IOE, Faculty of Education and Society

1958 National Child Development Survey

Age 62 Survey

User Guide (Version 1)

July 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

Brown Matt, Silvia Mendonca, Nureen Mohamad Zaki, Vanessa Moulton, Andrew Peters, George Ploubidis, Aida Sanchez-Galvez, Carole Sanchez, Michaela Sedovic, Vilma Agalioti-Sgompou, Richard Silverwood, Bożena Wielgoszewska, Liam Wright.

How to cite this guide

Brown, M., Mendonca, S., Mohamad-Zaki, N., Moulton, V., Peters, A., Ploubidis, G., Sanchez-Galvez, A., Sanchez, C., Sedovic, M., Agalioti-Sgompou, V., Silverwood, R., Wielgoszewska, B., Wright, L. (2025) *National Child Development Survey: Age 62 Survey (Version 1)*. UCL Centre for Longitudinal Studies.

You should cite the data and 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 incredibly rich resource for a wide range of research. The study data show the very long roots of childhood, how past experiences can reverberate through the years, and the interplay between the different facets of people's 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.

Important note about figures in this document

Figures that are presented in this document may vary in comparison with the shared data and figures reported in the NCDS Technical Report. This happens for various reasons: requests for data deletion, resolution of duplicate cases, data editing and quality checking which can result in the removal of cases.

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.

A full account of the survey development and fieldwork procedures can be found in the National Child Development Study technical report and appendices produced by NatCen Social Research, which accompanies this data.

This user guide provides information about the data collected in the 1958 National Child Development Study Age 62 Survey and accompanies the data at the UK Data Service. In addition to this user guide, the Age 62 Survey data documentation includes:

- NCDS - Age 62 Survey: Questionnaire
- NCDS - Age 62 Survey: Childhood self-completion questionnaire (also known as – the Life History Questionnaire/LHQ)
- NCDS – Age 62 Survey: Your Life Now self-completion questionnaire
- NCDS - Age 62 Survey: Technical Report
- NCDS Age 62 Survey: Technical Report – Appendix A – Fieldwork Documents
- NCDS – Age 62 Survey Appendix B Biomedical Protocols.

Data, questionnaires and user guides for all previous sweeps are also available at UKDS. All datasets use a common ID – NCDSID.

2. History and background

The 1958 National Child Development Study (NCDS) is a renowned national longitudinal birth cohort study run by the Centre for Longitudinal Studies (CLS) at the UCL Social Research Institute.

NCDS commenced as the Perinatal Mortality Survey, which was designed to examine the social and obstetric factors associated with stillbirth and infant mortality. Since this initial birth sweep, there have been eleven major follow-ups across childhood, adolescence, and adulthood (including this current survey), gathering information from respondents living in England, Scotland and Wales. With each successive attempt, the scope of enquiry has broadened from a strictly medical focus at birth, to encompass physical and educational development at ages seven (1965), 11 (1969) and 16 (1974), and then to include economic development, children and other wider factors at ages 23 (1981), 33 (1991), transitioning into midlife and further focus on issues such as health, care responsibilities and pension planning at 42 (2000), 44 (2002), 46 (2004), 50 (2008), 55 (2013) and 62 (2020).

In addition to these core sweeps, a series of three web surveys were conducted between 2020 and 2021 which explored the impact of the COVID-19 pandemic on participants.

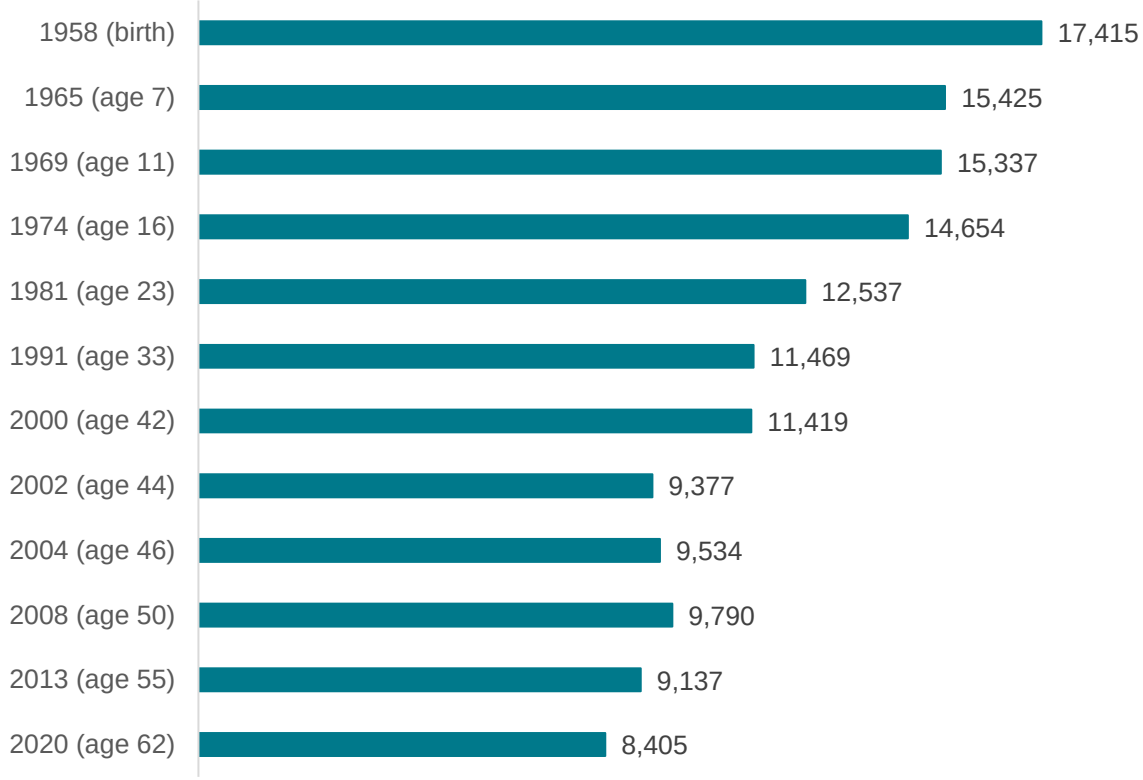
NCDS has equipped policymakers with robust evidence in areas as diverse as smoking in pregnancy, educational inequalities, adult basic skills, and social mobility. Today, with the cohort now in their sixties, the study is casting light on how people experience retirement and ageing in the 21st century.

The Centre for Longitudinal Studies (CLS) at the UCL Social Research Institute has managed the study since 1991, with funding from the Economic and Social Research Council (ESRC).

The chart below shows the number of interviews completed in each NCDS survey. The majority of sweeps were conducted in-person with the exception of the Age 46 sweep (telephone survey) and the Age 55 sweep, which was a mixed mode survey

using both web and telephone. The Age 44 sweep was a biomedical sweep, conducted in-person with a nurse.

Figure 1: Number of interviews per sweep of NCDS



3. The Age 62 Survey

The main aim of the Age 62 Survey was to gather information to understand lifelong factors affecting retirement and ageing and their determinants over the life course and for the first time to collect biomedical data since Age 44 sweep. This survey builds on the extensive data collected since the participants' birth, allowing comparisons with other generations including the 1946 National Survey of Health and Development (NSHD) cohort at age 60-64, allowing for study of social change.

The Age 62 Survey collects information on many aspects of cohort members' lives including employment, relationships, health and wellbeing, cognition, politics, attitudes, caring responsibilities, retirement, as well as capturing pensions and finances. The Age 62 Survey involved the following elements:

- main interview (conducted in-person or by video) - which included cognitive assessments, collection of data linkage consents and a self-completion section containing the most sensitive questions
- two paper self-completion questionnaires
- an Online Dietary Questionnaire (ODQ)
- health visit – which included seated and standing blood pressure, grip strength, collection of blood samples, anthropometry (weight and body fat, waist and hip circumference), timed normal walking speed and standing balance (leg raise).

In addition to core funding provided by the Economic and Social Research Council (ESRC), additional funding for this sweep was also provided by The Department for Work and Pensions (DWP), The Medical Research Council (MRC) and U.S. National Institutes of Health (NIH).

The lead fieldwork contractor was NatCen, who subcontracted Verian (formerly Kantar) to jointly conduct interviewer fieldwork. NatCen also worked in collaboration with INUVI on the biomedical fieldwork.

Ethical approval was provided by London –Fulham Research Ethics Committee. REC reference –18/LO/2034.

4. Fieldwork

Full details regarding the conduct of fieldwork are provided in the NCDS Age 62 Survey Technical report. A summary of the key details are provided here.

4.1 Issued sample

In the initial ‘birth’ sweep of NCDS, all babies born in the UK during one week in 1958 were selected. During the surveys at ages 7, 11 and 16, the cohort was augmented by additional children who were born outside Great Britain, but within the target week in 1958, and subsequently moved to and were educated within Britain.

The initial issued sample for the Age 62 Survey consisted of 11,493 cohort members in total. Participants were issued to the survey unless they were either: known to be deceased; had permanently withdrawn from the study; lived outside of Great Britain; were long-term untraced; or in prison.

