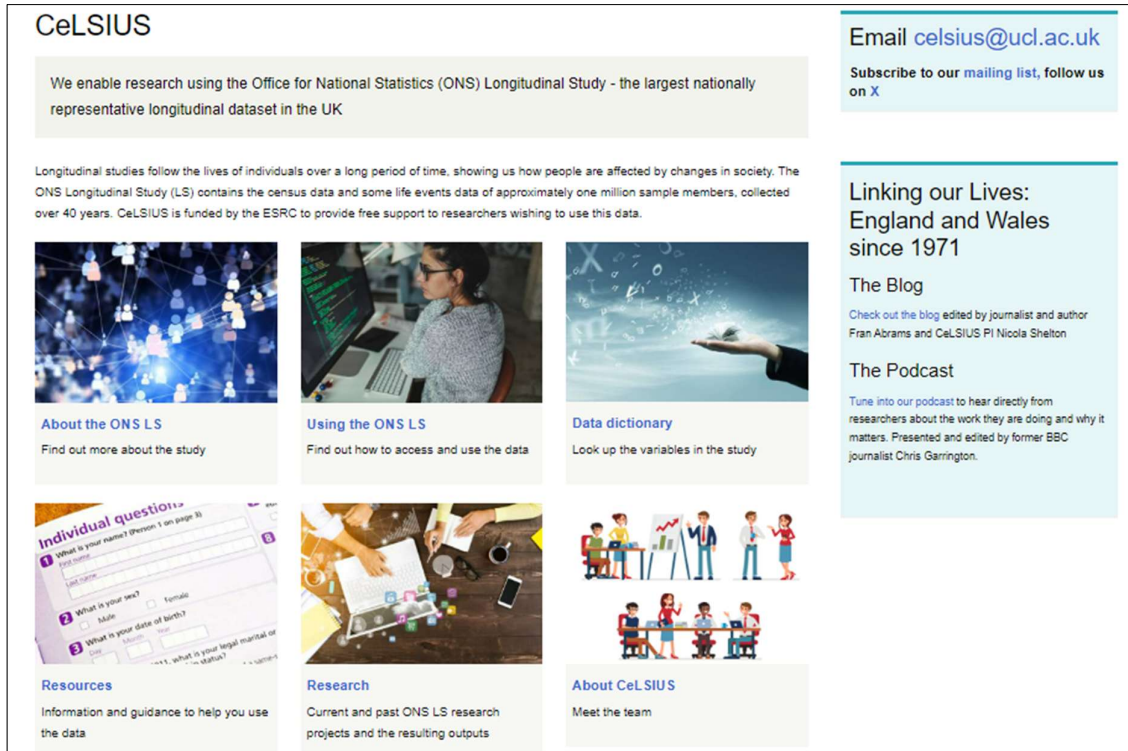
Once an application to use the ONS-LS has been approved and the data has been extracted, there are four routes through which researchers can access their project data. The first of these is through one of the ONS safe researcher settings at their offices in Titchfield (Hampshire) or Newport (Wales). Since August 2021 ONS has allowed researchers to securely access their ONS-LS data through one of a network of SafePods<sup>22</sup> around the UK. Currently there are 23 SafePods, 18 in England (including two in London), four in Scotland and one in Northern Ireland (University of St. Andrews, 2021). Some organisations have an Assured Organisational Connectivity (AOC) agreement with ONS, which enables researchers to access their data via the SRS from their office. Subject to further ONS approval, researchers may also be able to access their data from their home via their institution-owned computer (Office for National Statistics, 2024a). Finally, for researchers who are unable to access their data in any other way, CeLSIUS User Support Officers will run researcher written syntax and then send them the output.

