IOE, Faculty of Education and Society

# 1958 National Child Development Study

## Age 62 Geographical Identifiers

User Guide (Version 1)

June 2025

CENTRE FOR LONGITUDINAL STUDIES

UKRI Economic and Social Research Council

## Contact

Data queries: help@ukdataservice.ac.uk

Questions and feedback about this user guide: clsdata@ucl.ac.uk.

## Authors

David Church, Liam Curran, George Ploubidis.

## How to cite this guide

Church, D., Curran, L., Ploubidis, G.B. (2025) *1958 National Child Development Study: Age 62 Geographical Identifiers*. London: UCL Centre for Longitudinal Studies

This guide was published in June 2025 by the UCL Centre for Longitudinal Studies.

## Data citation and acknowledgement

You should cite the data and also acknowledge CLS following the guidance from cls.ucl.ac.uk/data-access-training/citing-our-data/

## Centre for Longitudinal Studies

Centre for Longitudinal Studies (CLS)

UCL Social Research Institute

University College London

20 Bedford Way, London WC1H 0AL

cls.ucl.ac.uk

The UCL Centre for Longitudinal Studies (CLS) is an Economic and Social Research Council (ESRC) Resource Centre. It is home to a unique series of UK national cohort studies. It is part of the UCL Social Research Institute, based at the IOE, UCL's Faculty of Education and Society.

This document is available in alternative formats. Please email the Centre for Longitudinal Studies at clsdata@ucl.ac.uk

# Contents

# About the 1958 National Child Development Study

The National Child Development Study (NCDS) is a longitudinal birth cohort study, following a nationally representative sample of over 17,000 people born in Britain in a single week in March 1958.

Cohort members have been surveyed throughout their lives, since birth, creating an exceptionally rich resource for a wide range of research.

The study data reveal the long lasting influence of childhood, illustrating how past experiences can resonate over time, as well as the interplay between the various aspects of individuals' lives.

NCDS has equipped policymakers with robust evidence in areas as diverse as smoking in pregnancy, educational inequalities, adult basic skills, and social mobility. Today, with the cohort now in their sixties, the study is casting light on how people experience retirement and ageing in the 21st century.

## 1. Introduction

The NCDS Age 62 Survey, (or 'Life in Your Early 60s' Survey as known to study members) was conducted between 2020 and 2024 when participants were aged 61-65 years. Prior surveys took place at birth, 7,11,16,23,33,42,44,46,50 and 55.

This sweep was designed and managed by the Centre for Longitudinal Studies (CLS) at the UCL Social Research Institute. Interviewer fieldwork was conducted by NatCen and Verian (formerly Kantar). Health visits were conducted by NatCen and INUVI.

The Age 62 Survey involved an interview, a health visit, two paper self-completion questionnaires and an online dietary questionnaire.

The broad aim of the Age 62 Survey was to collect information which would aid the understanding of the lifelong factors affecting retirement and ageing. This survey also had a biomedical focus with physical measurements and assessments being conducted for the first time since the Age 44 biomedical sweep. The data collection

built on the extensive data collected previously from birth and across the lifetime of study members and will facilitate comparisons with other generations as they reach the same life stage, allowing for study of social change.

The study was initially planned and designed to be conducted in-person. Fieldwork commenced in January 2020 but was subsequently paused in March 2020 due to the COVID-19 pandemic. As in-person interviewing was not feasible until early 2022, the protocol was adapted so that interviews could be conducted by video-call. Interviewer fieldwork restarted by video call in spring 2021 until April 2022 when it was feasible to return to in-person interviewing. The video mode option continued to be available if requested by a cohort member or was required due to interviewer capacity issues in a particular area.

Once mainstage fieldwork was complete, those who had not participated were invited to complete a short version of the questionnaire via web (known as the 'mop-up' survey).

Cohort members who completed the survey between January-March 2020, were also invited to take part in the mop-up survey in order establish how their circumstances might have changed since the pandemic.

Emigrants were not invited to take part in the main survey but were invited to take part in this short web-survey. For more information see the NCDS age 62 user guide.

# 2. Geographical indicators

## 2.1 Background

There has been an increasing awareness of the value of geographically linked data in social scientific research, especially since the 'GIS revolution' of the early 1990s (Longley and Batty, 1996). Spatial data can be approached from several directions. For example, "longitudinal studies are particularly valuable to geographers because they chart change, collect information across various domains and are spatially referenced" (Ekinsmyth, 1996: 364). On the other hand, economists are beginning to appreciate the value of spatially referenced data, for example in research into the economics of education (e.g. Gibbons et al, 2013 who used the National Pupil

Database to estimate the effects of neighbourhood composition on teenagers' behavioural and educational outcomes in England).