A further sample of 327 cohort members known to be living abroad (known as ‘emigrant’ cases) were issued to the ‘mop-up’ phase. These cases were not eligible for the pilots and main stages of the survey.

4.2 Fieldwork period

Mainstage fieldwork began in January 2020 but was paused due to the COVID-19 pandemic in March 2020. It restarted again in November 2021, and continued until November 2023. Health visit fieldwork finished in December 2023. The mop-up web survey was conducted from March to April 2024.

4.3 Fieldwork stages

The Age 62 Survey was conducted in several stages. The initial main stage fieldwork was preceded by three pilot stages conducted in-person: pre-pilot (2018, biomedical collection only), pilot (2019), dress rehearsal (2019). Main stage fieldwork initially took place between January and March 2020, using an in-person only approach.

As a result of the COVID-19 pandemic in-person fieldwork was paused and adaptions were made to the survey so that it could be conducted by video-call. A video-pilot study took place between April 2021 and July 2021.

Main stage interviewing was re-started from November 2021 using a video-only approach until it was feasible to re-launch in-person interviewing from spring 2022. Following the end of main stage interviewing in November 2023, an online 'mop-up' survey was conducted where non-respondents (and emigrants) were invited to complete an abbreviated web version of the survey. Those cohort members who took part in the initial main stage prior to COVID were also re-invited to complete the mop-up survey – in order to understand how their circumstances might have changed since the pandemic.

All stages (except the pre-pilot) were conducted with NCDS study members. Data collected during the pilots has been included in the data (except for the pre-pilot data, and any questions which were removed or significantly changed after the pilots). Each of the stages is described below.

Information about stage participation is provided by two variables. Variable '**n10survey_stage**' denotes the stage a case was completed in. Variable **'n10part**' provides information on whether a respondent has **taken part once** either as part of the main stage, dress rehearsal or pilots or once as a mop-up case, or **twice** as part of the main stage and then again in the mop-up survey.

4.3.1 Pre-pilot

The pre-pilot ran in September and October 2018 to test biomedical assessments.

Several biomedical assessments not conducted previously on NCDS were planned to be included in the Age 62 Survey, namely dynamic blood pressure (how blood pressure responds to exercise) and centrifuging blood samples in the home. The primary aim of the pre-pilot was to assess the feasibility of conducting these in the home. New protocols for collection of grip strength and balance were also evaluated.

The pre-pilot was not conducted with NCDS participants but with members of the public who had taken part in a previous NCDS pilot (as part of the Age 55 Survey) and some additional recruitment. 36 people took part.

Dynamic blood pressure measurement was dropped after the pre-pilot. Data from this stage is not included in the Age 62 data.

4.3.2 In-person pilot

This stage took place in February-March 2019 and was conducted in-person. The purpose of the pilot was to fully test the whole survey before the next phase and identify areas for improvement. As such, the pilot resembled the design of the main stage and health visit as closely as possible in terms of procedures. In this pilot, 105 NCDS cohort members were issued, 80 took part in the main interview and 72 in the health visit[1].

A small number of changes were made to the questionnaire and documents based on feedback from interviewers, healthcare professionals and cohort members. A key finding was that the health visit was longer than planned and as a result some amendments to protocols were agreed. It was initially planned that 3 seated blood pressure readings would be taken – following the pilot this was reduced to 2. Similarly, it was initially planned that 2 waist and hip measurements would be taken – which was reduced to 1. Height measurement was also dropped and instead this information was either fed forward from a previous data collection or based on self-reported height from the main interview.

4.3.3 Dress rehearsal

The dress rehearsal was conducted in-person and took place between June and September 2019 to test all survey components.

A total of 288 NCDS cohort members were invited to take part, 205 participated in the main interview and 181 in the Health Visit.

Due to the extensive piloting which took place prior to the dress rehearsal, only minimal changes were required following this stage to improve clarification of some

[1] Health visits continued beyond the 'official' end of fieldwork stage – up until the beginning of May.

question wording. A number of improvements were also identified for the training of fieldworkers.

4.3.4 Video Pilot

Video interview pilot fieldwork took place between 28th April 2021 and 6th July 2021.

It was conducted following the COVID-19 pandemic to assess the feasibility of using video call interviewing via MS Teams as an alternative mode of interview, when in-person interviewing was not feasible. While this had previously been trialled on the 1970 British Cohort Study (BCS70) using a small, selected sample of study members, the NCDS video pilot was carried out with a larger, representative sample – 310 cohort members were invited to take part and 142 participated.

The MS Teams video interview mirrored the CAPI interview, with adjustments made for the video format which included:

The sharing of showcards, questionnaire programme and leaflets via the 'share screen function'.

The self-completion section of the survey was administered as a short web survey during the video call (where possible) with the option of this being e-mailed to be completed after the survey.

The cognitive assessments were completed in a comparable format, the key adaption being the sending of the 'letter cancellation' document in advance of the survey in a sealed envelope to be opened when instructed by the interviewer.

The collection of data linkage consents also changed to be collected electronically during the interview (instead of via a signed paper consent form).

The video pilot confirmed that it was feasible to conduct the survey using a video interview mode and as such this mode was used to re-start the main stage.

4.3.5 Main Stage

The main stage took place in two phases: the ‘initial’ phase and the ‘re-start’ phase.

The ‘initial’ main stage was conducted prior to the COVID pandemic between January and March 2020 with 1377 cohort members taking part.

The ‘re-start’ phase of the main stage began in November 2021, with interviews being conducted by video mode until April 2022. From May 2022 in-person interviewing restarted and became the main mode of interview. Video interviews continued to be an option for those requesting this or in areas where there was a shortage of in-person interviewers. A total of 6,005 cohort members took part in the ‘re-start’ phase.

In the re-start main stage, there were slightly different protocols depending on whether the interview took part by video or in-person. The key differences are noted below:

- In in-person interviews, sensitive questions were completed privately by participants on the interviewer’s laptop during the interview – though they could request interviewer assistance if necessary). In video interviews they were completed by web (either during the interview or sent by e-mail to be completed after). Interviewer assistance was also possible. As noted below, this was not an option available during the video pilot stage.
- In in-person interviews, paper documents, including the paper questionnaire and letter cancellation document could be provided in person and collected by the interviewer (an option was still available to post these back). These could only be provided and returned by post in a video interview.
- During the video-only phase – all contact and tracing was remote (via telephone and email).

Video interviewing in the ‘re-start’ phase followed the protocols tested during the video pilot. One key adaption was to give study members who did not want to complete the sensitive questions via web questionnaire (during the interview) the option to complete these sensitive questions with assistance from the interviewer. If agreed, the interviewer would share their screen for the participant to read the questions and provide the interviewer with the number of their response (to be

entered into the CAPI). If they did not agree to taking part with assistance from the interviewer, the web survey was sent to them to complete after the interview.

In total 10,851 cases were issued to the main stage (initial pre-COVID and re-start phases), 2,139 interviews were conducted by video and 5,243 took place in-person.

4.3.6 Re-issues and reallocations

Between August 2022 and November 2023, cases that were invited to take part at the main and pilot stages but were unable to do so (unproductive cases), were re-issued or reallocated. Reallocated cases were those that were initially approached by a video interviewer and had requested an in-person interview. A total of 2,496 (22%) of cases were reissued (or reallocated) across two batches. A total of 842 productive interviews were achieved during this stage.

Cases who were approached by an in-person interviewer and requested a video call interview were reallocated through-out the in-person fieldwork period.

4.3.7 The Health Visit

A total of 7,774 cohort members who took part in either a video call or in-person interview were asked if they would be willing to be contacted by a healthcare professional to take part in a health visit. Of these 7,166 agreed to be contacted and 6,309 study members took part in this visit. It was initially intended health visits would take place shortly after the main interview (around 2 weeks). However, due to COVID restrictions this became unfeasible and the gaps between interviews and health visits were in many cases significant.

4.3.8 Mop-up Survey

Following completion of main stage fieldwork, a web survey was conducted from March to April 2024. Cohort members who had not responded to the main interview, including emigrants (who were not part of the main survey) were invited to participate. A total of 2,484 non-responders were invited to take part and 406 responded. A total of 327 emigrant cases were invited to take part and 190 responded.

Those who took part in the main survey before COVID-19, i.e. those interviewed between January and March 2020 as well as pilot and dress rehearsal cases, were also invited to take part in this mop-up survey. These cases were invited to participate in order to collect information on key changes in their circumstances since the pandemic. A total of 1,603 such participants were invited and 1,046 took part. For these participants data is available at two time points.

Further information on the breakdown of cases taking part in the mop-up survey is available in variable '**n10mu_samptype'**, such as whether the case was a non-responder to the main survey, an emigrant case or a pre-COVID case participating for a second time.

4.3.9 Fieldwork timescales and mode

Table 1 below provides a summary of the fieldwork stages.