### **The Centre for Longitudinal Study Information and User Support (CeLSIUS)**

Based at University College London (UCL), the Centre for Longitudinal Study Information and User Support (CeLSIUS) was set up in 2001 to support academic and voluntary sector users of the ONS-LS. This support includes helping researchers with their ONS-LS research applications, guiding and assisting them with their analysis and advising on publications and other research outputs that they might want to consider using to share their research findings, e.g. podcasts and blogs. CeLSIUS maintains a website that provides access to a wealth of information on the ONS-LS including the Data Dictionary; links to the census forms; a series of thematic guides to using the ONS-LS; blogs and podcasts showcasing research using the ONS-LS; and recordings of webinars delivered by the CeLSIUS team (Centre for Longitudinal Study Information and User Support, 2024a). CeLSIUS User Support

Officers (USOs) have also written syntax for deriving variables that many LS researchers use in their analysis, e.g. cause of death, and categorisations of country of birth. Figure 4 shows the CeLSIUS website Home page.

Figure 4: CeLSIUS website Home Page (Centre for Longitudinal Study Information and User Support, 2024a)



### Future directions of the ONS-LS

As stated above, access to ONS-LS data is strictly controlled to protect identifying information, e.g., the confidential inclusion dates. Although extensive user support from CeLSIUS is provided to guide researchers through application and data access processes, researchers can face challenges due to their inability to view data before approval. Despite the existence of the ONS-LS Data Dictionary, it lists thousands of variables, which can make identifying relevant variables difficult, particularly for new users, potentially resulting in delays to the application approval process. Additionally, researchers cannot begin writing scripts or testing code until after gaining access to their data in the secure environment, further slowing the research process. To address these issues, CeLSIUS is developing the Longitudinal Impossible Dataset (LIDS): an artificial, user-generated dataset that mimics the structure of ONS-LS data without posing disclosure risks. LIDS supports users in project planning and script development by allowing them to select variables and download artificial data generated from publicly available metadata via an interactive interface. As LIDS is generated by taking random draws from publicly available code lists, it remains informative with respect to the coding of individual variables while eliminating real data disclosure risks. Additional features, such as the inclusion of 'impossible' relationships between variables due to randomisation and adherence to the ONS-LS longitudinal linkage convention of tracking participants across censuses via unique identifiers, further facilitate data discovery and script development while minimising the risk of perceived disclosure (Mastrovasas & Shelton, 2025).

## The UK Longitudinal Studies and opportunities for research across the four nations of the United Kingdom

The ONS-LS, which covers England and Wales, is one of three sister studies that cover the four UK nations (England & Wales, Northern Ireland, Scotland). The other two cover Northern Ireland, the Northern Ireland Longitudinal Study (NILS), and Scotland, the Scottish Longitudinal Study (SLS). Although they are all based on linking the decennial census forms and comprise dynamic samples, they differ in terms of the years that they cover, the sample selection and the sample size. As stated above, the ONS-LS comprises a 1% sample of the population of England & Wales, which is based on being born on one of four birthdays, and started with the 1971 Census. NILS is based on a 28% sample of the Northern Ireland population, starting with the 1981 Census, selected using 104 birthdays. Finally, SLS is based on a 5.3% sample of the Scottish population, starting with the 1991 Census, selected using 20 birthdays. The three studies also differ in terms of the additional data that they are linked to. All the studies are linked to birth and death registrations, immigrations and embarkations, but there are additional linkages that differ between the studies. The ONS-LS is also linked to cancer registrations. NILS can also be linked to information on the type and value of accommodation from the Land and Property Services, and with special approval can be linked to health data, including breast cancer screening, dental treatment and the prescription of medicines. The SLS also includes marriages from the Civil Registration System, weather and pollution data, school level data from the Scottish Government Directorate, and with additional permission, linkage to hospital episodes, maternity data and cancer data.

The three UK Longitudinal Studies are supported by different agencies, and have separate websites with information about the study, research using it and how to apply for access to each study's data. CeLSIUS supports current and prospective researchers using the ONS-LS<sup>1</sup>; the Scottish Longitudinal Study Development and Support Unit (SLS-DSU) supports researchers using the SLS<sup>2</sup>; and the Northern Ireland Longitudinal Study Research Support Unit (NILS-RSU) supports researchers using NILS<sup>3</sup>. The websites also include information on how researchers can access their data<sup>4</sup>.