Epidemiology and its associated disciplines are perhaps most consistently associated with investigating the spatial effects of the type of data collected across longitudinal cohort studies. For example, Christakis and Fowler (2007) used data from the Framingham Heart Study in the US to examine the spread of obesity in a large social network over 32 years, while Tunstall et al (2010) used data from the Millennium Cohort Study to analyse the health outcomes of pregnant women who moved house. Two particularly fruitful fields are, firstly, the investigation of so-called 'neighbourhood effects' across a number of socio-economic domains (e.g. Lupton and Kneale (2012) used data from the 1970 British Cohort Study to investigate neighbourhood influences on teenage parenthood) and, secondly, network-based analyses of particular issues such as obesogenic environments (e.g. Burgoine et al, 2014), accessibility to health-promoting community resources (e.g. Wolch et al, 2011) and the impact of built environment (morphological) characteristics on health and well-being (e.g. Sarkar et al, 2014).

CLS takes the approach that access to geo-referenced data below Government Office Region (GOR) level should be subject to increasing access restrictions the more likely the data is to reveal the identity of cohort members. Other data limitations include non-uniformity of geo-identifiers used across different sweeps of the various cohort studies and varying levels of accuracy in terms of the geo-identifiers collected (a particular problem of early sweeps before the standardisation of unit postcodes).

## 2.2 Extent and nature of the geographical data

NCDS cohort members live throughout England, Scotland and Wales, therefore a short explanation of the extent and nature of the data, and differences across the constituent nations of Great Britain may be helpful. England and Wales use the same naming conventions across different geographies. Post-devolution, Scotland has adopted slightly different naming conventions. For example, in both 2001 and 2011 Census geography, what are known as 'Lower Super Output Areas' and 'Middle Super Output Areas' in England and Wales are called 'Data Zones' and

‘Intermediate Geographies’ respectively in Scotland and the mean populations used to create these areal units also varies between England &Wales and Scotland. The projected coordinate system used to display geo-referenced data across Great Britain (i.e. England, Wales and Scotland) is the British National Grid (espg:27700). The range of Ordnance Survey products (e.g. MasterMap, AddressBase, OpenData) is available for Great Britain (i.e. excluding Northern Ireland). Northern Ireland uses its own equivalent to the ONS Postcode Directory, called the Central Postcode Directory, and the process of spatialising geo-referenced data works in the same way as with Great Britain data.

# 3. Linkage methods

To enable the spatial analysis of longitudinal cohort study data, unit postcodes are gathered from the addresses collected during interview, which are then validated by CLS using a range of specialist software products from AFD[1]. This postcode data is then used to generate point data, usually within a GIS. There are several licensed and open-source GIS packages available (e.g. ArcGIS[2], MapInfo[3] and QGIS[4]). The primary data source for spatialising longitudinal cohort study data within this software is the ONS Postcode Directory, available from the Office for National Statistics website[5]. This dataset has been released quarterly since 2004 (every February, May, August and November) and contains Ordnance Survey eastings and northings for each unit postcode centroid. These eastings and northings are spatialised in GIS in the form of 'x', 'y' points, usually to an accuracy of within 1 metre of the mean postcode centroid[6]. The August 2024 ONS Postcode Directory was used to link NCDS sweep 10 to the set of geographical indicators described in Section 4.5. The NCDS sweep 10 data was linked to the August 2024 ONSPD on postcode (specifically, PCDS) by means of a 'one-to-first' join in ArcGIS Pro (i.e. the participant's postcode will be joined to the first matching postcode in the ONSPD data).

---

[1] http://www.afd.co.uk/

[2] Licensed software, available from http://www.esriuk.com/

[3] Licensed software, available from http://www.mapinfo.com/

[4] Open-source software, available from https://www.qgis.org/

[5] https://www.ons.gov.uk/methodology/geography/geographicalproducts/postcodeproducts

[6] There are, however, a range of 'grid reference positional quality indicators', ranging from 1 ('within the building of the matched address closest to the postcode mean' to 9 ('no grid reference available'). Only 0.68% of current postcodes in the August 2024 ONSPD have a positional quality indicator greater than 1.

# 4. Description of the linked geographical data

## 4.1 Licensing and data access

The NCDS age 62 geographical data have been processed by CLS and supplied to the UK Data Service. All data users need to be registered with the UK Data Service and sign the UK Data Service End User Licence. Details of how to do this are available at ukdataservice.ac.uk/get-data/how-to-access/registration.