Table 1: All stages and fieldwork dates for first issue and reissue cases

	Fieldwork mode	Start Date	End Date
Pre-pilot (not with NCDS participants)	In-person - including health visit only	12/09/2018	09/10/2018
Pilot	In-person - including health visit	12/02/2019	31/03/2019
Dress rehearsal	In-person - including health visit	27/06/2019	30/09/2019
Wave 1	In-person - including health visit	07/01/2020	17/03/2020
Wave 2	In-person - including health visit	03/03/2020	17/03/2020
Video pilot	Video-only	28/04/2021	06/07/2021
Wave 2.5 (cases from wave 1 and 2 not completed due to COVID pause)	Video-only	01/11/2021	07/11/2023
Wave 3	Video-only	17/02/2022	07/11/2023
Wave 4	In-person - including health visit*	12/04/2022	07/11/2023
Wave 5	In-person - including health visit*	12/10/2022	07/11/2023
Wave 6	In-person - including health visit*	09/12/2022	07/11/2023
Reissue / Reallocation (Batch 1)	In-person - including health visit*	17/08/2022	07/11/2023
Reissue / Reallocation (Batch 2)	In-person - including health visit*	27/04/2023	07/11/2023
Mop-up	Web only	18/03/2024	12/04/2024

*Video interviews available on request and where no in-person interviewer available

The interview mode can be identified in the data by variable **'survey_mode'**. The data also provides information on the date of interview **n10intdate_month**' (month) and '**n10intdate_month**' (year).

The data from the main-stage and the mop-up survey are contained in the same dataset. For those who took part in the main stage prior to COVID and took part again during the mop up phase, variables related to the second participation have the suffix 'r2'. For example, for such cases the date of the main stage interview will be provided in variables **n10intdate_month'** (month) and **n10intdate_year'** (year), and the date of the mop-up interview in variables '**n10intdate_month_r2'** (month) and **n10intdate_year_r2'** (year).

4.4 Contact Strategy

At the start of each wave (including the pilots and dress rehearsal), NatCen sent advance mailings by post and email to all cohort members informing them of the latest survey. For video only waves, cohort members were invited to have a video interview although the invite noted if they would prefer an in-person interview this could happen later. The letters used in waves 5 to 6 outlined that the interview could take part in-person or by video interview.

The advance mailings included a booklet which provided more detailed information about the survey. Emails included a link to an electronic version of the leaflet.

Following the advance mailing, interviewers would then contact the study members. How contact was made would vary depending on when they last took part in an interview, what contact information was available and whether the case was allocated to a video call only wave or in-person wave. In the video call waves interviewers first contacted cohort members by telephone, making at least three attempts before then sending a text or email. Approximately 91% of cases were first contacted by telephone or email.

For the remaining 9% of cases, initial contact was made through personal visits. Interviewers left calling cards at unattended homes, offering a Freephone number for rescheduling or opting out.

For reissue waves, no new advance letters or emails were sent initially, as cases were allocated gradually. However, from spring 2023, interviewers were provided with advance letters to send before making contact, following feedback from study members.

For the mop-up survey, study members were also sent an advance letter and e-mail, including a booklet/link to a booklet in advance of the survey. Those living abroad were only sent an e-mail invite. This correspondence was tailored to different respondent types (as 'emigrants' had not been invited to take part in the main stage, and pre-COVID responders, had already completed the survey). Study members were provided with a link to the survey and could also access a short video (via a link) to the Study Director explaining more about the importance of participating.

A thank you letter, and e-mail was sent to those who had completed the survey. This included confirmation of data linkage consents provided during the interview.

4.4.1 Tracing Strategy

Whenever interviewers identified that cohort members had moved from the issued address, they carried out tracing activities, which primarily consisted of making calls to partners and stable contacts or asking current occupiers of the cohort member's address if they had a forwarding address. For video call only waves, all tracing had to be conducted remotely (no visits could be made at addresses). Participants that interviewers were unable to trace were sent to CLS who attempted to find new contact details via office-tracing.

All tracing activities were completed before marking cases as non-contact.

4.4.2 Reminder Strategy

Once an interview was arranged an appointment letter was sent (or provided) by interviewers to the study member (this letter also contained any further documents needed for the interview which varied depending on whether it was a video call or in-person interview).

An appointment reminder was also sent the day before the interview by e-mail and text (depending on contact information available). When an appointment was made

for a video call interview, the interviewer would telephone the study member the day before the interview to confirm they had received the documentation needed and had all they needed to access the MS Teams call.

Reminders were also sent following the interview by e-mail and text to remind the study member to:

- ask their partner to complete the data linkage questionnaire (reminders to complete this were also sent directly to the partner)
- to complete the sensitive questions web survey (if not already done so during the interview – video only cases)
- to complete the Online Dietary Questionnaire
- to return the Paper Self Completion Questionnaire and Letter Cancellation document (if not collected by the interviewer)
- to remind them of the Health visit appointment.

For the mop-up survey – three electronic reminders were sent (by e-mail and text simultaneously) during the fieldwork period as well as one reminder letter.

Further detail on contact procedures can be found in the NCDS Age 62 Survey: Technical Report.

5. Questionnaire

5.1 Overview

The Age 62 Survey questionnaire contained a number of components including cognitive testing, a self-completion section consisting of sensitive questions and consent to data linkage.

Cohort members were also asked to complete two paper self-completion questionnaires which were either collected by the interviewer (at the end of the interview), the health professional (at the end of the Health Visit) or posted back.

The in-person and video interviews used the same questionnaire and Blaise 4 programme.

Slight adaptions were made to accommodate video interviewing, including:

- adaptions to interviewer instructions – including use of ‘show screen’ when required
- new section for video interviewers to record if and how the web survey containing sensitive questions would be administered e.g. independently during the interview, with assistance from the interviewer (showing their screen) or sent in an email to be completed after the interview
- instructions to use the letter cancellation document which was sent in advance.

The interview lasted around 95 minutes on average.

The mop-up web survey was a shortened version of the main stage questionnaire, taking an average of 25 minutes to complete and did not include any of the additional elements (cognitive assessments and consent to data linkage). A selection of the ‘sensitive’ questions were included in this survey.

5.2 Proxy, interpreter and partial interviews

The final data includes both partial and fully completed interviews. An interview was classed as partial if the respondent did not complete the interview but answered the last question in the family section. Partial interviews can be identified using the variables '**n10outcome'** (for the first or only interview) or '**n10outcome_r2'** (for those who took part in a second interview). 'Full' interviews are contingent on the completion of the core interview but do not take into account completion of the various 'Special elements' noted in section 5.4.

In cases where the cohort member was unable to understand survey questions or communicate the answers for themselves, a proxy interview could be conducted with a carer. This took them through a shorter route of the questionnaire lasting approximately 45 minutes. The survey was tailored to account for someone else participating on behalf of the cohort member, and did not include the sensitive questions self-completion module, cognitive testing or data linkage consent requests. The self-completion paper questionnaires were also not provided for completion.

If a respondent was able to participate themselves but just needed some assistance with communication (e.g. due to hearing or speech disability), an interpreter could be used, and the full interview would be conducted (unlike in a proxy case).

Proxy and interpreter interviews can be identified in the data using variable '**n10intwho**'.

5.3 Questionnaire Content

5.3.1 Main interview questionnaire content

Outlined below is a summary of the questionnaire content.

Table 2: Summary of Main Interview and Mop-Up questionnaire content

Questionnaire Module	Content	Included in the Mop Up Survey
Household grid	Cohabiting partnerships, children (including those not in the household) and any other household members.	This section was largely the same as the main interview but included a small additional section for emigrants about when and why they left Great Britain
Family	Non-cohabiting relationships, children, grandchildren, parents, and social contact.	No
Housing	Cohort member's housing history and current accommodation.	A short section collecting information on cohort member's current accommodation (housing history was not collected)
Employment	Cohort member's current and previous economic activity (including economic activity history, education or training, unemployment, retirement, sickness or disability and looking after the home or family) hours and pay and any changes in working practices due to the pandemic. Cohabiting partner's current economic activity income from employment, and any changes to working practices due to the pandemic.	Collected information on cohort member's current economic activity (including employment, education or training, unemployment, retirement, sickness or disability, and looking after the home or family) Cohabiting partner's current economic activity was also recorded. (Previous economic activity was not collected)
Income	Income from benefits, tax credits/allowances, pensions, investments and savings, inheritances and gifts received and debt for both the cohort member and their cohabiting partner.	Collected details on total income of the household from earnings, benefits, and any other form of earnings. One question replaced the detailed financial questions asked in the main stage
Cognitive Function	Four short memory, concentration and knowledge tasks (see section 5.4)	No
Lifelong Learning	Any new academic or vocational qualifications gained by the cohort member since the last interview (or 1st Jan 2008). Partner's highest qualification and age left education.	No

Questionnaire Module	Content	Included in the Mop Up Survey
Health	Physical and mental health, wellbeing, falls and broken bones, eyesight, private health insurance, exercise, drinking, smoking, self-reported height and weight	Asked general questions around physical and mental health and wellbeing and collected a self-reported weight measurement
COVID-19	Whether experienced symptoms of COVID-19 and long COVID, if they had had a positive test, and if they had been vaccinated.	Shortened version covering same core areas, excluding more detailed information such as how tested for COVID_19/when had positive test/type of vaccine received
Sensitive Questions	Political attitudes and voting, voluntary work, financial literacy, mental health, relationships with partner, relatives and friends, partner's health, gynaecological problems, menstruation, symptoms of the menopause and life satisfaction.	Included questions on mental health and life satisfaction only
Data linkage consent	Consent was sought to link survey data to government health and economic records where not obtained at the Age 50 survey. Permission from both the cohort member and their cohabiting partner was sought.	No
Contact Information	Updating contact details for the cohort member, partner and stable contacts. New contact details were also collected if the cohort member was planning to move. If the interview was carried out by proxy, the contact details of the person who acted as proxy were collected.	This section was largely the same as the section included in the main stage survey

5.3.2 Event Histories

There were three event histories included in the main interview: a relationship history, a housing history, and an economic activity history.