### Cross-nation research and analysis

While it is possible to undertake comparative research using data from the three sister studies, owing to disclosure concerns, it is not possible to fully combine the three datasets to create a UK-wide sample. However, there are alternative possibilities, one of which is to undertake analysis on the three datasets separately and then compare the results or combine them on an ad hoc basis. However, this approach has several disadvantages. First, it is not possible to be sure that the datasets and variables which appear to be similar are comparable. Second, an analysis that adjusts for covariates in each individual study will not be identical to one that would be obtained if the raw datasets were combined. Finally, tests of region by covariate interactions cannot be carried out easily through the comparison or combination of the region-specific results (Raab & Diben, 2020).

---

<sup>1</sup> Website: <https://www.ucl.ac.uk/population-health-sciences/epidemiology-health-care/research/ucl-research-department-epidemiology-public-health/research/health-and-social-surveys-research-group/studies/celsius>; email: [celsius@ucl.ac.uk](mailto:celsius@ucl.ac.uk).

<sup>2</sup> Website: <https://sls.lscs.ac.uk/>; email: [sls@lscs.ac.uk](mailto:sls@lscs.ac.uk).

<sup>3</sup> Website: <https://nils.ac.uk/>; email: [nils@qub.ac.uk](mailto:nils@qub.ac.uk).

<sup>4</sup> ONS-LS data can be accessed in four ways : safe researcher settings in the ONS offices in Titchfield and Newport; SafePods; through an AOC agreement with ONS; or via a home access AOC agreement. NILS and SLS are only accessible through workspaces in a safe setting room. For NILS the safe setting room is at the Northern Ireland Statistics and Research Agency (NISRA) Research Support Unit in Belfast. For SLS the safe setting room is at National Records Scotland in Edinburgh.



In response to these disadvantages, the eDataSHIELD (E-mail Data Aggregation Through Anonymous Summary statistics from Harmonised Individual level Databases) system was developed by researchers at the Scottish Longitudinal Study. This system is an adaptation of the DataSHIELD (Data Aggregation Through Anonymous Summary statistics from Harmonised Individual level Databases) system (see <https://www.datashield.org/>), whereby joint analysis of data from separate data centres is undertaken iteratively by linking a computer in each data centre to an analysis computer. The analysis computer, which does not hold any data, receives summary statistics from generalised linear models from each data centre and then combines them and passes them back to the individual data centres. The interface between the analysis computer and the individual data centres prevents any exchange of the raw data so addressing ethics-related data-sharing concerns related to privacy, confidentiality and the protection of study participants rights (Budin-Ljøsne et al., 2015; Wolfson et al., 2010). However, it is not possible to link the three UK Longitudinal Studies through an analysis computer in this way owing to potential disclosure concerns. The eDataSHIELD protocol provides a work-around for this by allowing the exchange of the summary statistics by email, and so enabling generalised linear models or ordered logistic regression models to be fitted iteratively (Raab, 2013).

Research that has been carried out using data from two or more of the studies, enabled by the eDataSHIELD protocol, includes examining levels of “excess” mortality in Scotland and Glasgow compared to elsewhere in the UK (Ralston et al., 2017), and examining the effect of religion on mortality in Scotland and Northern Ireland (Wright et al., 2017). Ralston et al. (2017) used the SLS and ONS-LS to examine the risk of all cause mortality between 2001 and 2010 and compared it for residents aged 35-74 between Scotland and England and Wales, and between residents of Glasgow (Scotland) and Liverpool/ Manchester (England). They found that after adjusting for age, gender and socio-economic status, all-cause mortality was 9% higher in Scotland versus England and Wales, and 27% higher in Glasgow versus Liverpool/ Manchester (Ralston et al., 2017). Using the SLS and NILS, Wright et al. (2017) compared socio-economic status and mortality among Protestants and Catholics. The researchers found that in fully-adjusted models Catholic men and women in Scotland experienced higher mortality risk compared to Protestants, whereas no such difference was observed in Northern Ireland (Wright et al., 2017).