The NCDS age 62 geographical identifiers are potentially disclosive and can be accessed as controlled data from the UK Data Service SecureLab. Applicants wishing to access this data need to abide by the terms and conditions of the UK Data Service Secure Access licence.

To gain access to the geographical identifiers, researchers must submit a Secure Access application to the UK Data Service detailing the intended analysis and provide a justification as to why this data is requested. Application guidance can be found at ukdataservice.ac.uk/find-data/access-conditions/secure-application-requirements/apply-to-access-non-ons-data/

Subject to approval, researchers can upload their own datasets to their UK Data Service SecureLab account to link to these geo-identifiers (e.g. linking Census, meteorological, crime, housing or other socio-economic data to the cohort studies).

## 4.2 List of datasets

Datasets are flat in structure, with one row per cohort member.

There are two versions of NCDS age 62 geographically linked data available, one set based on 2011 Census boundaries and the other set based on 2021 Census boundaries.

**Table 1: List of available datasets**

| Name of the dataset | Content summary |
|---|---|
| ncds10_age62_geographical_identifiers_2011_restricted | NCDS age 62 geographical identifiers: 2011 Census boundaries |

| | |
|---|---|
| ncds10_age62_geographical_identifiers_2021_restricted | NCDS age 62 geographical identifiers: 2021 Census boundaries |

## 4.3 Data documentation

The following data dictionaries contain the following metadata: variable names and labels, value labels, missing values. The order of the variables in the metadata documents is consistent with the variable ordering in the datasets.

**Table 2: Data documents**

| Name of the document | Content summary |
|---|---|
| ncds10_age62_geography_2011_data_dictionary.xlsx | Variable names and labels, value labels, missing values. |
| ncds10_age62_geography_2021_data_dictionary.xlsx | Variable names and labels, value labels, missing values. |

## 4.4 Identifiers

### Individual identifiers

The data are identified with the variable NCDSID. This is the same research ID used for the rest of the cohort data available at the UK Data Service. This enables the data to be easily merged with one another.

## 4.5 Dataset description

**Table 3: NCDS Age 62 Geographical Identifiers (2011)**

| Variable Name | Description |
|---|---|
| NCDSID | NCDSID Serial Number |
| N10COUNTRY | Country of interview |
| N10RGN | December 2020 Region of interview |
| N10OSWARD | May 2023 Ward |
| N10CASWARD | Census Area Statistics Ward |

| Variable Name | Description |
|---|---|
| N10OSLAUA | April 2023 Local Authority |
| N10PCON | July 2024 Westminster Parliamentary Constituency |
| N10IMD | IMD Overall Rank - Eng 2019, Sco 2020, Wal 2019, NI 2017 |
| N10IMDD | IMD Overall Rank Decile |
| N10RU11IND | 2011 Census Rural-Urban Classification |
| N10WZ11 | 2011 Census Workplace Zone |
| N10OA11 | 2011 Output Area |
| N10LSOA11 | 2011 Lower Super Output Area |
| N10MSOA11 | 2011 Middle Super Output Area |
| N10OAC11 | 2011 Output Area Classification |

**Table 4: NCDS Age 62 Geographical Identifiers (2021)**

| Variable Name | Description |
|---|---|
| NCDSID | NCDSID Serial Number |
| N10COUNTRY | Country of interview |
| N10RGN | December 2020 Region of interview |
| N10OSWARD | May 2023 Ward |
| N10CASWARD | Census Area Statistics Ward |
| N10OSLAUA | April 2023 Local Authority |
| N10PCON | July 2024 Westminster Parliamentary Constituency |
| N10IMD | IMD Overall Rank - Eng 2019, Sco 2020, Wal 2019, NI 2017 |
| N10IMDD | IMD Overall Rank Decile |
| N10OA21 | 2021 Output Area |
| N10LSOA21 | 2021 Lower Super Output Area |
| N10MSOA21 | 2021 Middle Super Output Area |

See the Appendix for the data sources from which the NCDS Age 62 Geographical Identifiers are derived.

## 4.6 Output disclosure control

Access to this controlled data is only available via the UK Data Service SecureLab.