- Cohort members that had been interviewed in the last three sweeps (at either the age 46, 50 or 55 sweeps) were asked to update their cohabiting relationship history from the date of their last interview. Cohort members that had not been interviewed in the last three sweeps were asked to update their situation from 1st January 2004.

- Cohort members that had been interviewed in the last two sweeps (at either the age 50 or age 55 sweeps) were asked to update their housing or economic situation from the date of their last interview. Cohort members that had not been interviewed in the last two sweeps were asked to update their situation from 1st January 2008.

5.4 Special elements

The special elements noted below were only asked of those taking part in the main and pilot in-person and video interviews and were not part of the mop-up web survey.

5.4.1 Cognitive function tasks

Cohort members were asked to undertake four different cognitive assessments. The interviewer asked for consent prior to each assessment being administered.

The tasks were designed to measure different aspects of cognition and have been included in various other studies such as the 1970 British Cohort Study (BCS70) and the English Longitudinal Study of Ageing (ELSA) as well as in the NCDS Age 50 Survey.

Immediate word-list recall

This tested verbal learning and recall. Cohort members listened to a list of 10 words. They were then asked to recall the words immediately. In most cases, the list was read out by the computer using a recorded voice.

In video interviews the interviewer shared their screen with sound over Microsoft Teams so the cohort member could hear the recording of the 10 words they needed to recall.

In some cases, where the cohort member could not hear the recorded voice, the interviewer read out the list.

Variable name	Variable label
n10cflisen	Word recall - number *Number of words recalled (0-10)*

Animal naming

This tested how quickly cohort members could think of words from a particular category. Cohort members were asked to name as many different animals as they could think of in one minute. The timing was controlled by the computer. Interviewers recorded the number of animals the cohort member said, not counting any repetitions.

Variable name	Variable label
n10cfani	Number of animals named *Number named within one minute*

Letter cancellation

This tested attention, mental speed, and visual scanning. Cohort members were given a page of random letters of the alphabet arranged in a grid and were asked to cross out as many “P’s and “W’s as possible in one minute. They were then scored on both how accurately they completed the task, and how far along the grid they managed to get within one minute.

In video interviews, the letter cancellation sheet was placed in a sealed envelope and posted out before the interview. On the envelope were clear instructions not to open the envelope before being asked to by the interviewer during the interview. The cohort member then posted back the sheet.

Scoring of this assessment was conducted in the office, rather than by interviewers.

Variable name	Variable label
n10cfrc	Letter cancellation (Computed) – speed score (number of letters scanned)
n10cflot	Letter cancellation: (Computed) Total number CORRECT + MISSED by respondent
n10cfmis	Letter cancellation (Computed) – accuracy score (number of Ps and Ws missed)

Delayed word-list recall

This tested short term memory. Cohort members had 2 minutes to recall as many words as they could remember from the list they heard during the first word recall test. They were not permitted to listen to the list again. Interviewers made a note of

each word correctly recalled and entered the total number into the questionnaire programme.

Variable name	Variable label
n10cflisd	Delayed recall words: number *Number of words recalled (0-10)*

In the video pilot stage, the interval between the immediate and delayed word recall task was slightly shorter as this stage did not include the Letter Cancellation test prior to this cognitive test. Study members who opted out of taking part in the animal naming or letter cancellation test would also have a shorter time period between the immediate and delayed recall tests.

5.4.2 Sensitive questions

During the main interview, the cohort member was asked to complete a self-completion section which lasted for approximately 10 minutes and covered sensitive questions. This was administered differently in the in-person and in video interviews.

During in-person interviews the interviewer passed their laptop over to the respondent so that they could answer the sensitive questions themselves. The option was available for the interviewer to read out the questions should the cohort member request this.

In video interviews, the sensitive questions were programmed in a web survey. During the interview, the interviewer pasted the cohort member's unique link to the web survey in the chat function on MS Teams. The interviewer stayed on the video call while the cohort member completed the web survey.

If the cohort member could not access or complete the web survey themselves, the interviewer had the self-completion section on their interview programme and could share their screen so that the participant could read the questions and tell the interviewer the number of the response option they wished to choose. This option was only available after the soft launch.

The link to the web survey could also be emailed out following the survey if necessary. The web survey was programmed to be as similar as possible to the self-completion section used during in-person interviews to reduce any mode effects.

A ‘don’t know’ or ‘prefer not to say’ answer option was not available at first. It was made clear that the cohort member could skip questions they did not want to answer by leaving them blank. On doing so the options of ‘don’t know’ and ‘prefer not to say’ would then appear on their screen.

All variables asked as part this element have ‘SC’ at the beginning of the variable label. The responses completed as part of the web survey have been merged in the data with those completed in-person or with the interviewer. Variable ‘**nd10sc_mode’** provides further information on the mode of completion for this element.

5.4.3 Paper Self Completion questionnaires

‘Your Life Now’

Cohort members were asked to complete this paper self-completion questionnaire before their interview. The questionnaire was posted or given to the cohort member by the interviewer when making an appointment to conduct the interview. The questionnaire was either collected by the interviewer at the end of the interview (if an in-person interview) or posted back directly to the fieldwork agency. While it was mainly those taking part by video who posted back the questionnaire, this was an option in some cases where the questionnaire had not been completed in advance of an in-person visit. The paper self-completion questionnaire included the following sections:

- Activities
- Leisure time
- Activism
- Attitudes
- Neighbourhood
- Participation in organisations
- Religion
- Health and mental health

- Sleep
- Hearing and eyesight
- Teeth
- Alcohol consumption
- Risk taking/preferences
- Retirement
- Expectations for the future.

All variables in the data which are part of this paper self-completion have the variable labels prefixed with '(YLN PSC)'.

Mop-up respondents were not asked to complete this paper-self-completion questionnaire, though 29 mop-up cases did so and are included in the data. These cases would have received (and completed) the paper questionnaire when they were first invited to take part in the main survey (but did not take part at this stage and instead took part when issued to the 'mop-up').

Life History Questionnaire (LHQ) – also known as the 'Childhood Questionnaire'

Cohort members were initially provided this questionnaire at the end of their main interview to complete in advance of the Health Visit where it would be collected by the healthcare professional. Due to the impact of the pandemic and delays with the Health Visits, the protocol was changed so that this questionnaire was provided alongside the 'Your Life Now' self-completion questionnaire – to be completed in advance of the main interview (to be collected by the interviewer or returned by post).

This questionnaire asked cohort members to think back to their childhood and answer questions about their health, education, and family life as a child. The questionnaire asked about when cohort members were aged 7, 11, and 16, which is when the NCDS childhood surveys took place; as well as collecting a general overview of their childhood.

All variables in the data which are part of this paper self-completion have the variable labels prefixed with '(LHQ PSC)'.

As with the ‘Your Life Now’ questionnaire, mop-up respondents were not asked to complete this paper-self-completion questionnaire but 30 mop-up cases did so and are included in the data (following the receipt of this questionnaire when invited to take part at an earlier stage).

5.4.4 Data linkage consents

NCDS collects consent from cohort members and their cohabiting partners to link data collected in the study with records held by the National Health Service (NHS), His Majesty’s Revenue and Customs (HMRC), and the Department for Work and Pensions (DWP). The information contained in the health records provides details of hospital visits, long-lasting health conditions, treatments received, and medications prescribed. The economic records from DWP and HMRC includes details of benefits being received, national insurance and tax payments, and a full employment history.

Consent to data linkage was first asked about in the Age 50 survey. In the Age 62 Survey cohort members were asked to give consent if they had not taken part in the Age 50 survey or if they had taken part but refused one or more of the consents. If they had consented to link their data to some but not all of their records previously, they were only asked about those records they had refused.

Similarly, the cohort member’s partner was asked for consent if the cohort member had not taken part in the Age 50 Survey, they were a new partner or they were the same partner, but they had refused consent to one or more of their records before.

Prior to the pandemic, cohort members and their partners were provided with a paper consent form to record their consents. Due to the pandemic, changes were made so that consent to data linkage was only recorded electronically in the interview programme. Partners who were not present during the interview were sent a web survey asking them to provide their consents.

After the interview, cohort members and partners were sent an email or a letter which detailed what consents they had agreed to or refused. This correspondence also detailed the process to follow if they changed their minds.

A full description of the consent process and consent rates obtained is provided in the NCDS - Age 62 Survey: Technical Report.

For up-to-date information about the availability of linked data for research visit: www.cls.ucl.ac.uk/data-access-training/linked-data.