Researchers who are interested in undertaking comparative research using data from the two or more of the three UK longitudinal studies will need to apply for accredited research projects for each of the studies that they want to include in their research. Therefore, if a researcher wanted to undertake research using data from the ONS Longitudinal Study and NILS, they would need to submit separate research applications for the ONS-LS and NILS and submit them to their respective support units (CeLSIUS and NILS-RSU).

### **International context of the ONS-LS**

The ONS-LS and its sister studies, SLS and NILS, are among several other longitudinal census-based datasets. Similar studies include IPUMS USA (Integrated Public Use Microdata Series USA) and the Australian Census Longitudinal Study. IPUMS USA collects, preserves and harmonises US census microdata and provides access to it. It includes 50 high-precision samples of the American population drawn from fifteen federal censuses. They draw on every surviving census from 1850-2010, and each sample case represents between 20 and 1,000 individuals in the whole US population meaning that weights need to be used in analysis. The data includes harmonised income and occupation variables, but unlike the ONS-LS and its sister studies, there is no guarantee that an individual in one sample is in the one for the next wave of data and there is limited geographic information (IPUMS USA, 2025).

The Australian Census Longitudinal Study (ACLS) comprises four waves of census data (2006-2021: 2006-2011, 2011-2016, 2016-2021). Three ACLS data panels, representing a 5% sample of individuals enumerated in Australia on the 2006, 2011 and 2016 census nights. It began with a 5% panel sample

from the 2006 census, which was linked to the subsequent censuses, with samples from later censuses being formed using the same sample selection strategy. This was designed to maintain a linked sample size of 5% but also introduce new records to subsequent panels to account for new births, migrants and missed links in previous panels. Similarly to the ONS-LS, the ACLS includes people living in both private and communal establishments, but it does not include linkage to administrative data (e.g., births and deaths). Currently it has been used to investigate how family structure has changed over time, the characteristics and changes resulting from single parenthood and whether Australians who were unemployed in 2011 and moved regions by 2016 were more likely to be employed than those who remained in the same area (Australian Bureau of Statistics, April 2024).

A feasibility study for the establishing of a New Zealand Longitudinal Census (NZLC) has also been done (Didham, Nissen, & Dobson, 2014). It examined linking the 1981, 1986, 1991, 1996, 2001 and 2006 Censuses, starting with the linkage of five census pairs (1981-1986, 1986-1991, 1991-1996, 1996-2001, 2001-2006) and then linking the five census pairs to each other, with linkage done using date of birth, sex and area unit of usual residence. The study also examined linking birth and death registrations and international migration data. Despite demonstrating the feasibility of such a longitudinal census and making a series of recommendations including developing confidentiality rules to enable the secure remote access to the NZLC by approved researchers, finalising the linkage of birth and death registrations, and investigating linkage to other sources of administrative data, the NZLC remains as an experimental dataset.

### Challenges to the ONS-LS and other census-based longitudinal studies

A generic threat to census-based longitudinal studies including the ONS-LS is the possible cessation of traditional census approaches, in favour of methods that derive equivalent data from administrative sources. ONS (2023) set out an ambitious intention to create a *"system [that] would primarily use administrative data like tax, benefit and border data, complemented by survey data and a wider range of data sources"* (Office for National Statistics, 2023). This would most likely have removed the requirement for a further traditional census. The potential for use of administrative data has also been considered elsewhere, for example, Australia Bureau of Statistics (2020) (Australian Bureau of Statistics, October 2020) and Stats New Zealand (2022) (Stats NZ, 2022).