The UK Data Service will always perform a certain level of disclosure control on the outputs generated by researchers, as outlined in their SDC Handbook which can be downloaded from: securedatagroup.org/sdc-handbook/

The two UK Data Service Secure Lab rules of thumb that will be applied to all outputs are:

- Threshold rule: No cells should contain less than 10 observations
- Dominance rule: No observation should dominate the data to a huge extent

# 5. References


Burgoine, T., N. G. Forouhi, et al. (2014). "Associations between exposure to takeaway food outlets, takeaway food consumption, and body weight in Cambridgeshire, UK: population based, cross sectional study." Bmj **348**.

Christakis, N. A. and J. H. Fowler (2007). "The spread of obesity in a large social network over 32 years." New England journal of medicine **357**(4): 370-379.

Ekinsmyth, C. (1996). "Large-scale longitudinal studies: their utility for geographic enquiry." Area: 358-372.

Gibbons, S., O. Silva, et al. (2013). "Everybody needs good neighbours? Evidence from students' outcomes in England." The Economic Journal **123**(571): 831-874.

Longley, P. A. and M. Batty (1996). Spatial analysis: modelling in a GIS environment, John Wiley & Sons.

Lupton, R. and D. Kneale (2012). Theorising and measuring place in neighbourhood effects research: The example of teenage parenthood in England. Neighbourhood Effects Research: New Perspectives. M. van Ham, D. Manley, N. Bailey, L. Simpson and D. Maclennan. Berlin, Springer.

Sarkar, C., C. Webster, et al. (2014). Healthy Cities: Public Health Through Urban Planning, Edward Elgar Publishing.

Swanson, J., T. Vincent, et al. (2014). "Relative accuracy of grid references derived from postcode and address in UK epidemiological studies of overhead power lines." Journal of Radiological Protection **34**(4): N81.

Tunstall, H., K. Pickett, et al. (2010). "Residential mobility in the UK during pregnancy and infancy: Are pregnant women, new mothers and infants 'unhealthy migrants'?" Social Science & Medicine **71**(4): 786-798.

Wolch, J., M. Jerrett, et al. (2011). "Childhood obesity and proximity to urban parks and recreational resources: A longitudinal cohort study." Health & place **17**(1): 207-214.

# Appendix: Data sources for geographical identifiers

**(Please see overleaf)**

| Dataset | Description | Boundary Data Source (ONS Postcode Directory August 2024) |
|---|---|---|
| Interview Wards | 2023 Ward Boundaries<br><br>2003 Census Area Statistic Ward[7] | English Electoral Wards, May 2023<br>Welsh Electoral Wards, May 2023<br>Scottish Electoral Wards, May 2023<br><br>English Census Area Statistic Wards, 2003<br>Welsh Census Area Statistic Wards, 2003<br>Scottish Census Area Statistic Wards, 2003 |
| Interview Output Area (OA) | 2011 Output Area<br><br>2021 Output Area | English Output Areas, 2011<br>Welsh Output Areas, 2011<br>Scottish Output Areas, 2011<br><br>English Output Areas, 2021<br>Welsh Output Areas, 2021<br>Scottish Output Areas, 2021 |
| Interview Lower Super Output Area (LSOA) | 2011 Lower Super Output Area<br><br>2021 Lower Super Output Area | English Lower Super Output Areas, 2011<br>Welsh Lower Super Output Areas, 2011<br>Scottish Datazones, 2011<br>English Lower Super Output Areas, 2021 |

[7] Please see the ONS publication 'A Beginner's Guide to UK Geography 2023' for definitions and an overview of the various geographical units.

| Dataset | Description | Boundary Data Source (ONS Postcode Directory August 2024) |
|---|---|---|
| | | Welsh Lower Super Output Areas, 2021<br>Scottish Data Zones , 2022 |
| Interview Middle Super Output Area (MSOA) | 2011 Middle Super Output Area | English Lower Super Output Areas, 2011<br>Welsh Lower Super Output Areas, 2011<br>Scottish Intermediate Geographies, 2011 |
| | 2021 Middle Super Output Area | English Lower Super Output Areas, 2021<br>Welsh Lower Super Output Areas, 2021<br>Scottish Intermediate Geographies, 2022 |
| Interview Local Authority District | 2023 Local Authority District/Unitary Authority | English Administrative Districts, April 2023<br>English Unitary Authorities, April 2023<br>Welsh Unitary Authorities, April 2023<br>Scottish Council Areas, April 2023 |
| Interview Westminster Parliamentary Constituency | 2014 Westminster Parliamentary | English Westminster Parliamentary Constituencies, December 2014<br>Welsh Westminster Parliamentary Constituencies, December 2014<br>Scottish Westminster Parliamentary Constituencies, December 2014 |