5.4.5 Occupation coding

Participants were asked to provide details about their current job, as well as any previous jobs they had since the last interview (or since 1st January 2008 if they had not taken part in the last two sweeps). Those with a cohabiting partner were also asked to provide details about their partner's job.

All occupations were coded to the four-digit standard occupation coding frame (SOC 2020). The SOC2020 codes are included in the data. To minimise disclosure risk, 3-digit SOC codes are included in the safeguarded dataset (EUL). The 4-digit SOC codes are available as controlled data under Secure Access (see Section 8.1.2).

5.4.6 Online dietary questionnaire

Cohort members were asked during their Health Visit if they would be willing to complete an online dietary questionnaire about two randomly allocated days from the seven-day period following their health visit. They were provided with a leaflet containing a link to the questionnaire and a unique login code.

The questionnaire (The Oxford WebQ © the University of Oxford (2009)) was originally developed by Professor Dame Valerie Beral, Professor Tim Key, Dr Bette Liu, Vicky Benson, Dr Elizabeth Spencer and Heather Young, University of Oxford.

The Oxford WebQ was developed for repeated implementation in large prospective studies, e.g., the UK Biobank and the Million Women Study. It asks about consumption of about 200 commonly consumed food and beverage items during the previous 24 hours. The quantity of each food or drink consumed during the reference period is calculated by multiplying the assigned portion of each food or beverage by the amount consumed. The nutrient intake is calculated by multiplying the quantity consumed by the nutrient composition of the food or beverage (using McCance and Widdowson nutrient database). For further information please visit the UK Biobank Showcase pages for the Oxford WebQ or www.ceu.ox.ac.uk/research/oxford-webq.

Data from the OxfordWebQ is not included in this first version of the data. The data will be made available later in 2025.

5.5 Scales

The NCDS Age 62 Survey included several established scales which are listed below. Overall scores for each scale have been derived and included within the data dataset **ncds10_age62_main_interview.** Further details regarding the derivation of the scores can be found in Appendix 1 – 'Derived Variables'.

5.5.1 Health module: ONS long lasting health conditions and illnesses: Impairments and Disability (ONS, 2015)

The Age 62 Survey included a sub-set of the ONS harmonised set of questions on Long-lasting Health Conditions and Illnesses including Impairments and Disability. The three items listed below are used to derive variables indicating whether cohort members are disabled using the Equality Act 2010 definition and whether they have a long-standing illness or condition using the European Union's Statistics on Income and Living Conditions (EU-SILC) definition (ONS, 2015). Nd10diseq identifies individuals as disabled or not, **nd10disls** identifies individuals as having no long-standing health condition, having a condition which hampers daily activities to an extent or having a condition which severely hampers daily activities.

Variable name	**Variable label**
n10loil/ n10loil_r2	Whether has longstanding illness/conditions
n10lolm/ n10lolm_r2	Whether illness/conditions reduce ability to carry out day to day activities
n10lolp/ n10lolp_r2	Length of time ability to carry out day to day activities has been reduced
nd10diseq/ n10diseq_r2	(Derived) Disability classification Equality act (2010)
nd10disls/ nd10disls_r2	(Derived) Disability classification EU-SILC

According to the Equality Act 2010 definition, a cohort member is considered to be disabled if they report a longstanding illness (**n10loil/ n10loil_r2**) and have a

reduced ability to carry out day-to-day activities as a result of their illness (**n10lolm/ n10lolm_r2**).

According to the EU-SILC definition, a cohort member is considered to be disabled if they report a longstanding illness (**n10loil/ n10loil_r2**)), have a reduced ability to carry out day to-day activities as a result of their illness (**n10lolm/ n10lolm_r2**), and this reduced ability has lasted for more than 6 months (**n10lolp/ n10lolp_r2**). This variable also distinguishes between those that are disabled to some extent, and those that are severely hampered (**from n10lolm/ n10lolm_r2**).

5.5.2 Self-completion module: Social provisions

Cutrona CE, Russell DW. The provisions of social support and adaptation to stress. Advance in Personal Relationships. 1987;1:37–67

Three items were included from the 10-item Social Provisions Scale (Cutrona 1987). These were included in the video pilot and re-start main stages only and were part of the section covering 'sensitive questions'.

The Social Provisions Scale measures the availability of social support. Cohort members were asked to think about their current relationships with friends, family members, community members and so on. They were asked to indicate the extent to which each statement: a) I have family and friends who help me feel safe, secure and happy b) There is someone I trust whom I would turn to for advice if I were having problems c) There is no one I feel close to…. described their current relationship with other people from the following responses:

1. Very true
2. Partly true
3. Not true at all.

Variable Name	Variable label
n10socprova	(SC) Have family and friends who help feel safe, secure and happy
n10socprovb	(SC) Have someone to turn to for advice
n10socprovc	(SC) No one close to

These questions were asked as part of the sensitive questions section in the soft launch and main stages only (they were not included in the pilots and mop-up).

5.5.3 Self-completion module: GAD2 (Generalised Anxiety Disorder 2-item)

Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med. 2007;146:317-25.

The GAD-2 was based on the GAD-7, which was developed by Drs. Robert L. Spitzer, Janet B.W. Williams, Kurt Kroenke and colleagues, with an educational grant from Pfizer Inc. No permission was required to reproduce, translate, display or distribute. The Generalized Anxiety Disorder 2-item (GAD-2) is a brief initial screening tool for generalized anxiety disorder.

Respondents are asked how often they have been bothered by problems over the last 2 weeks: a) “Feeling nervous, anxious or on edge”; and b) “Not being able to stop or control worrying”, with the following response options:

1. Not at all
2. Several days
3. More than half the days
4. Nearly every day.

The GAD-2 score is obtained by adding the score for each question (Total points). The score for each question is:

0 = Not at all

1 = Several days

2 = More than half the days

3 = Nearly every day

Variable name	Variable label
n10gad2phq2a/ n10gad2phq2a_r2	(SC) Whether nervous, anxious or on edge over the last 2 weeks
n10gad2phq2b/ n10gad2phq2b_r2	(SC) Whether not able to stop or control worrying in the last 2 weeks
nd1011gad/ nd1011gad2_r2	(Derived) Generalised Anxiety Disorder 2-item

These questions were asked as part of the sensitive questions section in the video pilot and re-start main stage as well as being included in the mop-up.

5.5.4 Self-completion module: PHQ2 (Patient Health Questionnaire 2-item)

Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: Validity of a Two-Item Depression Screener. Medical Care. 2003;41:1284-92.

The PHQ-2 enquires about the frequency of depressed mood and anhedonia over the past two weeks. The PHQ-2 includes the first two items of the PHQ-9.

Respondents are asked how often they have been bothered by problems over the last 2 weeks: c) “Little interest or pleasure in doing things”; and d) “feeling down, depressed or hopeless”, with the following response options:

1. Not at all
2. Several days
3. More than half the days
4. Nearly every day.

The PHQ-2 score is obtained by adding the score for each question (Total points). The score for each question is:

0 = Not at all

1 = Several days

2 = More than half the days

3 = Nearly every day

Variable name	Variable label
n10gad2phq2c/ n10gad2phq2c_r2	(SC) Whether had little interest or pleasure in doing things in the last 2 weeks
n10gad2phq2d n10gad2phq2d_r2	(SC) Whether feeling down, depressed or hopeless in the last 2 weeks
nd10phq2 nd10phq2_r2	(Derived) Patient Health Questionnaire 2-item

These questions were asked as part of the sensitive questions section in the video pilot, re-start main stage and mop-up.

5.5.5 Self-completion module: UCLA loneliness 3 item

Daniel W. Russell (1996) UCLA Loneliness Scale (Version 3): Reliability, Validity, and Factor Structure, Journal of Personality Assessment, 66:1, 20-40, DOI: 10.1207/s15327752jpa6601_2

Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A Short Scale for Measuring Loneliness in Large Surveys: Results From Two Population-Based Studies. Res Aging. 2004;26(6):655-672. doi: 10.1177/0164027504268574. PMID: 18504506; PMCID: PMC2394670.

Three items from the 20-item UCLA loneliness scale were asked of all cohort members in the in-person pilot, main stage and mop-up stages. These questions were not included in the video pilot. They were asked to give the frequency in response to questions about current loneliness and related emotional states from the following response options:

1. Hardly ever
2. Some of the time
3. Often.

A combined score is obtained by adding the score for each question (Total points). The score for each question is:

1 = Hardly ever

2 = Some of the time

3 = Often

In addition, a fourth item (How often do you feel lonely?) was included in the NCDS paper questionnaire and as part of the mop-up survey. This is not part of the UCLA scale.

Variable name	Variable label
n10lonelya/ n10lonelya_r2	(SC) Feeling lack of companionship
n10lonelyb n10lonelyb_r2	(SC) Feeling left out
n10lonelyc n10lonelyc_r2	(SC) Feeling isolated from others
nd10loneliness nd10loneliness_r2	(Derived) UCLA loneliness 3 item

5.5.6 Self-completion module – MALAISE Inventory

Rutter, M., Tizard, J., & Whitmore, K. (1970). Education, health, and behaviour. London: Longman.McGee, R., Williams, S., and Silva, P. A. (1986) 'An evaluation of the Malaise Inventory', Journal of Psychosomatic Research, 30(2), pp.147-152.