Whilst it is now recommended that full censuses are taken in the UK in 2031, a switch to administrative data would be a significant challenge. Administrative records could continue to be attached to the ONS-LS base as long as sufficient individual level data were available to carry out that linkage. However, upstream suppliers of data (government departments, etc.) may be unwilling to provide such data, or legally restricted from doing so. Furthermore, characteristics collected in census results but not readily available through administrative data would not be included. For the purposes of cross-sectional reporting, administrative data could be supplemented with data from regular specialised population surveys. However, any such surveys would have a different sampling framework to the ONS-LS (and indeed could not replicate the ONS-LS sample without disclosing the ONS-LS confidential inclusion dates), and thus could not be appended to the ONS-LS database. The extent to which this challenge is faced by census-based longitudinal studies beyond the UK would depend on the sampling frameworks used, and the linking fields available. If there are common individual identifiers, such as ID numbers, used in different data sources, then linkage would be, at least in principle, straightforward. However, where multiple non-comprehensive identifiers are used, linkage becomes harder.

Other challenges are economic. Whilst such studies offer uniquely rich data, they are also complex and tend to be used only by data and subject specialists. They therefore run the risk of being seen as expensive to maintain, and thus liable to funding cuts.

## Conclusion

The ONS Longitudinal Study, follows a 1% sample of the England & Wales population from the decennial census data (1971 – 2011), linked to births, deaths and cancer registration data. Sample members are selected on the basis of four confidential birthdays, with new study members entering the study through birth on one of the four birthdays or immigration (and being born on one of the four birthdays) and leave through death or emigration. The main strength of the LS is its large sample size (>1 million), making it the largest nationally representative dataset in the UK, and allowing the analysis of small areas and specific population groups. Currently, the ONS-LS has 46 years of follow-up data 1971 – 2017, and it has been used to examine issues with social policy implications, for example social inequalities in health, employment and education; housing and geographical mobility; and family policy. The upcoming linkage of the 2021 Census data to the ONS-LS will extend the span of the study to 50 years, and will enable researchers to examine changes that have taken place in 2011 – 2021 period, a period which saw Brexit and the Covid 19 pandemic.

The ONS-LS is one of the three “sister” studies covering the United Kingdom. While there are similarities between the three studies there are important differences related to their sample size and the additional information that they linked to. That being said all three of the studies provide opportunities for analyses with large sample sizes and they are supported by teams of researchers all of whom have undertaken research using one or more of the studies. Research using data from two or more of the studies is possible enabled by the eDataSHIELD protocol. Indicatively, past research using eDataSHIELD, has examined levels of “excess” mortality in Scotland and Glasgow compared to England and Wales and Liverpool/ Manchester using data from the ONS-LS and SLS (Ralston et al., 2017). It has also been used to examine the effect of religion on mortality in Scotland and Northern Ireland (Wright et al., 2017).

The ONS-LS and its sister studies are among several census-based longitudinal studies worldwide. However, such studies face the on-going threat of the possible cessation of traditional census approaches and the economic challenges posed by their perception as being expensive to run. That being said, census approaches do appear to be continuing, and funders continue to recognise the value such studies offer for research into social inequalities in health, education and employment, social exclusion and mobility, and family policy.

## Acknowledgement and disclaimer

The permission of the Office for National Statistics to use the Longitudinal Study is gratefully acknowledged, as is the help provided by staff of the Centre for Longitudinal Study Information & User Support (CeLSIUS). CeLSIUS is funded by the ESRC under project ES/V003488/1. The authors alone are responsible for the interpretation of the data.

The help provided by staff of the Longitudinal Studies Centre—Scotland (LSCS) is acknowledged. The LSCS is supported by the ESRC/JISC, the Scottish Funding Council, the Chief Scientist’s Office, and the Scottish Government. The authors alone are responsible for the interpretation of the data. Census output is Crown copyright and is reproduced with the permission of the Controller of HMSO and the Queen’s Printer for Scotland.

The help provided by the staff of the Northern Ireland Longitudinal Study/Northern Ireland Mortality Study (NILS/NIMS) and the NILS Research Support Unit is acknowledged. The NILS/NIMS is funded by the Health and Social Care Research and Development Division of the Public Health Agency (HSC R&D Division) and NISRA. The NILS-RSU is funded by the ESRC and the Northern Ireland Government. The authors alone are responsible for the interpretation of the data and any views or opinions presented are solely those of the author and do not necessarily represent those of NISRA/NILS.