Earlier sweeps of the study have included a set of 24 self-completion questions which combine to measure levels of psychological distress, or depression (Rutter et al, 1970). As per the Age 50 Survey, the Age 62 Survey used 9 of the original 24 items.

Variable name	Variable label
N10mal02/ N10mal02_r2	(SC) Whether feel tired
N10mal03 N10mal03_r2	(SC) Whether feel depressed
N10mal05 N10mal05_r2	(SC) Whether worried
N10mal09 N10mal09_r2	(SC) Whether gets enraged
n10mal12 n10mal12_r2	(SC) Whether gets scared
n10mal14 n10mal14_r2	(SC) Whether gets easily upset
n10mal16 n10mal16_r2	(SC) Nervousness
n10mal20 n10mal20_r2	(SC) Whether annoyed and worn out
n10mal21 n10mal21_r2	(SC) Whether heart races
nd10mal nd10mal_r2	(Derived) Total Malaise score (9 questions)

Variable name	Variable label
nd10malg nd10malg_r2	(Derived) Total Malaise score – grouped

These questions were asked as part of the sensitive questions section in all stages.

5.5.7 Paper Self Completion: – Mental Health Inventory (MHI-5)

Ware J.E., Snow K.K., Kosinsk,i M., & Gandek, B. (1993), SF-36 Health Survey Manual and Interpretation Guide. Boston, MA: New England Medical Center, The Health Institute.

The NCDS Age 62 Survey, has used a shortened version of SF_36 (Ware et al, 1993) – The Mental Health Inventory (MHI-5) which is often used as a screener for mood and anxiety disorders. This scale consists of five items – outlined below.

Variable name	Variable label
n10ylq16a	(YLN PSC) During the past 4 weeks whether been -a very nervous person
n10ylq16b	(YLN PSC) During the past 4 weeks whether felt that nothing could cheer you up
n10ylq16c	(YLN PSC) During the past 4 weeks felt calm and cheerful
n10ylq16d	(YLN PSC) During the past 4 weeks felt downhearted and low
n10ylq16e	(YLN PSC) During the past 4 weeks has been-a happy person
nd10emwb	(Derived) SF-36 Emotional Well-Being score

5.5.8 Paper Self Completion - AUDIT-PC

F. Babor, T., C. Higgins-Biddle, J., B. Saunders, J., & G. Monteiro, M. (2001). The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care. W. H. Organisation.

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1113176/Alcohol-use-disorders-identification-test-for-primary-care-AUDIT-PC_for-print.pdf

The AUDIT-PC consists of 5 questions covering alcohol consumption, problems and dependency. Responses to each question are scored from 0 to 4 giving a maximum score of 20 (nd10audit). Scores of 5 or more are considered AUDIT-PC positive and associated with increasing or higher risk drinking (nd10audg). It is an abbreviated version of the full AUDIT scale which was included in the Age 50 NCDS Survey.

Variable Name	Variable label
n10ylq28	(YLN PSC) How often CM has drink containing alcohol
n10ylq29	(YLN PSC) How many drinks containing alcohol has on a typical day when drinking
n10ylq30a	(YLN PSC) How often during the last year not able to stop drinking once had started
n10ylq30b	(YLN PSC) How often during the last year failed to do what was normally expected because of own drinking
n10ylq31	(YLN PSC) Whether a relative or friend or health worker has been concerned about CMs drinking
nd10audit	(Derived) Total AUDIT-PC Score
nd10audg	(Derived) AUDIT-PC Group

5.5.9 Paper Self Completion: Shortened version of Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS)

Michael T. McKay, James R. Andretta, Evidence for the Psychometric Validity, Internal Consistency and Measurement Invariance of Warwick Edinburgh Mental Well-being Scale Scores in Scottish and Irish Adolescents, Psychiatry Research, Volume 255, 2017,Pages 382-386, ISSN 0165-1781, https://doi.org/10.1016/j.psychres.2017.06.071.

Hanzlová, R. and Lynn, P., 2023. Item response theory-based psychometric analysis of the Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS) among adolescents in the UK. Health and Quality of Life Outcomes, 21(1), p.108.

Ringdal, R., Eilertsen, M.-E., Bjørnsen, H., Espnes, G., & Moksnes, U. (2018). Validation of two versions of the Warwick-Edinburgh Mental Well-Being Scale among Norwegian adolescents. Scandinavian Journal of Public Health, 46, 140349481773539. https://doi.org/10.1177/1403494817735391

Ng Fat, L., Scholes, S., Boniface, S., Mindell J., & Stewart-Brown S. (2017) Evaluating and establishing the national norms for mental well-being using the short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS): findings from the Health Survey for England. Quality of Life Research, 26(5), 1129-1144.

SWEMWBS is a short version of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS). SWEMWBS uses 7 of the WEMWBS's 14 statements about thoughts and feelings, which relate more to functioning than feelings therefore offering a slightly different perspective on mental wellbeing.

The seven positively worded items with five response categories are outlined below.

The SWEMWBS is scored by first summing the scores for each of the seven items, which are scored from 1 to 5. The total raw scores are then transformed into metric

scores using the SWEMWBS conversion table which can be found here: www.warwick.ac.uk/fac/sci/med/research/platform/wemwbs/using/howto/swemwbs_raw_score_to_metric_score_conversion_table.pdf

Scores range between 7 to 35 and higher scores indicate higher levels of mental well-being.

Rating scale:

1. None of the time
2. Rarely
3. Some of the time
4. Often
5. All of the time.

Statements:

1. I've been feeling optimistic about the future.
2. I've been feeling useful.
3. I've been feeling relaxed.
4. I've been dealing with problems well.
5. I've been thinking clearly.
6. I've been feeling close to other people.
7. I've been able to make up my own mind about things.

Variable Name	**Variable label**
n10ylq18a	(YLN PSC) In the last 2 weeks felt optimistic about the future
n10ylq18b	(YLN PSC) Over the last 2 weeks whether felt useful
n10ylq18c	(YLN PSC) Over the last 2 weeks whether felt relaxed
n10ylq18d	(YLN PSC) Over the last 2 weeks whether dealt with problems well
n10ylq18e	(YLN PSC) Over the last 2 weeks whether been thinking clearly
n10ylq18f	(YLN PSC) Over the last 2 weeks whether felt close to other people
n10ylq18g	(YLN PSC) Over the last 2 weeks whether felt able to make up your own mind about things
n10wemwb	(Derived) Warwick Edinburgh Mental Well-Being Scale

5.5.10 Paper Self Completion: Quality of Life - CASP6

Wiggins, R.D., Brown, M., Ploubidis G.B. (2017). A measurement evaluation of a six item measure of quality of life (CASP6) across different modes of data collection in the 1958 National Child Development Survey (NCDS) Age 55 years: Working Paper 2017/2: CLS-WP-20172.pdf

Wiggins, R. et al (2004) Quality of life in the third age: key predictors of the CASP-19 measure, Ageing & Society, 24, pp. 693–708.

Wiggins, R. et al (2008) 'The Evaluation of a Self-enumerated Scale of Quality of Life (CASP-19) in the Context of Research on Ageing: A Combination of Exploratory and Confirmatory Approaches', Social Indicators Research, 89, pp 61-77.

The original version of CASP contained nineteen items spanning four life domains (Hyde et al., 2003) – control, autonomy, self-realisation and pleasure, used to measure the quality of life. Due to space constraints, further research has been conducted to shorten the scale and both CASP-12 and CASP-14 were available in the NCDS Age 50 Survey (Wiggens et al, 2004). In both the Age 55 survey and this recent Age 62 survey, it was only possible to include six items from CASP-12, covering two of the domains – control and self-realisation (Wiggens et al, 2017).

Variable Name	Variable label
n10ylq19a	How often feels age prevents me from doing things
n10ylq19b	How often feels what happens to me is out of my control
n10ylq19c	How often feels left out of things
n10ylq19d	How often feels full of energy these days
n10ylq19e	How often feels that life is full of opportunities
n10ylq19f	How often feels that the future looks good to me
n10casp6	(Derived) CASP Quality of Life score 6-item

6. Biomeasures

The biomeasures were obtained in a follow-up Health Visit (also known as a nurse visit) conducted by a healthcare professional, that lasted around one hour. The health visit involved the following:

- drug coding of prescribed medications
- blood pressure (sitting and standing)
- grip strength
- waist and hip circumference
- weight and body fat percentage
- blood sampling
- timed walk
- balance measurements.

A summary of each measurement and associated key variables is provided below. Full protocols for all measurements can be found in NCDS – Age 62 Survey: Nurse Protocols.

Data from this visit is contained in: **ncds10_age62_biomeasures** (unless otherwise specified below).

6.1 Drug/medication coding

Participants were asked '*Are you taking or using any medicines, pills, syrups, ointments, puffers or injections prescribed for you by a doctor or nurse?'* and also about any long-acting medications (such as an injection or implant). The name of each medication was recorded and where possible health professionals asked to see the medication packaging to increase accuracy. The health professionals then coded each medication to sub-chapter level (4 digits) of the British National Formulary (BNF) edition 69. The coding of the medications could be completed during the blood pressure module (when participants are resting) or at the end of the health visit.

The data for medications is contained in the following files:

- **ncds10_age62_medicin_longf** (medications)
- **ncds10_age62_long_acting_medicin_longf** (long-acting medications).

Variable name	Variable label
n10drc6_tr*	BNF Drug Code
n10drc4_tr*	BNF Drug code subchapter
n10brc2_tr*	BNF Drug code chapter
n10drc_codes_tr*	BNF Drug code missing values

Variable name	Variable label
n10mlndrc6_tr*	Long Acting BNF Drug Code
n10mlndrc4_tr*	Long Acting BNF Drug code subchapter
n10mlnbrc2_tr*	Long Acting BNF Drug code chapter
n10mlndrc_codes_tr*	Long Acting BNF Drug code missing values

6.2 Blood pressure

Both seated and standing blood pressure measurements were taken using an Omron HEM 907 blood pressure monitor. Two blood pressure measurements were taken while the participant was seated (one minute apart), immediately after the participant is asked to stand for one minute and a measurement is taken while they are standing.

The health professional fits a cuff of the appropriate size (standard or large) to the respondent's left arm (where possible).

Variable name	Variable label
n10sys	first systolic reading (mmhg)
n10dias	first diastolic reading (mmhg)
n10pulse	first pulse reading (bpm)
n10sys2	second systolic reading (mmhg)
n10dias2	second diastolic reading (mmhg)
n10pulse2	second pulse reading (bpm)
n10standsys	standing systolic reading (mmhg)
n10standdias	standing diastolic reading (mmhg)
n10standpulse	standing pulse reading (bpm)

6.3 Grip Strength

Grip strength is measured in kilograms using a Jamar Plus digital dynamometer. The measurement is taken while the participant is seated. Respondents must squeeze the dynamometer as hard as they can. Two measurements are taken in each hand, starting with the left hand (regardless of dominant hand) and alternating between left and right.

Variable name	Variable label
n10mmgsl1	left hand, first measurement
n10mmgsr1	right hand, first measurement
n10mmgsl2	left hand, second measurement
n10mmgsr2	right hand, second measurement
n10mmgslav	average grip strength – left hand
n10mmgsrav	average grip strength – right hand
n10mmgsres	(Derived) Whether all grip measures obtained

6.4 Anthropometry

The anthropometric measurements included are:

- weight and body-fat percentage
- waist and hip circumference.

Height was not measured in this NCDS survey. Instead, the healthcare professional was provided information on height based on the measurement recorded in Age 44 biomedical survey. If this information was not available, height was then taken from the self-reported height recorded by the cohort member during their main interview. If height was not available from either of these sources or the healthcare professional recorded that the height provided was not accurate, they would be asked to enter their estimate of the cohort member's height.

Weight was measured in kilograms using Tanita BF - 522W scales. The scales can accurately measure up to 130 kilograms and those whose weight likely exceeded this were not weighed.

Body fat percentage is the total weight of the person's fat divided by the person's weight and is measured using the same scales by sending a weak electrical current

around the body from one foot to the other. Respondent's age, gender and height is entered into the scales prior to measurement, to ensure the accuracy of the body fat measurements.

Body Mass Index = (body mass divided by the square of body height)

Waist and hip measurements were taken using a EasyCheck (extra long 200cm) retractable tape measure. Measurements are taken to the nearest even millimetre.

One measure of waist circumference is taken followed by one measurement of hip circumference and recorded in the CAPI.

Variable name	Variable label
n10bfpc	Body fat percentage measurement
n10mweight	Weight (if body fat to be measured as well as weight)
n10wgtonly	Weight (no body fat measurement)
n10waist	Waist measurement
n10hip	Hip measurement
nd10hghtm	(Derived) Self-reported height in metres
nd10nhghtm	(Derived) Nurse height in metres
nd10wghtk	(Derived) Self-reported weight in kilograms
nd10bmi nd10bmi_r2	(Derived) Body mass index (based on self-reported data)
nd10bmic nd10bmic_r2	(Derived) Body mass index - classification (based on self-reported data)
nd10nbmi	(Derived) Body mass index (based on nurse measured data)
nd10nbmic	(Derived) Body mass index - classification (based on nurse measured data)

6.5 Collection of blood samples

If cohort members consented to all blood sample measurements, a maximum of five tubes of blood were taken. Participants were not required to be in a fasted state as health professionals made visits at any time in the day.

Two tubes were centrifuged in the respondent's home. Two tubes were sent to the Newcastle RVI lab for analysis of cholesterol (total and HDL), glycated haemoglobin, triglycerides and c-reactive protein (CRP). The remaining tubes were sent to

University of Bristol lab to be stored for future analyses including DNA extraction (where consent provided).

On completion of fieldwork additional analyses were conducted using samples stored by Bristol. Assays of Troponin I, NT-proBNP and GDF-15 were conducted by the BHF Glasgow Cardiovascular Research Centre at University of Glasgow.

Metabolomics analysis was conducted by Nightingale and further information on this and how to access the data can be found in the user guide: https://cls.ucl.ac.uk/wp-content/uploads/2025/07/NCDS-Metabolomics-User-Guide.pdf.

Table 3: Blood sample collection

Blood quantity and type of tube	Number of inversions[2]	Centrifuged	Lab	Analytes
5 ml / Rapid Serum (RST)	6	Yes	Bristol	Aliquoting and storage for future use / Metabolomics (Nightingale)
2.5 ml / SST	6	No	RVI	Total/HDL cholesterol, Triglycerides, C-Reactive-Protein
3 ml / K2 EDTA	10	No	RVI	HbA1c
6 ml / K2 EDTA	10	No	Bristol	DNA, Troponin I, NT-proBNP and GDF-15 (Glasgow University lab), other analytes
5 ml / PPT EDTA	10	Yes	Bristol	Aliquoting and storage for future use.

6.5.1 Blood Assays

Newcastle RVI laboratory conducted the following assays: Total cholesterol, HDL cholesterol, Triglycerides, High-sensitivity C-reactive protein and Glycated haemoglobin (HbA1c).

Only participants who provided signed consent for blood sample analysis and received by the lab were included. The laboratory returned results for 5125 samples, and 5117 results are retained after exclusions due to participant consent withdrawal.

[2] Inversions refers to the turning of the tube (by healthcare professional) upside down and back upright to mix the contents

Table 4: Methods used to conduct blood assays at Newcastle RVI

Code	Full name	Method principle	Manufacturer	Analytical limits
NCHOL	Total cholesterol	Enzymatic colorimetric: cholesterol esterase/cholesterol oxidase/peroxidase	Roche Cobas c702, generation 2 assay	0.1 – 20.7 (207 on dilution)
HDL	HDL cholesterol	Enzymatic colorimetric: cholesterol esterase/cholesterol oxidase/peroxidase	Roche Cobas c702, generation 4 assay	0.1-3.88
TRIG	Triglycerides	Enzymatic colorimetric: lipoprotein lipase/glycerol kinase/glycerol phosphate oxidase/peroxidase	Roche Cobas c702	0.1-10 (50 on dilution)
HSCRP	High-sensitivity C-reactive protein	Particle-enhanced immunoturbidimetry	Roche Cobas c702	0.6-350 (700 on dilution) 0.15-20.0
HBA1C	Glycated haemoglobin	Ion exchange HPLC	Tosoh G8 - Moved to Tosho G11 in January 24	Not quoted by manufacturer

6.5.2 Cardiovascular health markers

Three cardiovascular (CV) health markers - Troponin-I, NT-proBNP, and GDF-15 - were analyzed at the BHF Glasgow Cardiovascular Research Centre, University of Glasgow. Results for 4939 samples were returned, and 4931 samples were retained in the final dataset after exclusions due to participant consent withdrawal.

QC samples from the manufacturer were run during the assay, but no QC checks for individual samples as these tests are not as sensitive to moderate haemolysis. For this group of biomarkers, no outliers were found across the batch of samples.

Altogether, there are 5141 samples in the final dataset of combined panel of blood assays and cardiovascular health markers.

Table 5: Methods used to conduct blood assays at BHF Glasgow

Code	Full name	Method principle	Manufacturer
GDF15	Growth differentiation factor-15	Electrochemiluminescence immunoassay (ECLIA)	Roche GDF-15, cobas e801 analyser
NTproBNP	N-terminal prohormone of B-type natriuretic peptide		Roche proBNP II, cobas e801 analyser
hsTnI	High sensitivity Troponin-I	Chemiluminescent microparticle immunoassay (CMIA)	AbbotT hsTnI, Architect i1000 analyser

Table 6: Manufacturer limits of detection, limit of quantitation, and coefficient of variation for each assay

Code	Manufacturer limits of detection	%CV	Manufacturer Limit of Quantitation (CV >20% below this point)	Disease Thresholds
GDF15	400-20000 pg/mL	Low: 2.78 Med: N/A High: 2.81	400 pg/mL	In the Generation Scotland cohort, median GDF-15 in healthy and non-pregnant individuals was 808 pg/mL (IQR 608, 1,103 pg/mL), PMID 35976089. GDF-15 correlated positively with NT-proBNP and Troponin I.
NTproBNP	5-35000 pg/mL	Low: 3.55 Med: N/A Hi: 2.81	50 pg/mL	NT-proBNP of ≥ 125 pg/mL is corroborating evidence of heart failure with symptoms. In non-acute settings 125 pg/mL is the rule out threshold. UK NICE guidelines indicate NT-proBNP 400 pg/mL as rule out threshold for chronic heart failure.
hsTnI	1.1-50000 pg/mL	Low: 9.7 Med: 8.8 Hi: 9.4	1.3 pg/mL	NICE guidelines indicate the rule out threshold for acute MI at a hs cardiac troponin I of 15.6 pg/mL for women and 34.2 pg/mL for men.