## References

- Australian Bureau of Statistics. (April 2024). Microdata and table builder: Australian Census Longitudinal Dataset. Retrieved from <https://www.abs.gov.au/statistics/microdata-tablebuilder/available-microdata-tablebuilder/australian-census-longitudinal-dataset#cite-window1>
- Australian Bureau of Statistics. (October 2020). Assessing administrative data quality to enhance the 2021 Census. Retrieved from <https://www.abs.gov.au/statistics/research/assessing-administrative-data-quality-enhance-2021-census>
- Bourqin, P., Joyce, R., & Sturrock, D. (2021). *Inheritances and inequality over the life cycle: what will they mean for younger generations?* Retrieved from London, UK: [https://ifs.org.uk/sites/default/files/output\\_url\\_files/R188-Inheritances-and-inequality-over-the-lifecycle%252520%2525281%252529.pdf](https://ifs.org.uk/sites/default/files/output_url_files/R188-Inheritances-and-inequality-over-the-lifecycle%252520%2525281%252529.pdf)
- Budin-Ljøsne, I., Burton, P., Isaeva, J., Gaye, A., Turner, A., Murtagh, M. J., . . . Harris, J. R. (2015). DataSHIELD: An Ethically Robust Solution to Multiple-Site Individual-Level Data Analysis. *Public health genomics*, 18(2), 87-96. doi:<https://doi.org/10.1159/000368959>
- Centre for Longitudinal Study Information and User Support. (2018). *Events*. Retrieved from London, United Kingdom: <https://www.ucl.ac.uk/epidemiology-health-care/research/epidemiology-and-public-health/research/health-and-social-surveys-research-group/studies-21>
- Centre for Longitudinal Study Information and User Support. (2022). *Research Project Accreditation Application Guidance*. London: University College London.
- Centre for Longitudinal Study Information and User Support. (2024a, May 2024). CeLSIUS. Retrieved from <https://www.ucl.ac.uk/epidemiology-health-care/research/epidemiology-and-public-health/research/health-and-social-surveys-research-group/studies-10>
- Centre for Longitudinal Study Information and User Support. (2024b). Want to use the ONS LS for your research? Retrieved from <https://www.ucl.ac.uk/epidemiology-health-care/research/epidemiology-and-public-health/research/health-and-social-surveys-research-group/studies-14>
- Centre for Longitudinal Study Information and User Support, & Office for National Statistics. (Eds.). (2014). London, United Kingdom: University College London.
- Didham, R., Nissen, K., & Dobson, W. (2014). *Linking censuses: New Zealand longitudinal census 1981-2006*. Retrieved from Wellington, NZ: <https://www.stats.govt.nz/methods/linking-censuses-new-zealand-longitudinal-census-19812006/>
- Hattersley, L., & Creeser, R. (1995). *Longitudinal Study 1971-1991: History, organisation and quality of data*. London: HMSO Retrieved from <https://www.ucl.ac.uk/epidemiology-health-care/sites/epidemiology-health-care/files/LS7.pdf>
- IPUMS USA. (2025). US Census Data for Social, Economic and Health Research. Retrieved from <https://usa.ipums.org/usa/>
- Irastorza-Fadrique, A., Levell, P., & Parey, M. (2024). *Household responses to trade shocks*. Retrieved from London: <https://ifs.org.uk/publications/household-responses-trade-shocks-0>
- Mastrosavvas, A., & Shelton, N. (2025). *The Longitudinal Impossible Dataset: Helping Users Navigate the ONS Longitudinal Study*. Paper presented at the IASSIST, 2025, Bristol, UK. <https://zenodo.org/records/15611228>
- Murray, E. T., Lacey, R., Maughan, B., & Sacker, A. (2020). Association of childhood out-of-home care status with all-cause mortality up to 42-years later: Office of National Statistics Longitudinal Study. *BMC Public Health*, 20(1), 735. doi: <https://doi.org/10.1186/s12889-020-08867-3>

- Norman, P., & Boyle, P. (2014). Are health inequalities between differently deprived areas evident at different ages? A longitudinal study of census records in England and Wales, 1991-2001. *Health Place*, 26, 88-93. doi: <https://doi.org/10.1016/j.healthplace.2013.12.010>
- Office for National Statistics. (2001). *Census 2001: England Household Form*. United Kingdom: Office for National Statistics.
- Office for National Statistics. (2011). *2011 Census: Household Questionnaire England*. United Kingdom: Office for National Statistics.
- Office for National Statistics. (2012). 2011 Census Geographies. Retrieved from <https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies/2011censusgeographies>
- Office for National Statistics. (2022). Census 2021 Geographies. Retrieved from <https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies/census2021geographies>
- Office for National Statistics. (2023). *The future of population and migration statistics in England and Wales: A consultation on ONS proposals, 29 June 2023*. Retrieved from London: <https://consultations.ons.gov.uk/ons/futureofpopulationandmigrationstatistics/supportingdocuments/Future%20of%20Population%20and%20Migration%20Statistics%20Consultation%20Document.pdf>
- Office for National Statistics. (2024a). Access the data securely: Assured Organisational Connectivity (AOC). Retrieved from <https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/secureresearchservice/accessthe数据安全#assured-organisational-connectivity-aoc->
- Office for National Statistics. (2024b). Become an accredited researcher. Retrieved from <https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/secureresearchservice/becomeanaccreditedresearcher>
- Office of Population Censuses and Surveys. (1991). *1991 Census - England*. United Kingdom: Office of Population Censuses and Surveys.
- Raab, G. (2013). *e-DataSHIELD: E-mail Data Aggregation Through Anonymous Summary statistics from Harmonised Individuals level Databases*. Retrieved from Edinburgh:
- Raab, G., & Dibben, C. (2020). *eDataShield: Running an analysis of combined data when the individual records cannot be combined*. Paper presented at the Scottish Centre for Administrative Data Research, Edinburgh. <https://era.ed.ac.uk/handle/1842/36691>
- Ralston, K., Walsh, D., Feng, Z., Dibben, C., McCartney, G., & O'Reilly, D. (2017). Do differences in religious affiliation explain high levels of excess mortality in the UK? *Journal of Epidemiology and Community Health*, 71(5), 493-498. doi: <https://doi.org/10.1136/jech-2016-208176>
- Ritchie, F. (2017). *The "Five Safes": A framework for planning, designing and evaluating data access solutions*. Paper presented at the Data for Policy, London, UK.
- Sacker, A., Murray, E., Maughan, B., & Lacey, R. (2024). Social care in childhood and adult outcomes. Double whammy for minority children? *Longitudinal and Life Course Studies*, 15(2), 139-162. doi: <https://doi.org/10.1332/17579597Y2023D000000008>
- Stats NZ. (2022). *Experimental administrative population census: Data sources, methods, and quality (second iteration)*. Retrieved from Wellington, New Zealand: <https://www.stats.govt.nz/research/experimental-administrative-population-census-data-sources-methods-and-quality-second-iteration/#:~:text=This%20paper%20discusses%20the%20data%20and%20methods%20used,provides%20information%20about%20the%20quality%20of%20the%20data.>
- Sturley, C., Norman, P., Morris, M., & Downing, A. (2023). Contrasting socio-economic influences on colorectal cancer incidence and survival in England and Wales. *Social Science and Medicine*, 333. doi: <https://doi.org/10.1016/j.socscimed.2023.116138>

- UK Statistics Authority. (2020). *Research Code of Practice and Accreditation Criteria. Section B: Accreditation of Researchers and Peer Reviewers*. United Kingdom: Gov.UK Retrieved from <https://www.gov.uk/government/publications/digital-economy-act-2017-part-5-codes-of-practice/research-code-of-practice-and-accreditation-criteria#section-b-accreditation-of-researchers-and-peer-reviewers>
- UK Statistics Authority. (2024). *Ethics Self-Assessment Tool*. London, United Kingdom: UK Statistics Authority Retrieved from <https://uksa.statisticsauthority.gov.uk/the-authority-board/committees/national-statisticians-advisory-committees-and-panels/national-statisticians-data-ethics-advisory-committee/ethics-self-assessment-tool/>
- University of St. Andrews. (2021). SafePod Network. Retrieved from <https://safepodnetwork.ac.uk/>
- Wolfson, M., Wallace, S. E., Masca, N., Rowe, G., Sheehan, N. A., Ferretti, V., ... Burton, P. R. (2010). DataSHIELD: resolving a conflict in contemporary bioscience--performing a pooled analysis of individual-level data without sharing the data. *International Journal of Epidemiology*, 39(5), 1372-1382. doi: <https://doi.org/10.1093/ije/dyq111>
- Wright, D. M., Rosato, M., Raab, G., Dibben, C., Boyle, P., & O'Reilly, D. (2017). Does equality legislation reduce intergroup differences? Religious affiliation, socio-economic status and mortality in Scotland and Northern Ireland: A cohort study of 400,000 people. *Health & Place*, 45, 32-38. doi:<https://doi.org/10.1016/j.healthplace.2017.02.009>

---

## Endnotes

- <sup>1</sup> Alison Sizer is a CeLSIUS User Support Officer and Associate researcher in the Department of Information Studies, UCL. She can be reached by email: [a.sizer.11@ucl.ac.uk](mailto:a.sizer.11@ucl.ac.uk).
- <sup>2</sup> Andreas Mastrosavvas, Department of Epidemiology and Public Health, University College London. He can be reached by email: [a.mastrosavvas@ucl.ac.uk](mailto:a.mastrosavvas@ucl.ac.uk)
- <sup>3</sup> Oliver Duke-Williams is a Senior Advisor for CeLSIUS and a Professor in the Department of Information Studies, UCL.
- <sup>4</sup> 'Do you have any long-term illness, health problem or handicap which limits the activities or work that you can do?'
- <sup>5</sup> 'What is your religion?'
- <sup>6</sup> 'Do you look after, or give any help or support to family, friends, neighbours or others because of: long-term physical or mental ill-health or disability, or problems related to old age?'
- <sup>7</sup> 'Over the last twelve months would you say that your health has on the whole been: Good; Fairly good; Not good?'
- <sup>8</sup> 'How would you describe your national identity?'
- <sup>9</sup> 'What passports do you hold?'
- <sup>10</sup> 'If you were not born in the United Kingdom, when did you most recently arrive to live here?'

---

<sup>11</sup> 'Including the time you have already spent here, how long do you intend to stay in the United Kingdom?'

<sup>12</sup> The census was also held in Northern Ireland, but in Scotland it was delayed by a year to 21st March 2022.

<sup>13</sup> 'Is the gender you identify with the same as your sex registered at birth?'

<sup>14</sup> 'Which of the following best describes your sexual orientation?', with tick boxes for Straight/Heterosexual, Gay or Lesbian, Bisexual, Other sexual orientation.

<sup>15</sup> 'Who is (was) your legal marriage or registered civil partnership to?', with tick boxes for Someone of the opposite sex, or Someone of the same sex.

<sup>16</sup> 'Have you previously served in the UK Armed Forces?'

<sup>17</sup> Output areas are the lowest level of geographical area used for census statistics. They were created for the 2001 Census and are made up of 40-250 households (100 – 625 individuals).

<sup>18</sup> Lower Super Output Areas usually comprise 4 - 5 OAs and 400 - 1,200 households (1,000 – 3,000 individuals).

<sup>19</sup> Middle Super Output Areas usually comprise 4 – 5 LSOAs and 2,000 – 6,000 households (5,000 – 15,000 individuals).

<sup>20</sup> The “five safes” framework comprise a set of principles: safe data, safe projects, safe people, safe settings and safe outputs.

<sup>21</sup> Non-members are the household members who are living with an LS member at the time of the census.

<sup>22</sup> A SafePod is a standardised safe setting that provides the security and controls for data that requires secure access for research. A SafePod includes a door control access system, CCTV, a researcher area for dataset analysis, secure IT cupboard and a height adjustable desk.