Order	Variable name	Variable label
1	ncdsid	NCDS Research ID
2	n10gdf15	Growth differentiation factor-15, LOD: 400 (pg/mL)
3	n10ntprobnp	N-terminal prohormone of B-type natriuretic peptide, LOD: 5 (pg/mL)
4	n10hstni	High sensitivity Troponin-I, LOD: 1.2 (pg/mL)
5	n10chol	Total cholesterol, LOD: 0.1 (mmol/L)
6	n10cholq	QC flag: Total cholesterol
7	n10hdl	HDL cholesterol, LOD: 0.1 (mmol/L)
8	n10hdlq	QC flag: HDL cholesterol
9	n10trig	Triglycerides, LOD: 0.1 (mmol/L)
10	n10trigq	QC flag: Triglyceride
11	n10hscrp	High-sensitivity C-reactive protein, LOD: 0.6 (mg/L)
12	n10hscrpq	QC flag: High-sensitivity C-reactive protein
13	n10hba1c	Glycated hemoglobin, LOD: unknown (mmol/mol)
14	n10hba1cq	QC flag: Glycated hemoglobin

6.6 Timed Walk

Cohort members were asked to walk a distance of 8 feet/244 cm and the time taken to do this was recorded (this exercise was completed twice). The purpose of this measurement is to objectively measure the overall health and level of disability in NCDS respondents. Walking speeds of those aged 60 and over have been known to predict: levels of disability; future use of health care and mortality.

Variable name	Variable label
N10MmTrya	Outcome of first timed walk
N10MMWlkA	Time of first walk in seconds
N10MmTryb	Outcome of second timed walk
N10MMWlkB	Time of second walk in seconds

6.7 Balance measurements

NCDS respondents balance was measured through the ‘leg raise protocol’. This involved asking them to stand on one foot with the other foot raised a few inches off the ground, with their arms folded across their chest. They were asked to hold this position, for as long as possible, up to 30 seconds. Timing would stop when this balance was lost (such as their feet moving position or their arms moving away from

their chest). The test was conducted with eyes open and then again with their eyes closed.

Variable name	Variable label
N10MMLORE	Outcome of eyes open leg raise
N10MMLOTI	Time of eyes open leg raise
N10MMLSRE	Outcome of eyes closed leg raise
N10MMLSTI	Time of eyes closed leg raise

7. Response

7.1 Overall response

Overall, 11,493 cohort members were initially invited to take part in this sweep of NCDS. An additional 327 cases known to be living outside of Great Britian (emigrant cases) were invited to take part in the mop-up web, boosting the total issued sample to 11,820. Emigrant cases were not considered as 'eligible' to take part in the main stage or pilots which were conducted in-person or by video as these cases are not formally considered part of the target population.

A total of 8,405 study members participated in either the pilots, dress rehearsal, main stage or mop-up web survey (including partial interviews, proxy cases and emigrants).

Due to the addition of 'emigrant' cases, which were not part of the initial 'target' population, two separate response rates are outlined below:

Response rate A excludes the 'emigrant' cases from interviews achieved and issued sample, as these cases were deemed as ineligible for the main and pilot stages.

Response rate B includes these 'emigrant' cases as eligible in both achieved interviews and issued sample.

A total of 285 confirmed ineligibles cases have been removed from both response rates (for example those found to have died/were in prison).

- Response rate A (emigrants ineligible) – 8,215 productive interviews and a response rate of 73%[3].

[3] Response = 100*(8215 productive interviews/(11493 original cases issued minus 285 confirmed ineligibles))

- Response rate B (emigrants eligible) – 8,405 productive interviews and a response rate of 73%[4].

Out of the 8,405 productive interviews, 26 were partially productive in-person/video interviews and 31 were partially completed mop-up web surveys (2 of these cases were 'emigrant' cases).

Eleven cases took part via proxy and six with the aid of an 'interpreter' (see section 5.2 for further information).

7.2 Response by fieldwork stage

7.2.1 Main stage and pilots

A total of 7,382 cohort members were interviewed during main stage fieldwork, between January 2020 and November 2023, a response of 69%. An additional 488 interviews were achieved during the pilots and dress rehearsal phases boosting the total number of interviews achieved (prior to the mop-up) to 7,809, an overall response of 70%.

7.2.2 Mop-Up Survey Non Responders (excluding emigrants)

In the web-based Mop-Up Survey which took place from March to April 2024, a further 2,484 cohort members who had been invited to the main stage pilots and dress rehearsal but had not participated were invited to take part and 406 interviews were achieved, giving a total of 8,215 interviews. The mop-up survey increased the response rate to 73%. Table 7 provides a breakdown of response rate A.

[4] Response=100*(8405 productive interviews including emigrants/(11,493 original cases issued plus 327 emigrant cases issued minus 285 confirmed ineligibles))

Table 7: Overview of response rates from pilots, main stage survey and mop-up (excluding emigrants)

	N	*%*
Productive	8,215	71
by video interview	*2,281*	*20*
by in-person interview	*5,528*	*48*
by web in mop-up	*406*	*4*
Non-contact	417	4
Refusal	1829	16
Other unproductive	348	3
Unknown eligibility (no contact)	399	3
Ineligible[5]	285	3
Total	*11,493*	*100*
Response rate[6]		***73***

7.2.3 Mop-Up Survey Pre-COVID Responders (excluding emigrants)

As outlined in section 4.3.8, participants who took part in the main survey, prior to the COVID lockdown (between January and March 2020), were also invited to take part in the mop-up survey. For these cases we have data collected at two time points – before and after the COVID_19 outbreak.

1,603 pre-COVID responders were invited to take part and 1,046 surveys were completed giving a response rate of 65%.

7.2.4 Mop-Up Survey (including emigrants)

A total of 327 study members living outside of Great Britian were invited to take part in the mop-up survey and 190 participated (response 58%). Table 8 below provides further information on the response to the mop-up survey.

[5] This 'ineligible' code includes 235 cases identified during fieldwork as 'deceased'.
[6] The response rate is the percentage of productive interviews from the issued sample, excluding those confirmed as ineligible.

Table 8: Mop-up Survey response

	Sample size	Productive completes		Productive partials		All productives	
		N	%	N	%	N	%
Non-responders to main survey	2,484	377	15	29	1	406	16
Emigrants	327	188	57	2	1	190	58
Pre-COVID responders	1,603	1,017	63	29	2	1,046	65
Total	4,414	1,582	36	60	1	1,642	37
Base: All cases issued to mop-up survey (4,414)							

The tables 9 provides an overview of each stage and the response achieved.

Table 9: Survey response at each stage of fieldwork

	Main stage (initial and re-launch)	Main stage with pilots/DR	Main stage/ pilots mop-up cases (excluding emigrants)	Main stage/ pilots mop-up cases (including emigrants)	Main stage and mop-up – 2 time points
Number of completed interviews (N)	7,382	7,809	8,215	8,405	1,046
Response rate (%)	66	70	73	73	65
Base: All known eligible cases issued to main stage and mop-up survey (base 11,208 – excludes 'emigrant' cases and ineligibles, base 11,535 includes 'emigrant' cases, excludes ineligibles)					

7.3 Mode of Completion

Of the 8,405 total number of interviews achieved in the pilot, video pilot, main stage and mop-up survey (including emigrants), 66% were completed in-person, 27% were completed by video and7% were completed by web. The Age 62 sweep of NCDS utilised a data collection design similar to that employed in the 1970 British Cohort Study (BCS70) Age 51 Survey, which involved a mixed mode approach with the majority of interviews conducted either face-to-face or via video call. Findings from BCS70 indicate that differences observed between in persons vs video interviews are largely due to selection effects—that is, differences in who chooses each mode—rather than mode effects arising from measurement error. This suggests that mode-related measurement bias is minimal, assuming no NCDS specific mode

effects. Empirical evidence for the NCDS Age 62 sweep will be provided in the second version of this User Guide to probe this assumption. In the meantime, consistent with BCS70 results, it is reasonable to conclude that mode differences in the NCDS data are primarily driven by respondent selection rather than the mode of data collection itself, meaning that data can be analysed across modes. Mode effects may be present when comparing with the online mop-up questionnaire (though few individuals responded by this mode, which limits the extent to which web completion related mode effects may bias analyses). If mode effects are suspected, it is advisable to follow the guidance in the Handling Survey Mode in CLS Cohorts User Guide (Wright, 2024).

Survey mode is denoted by '**n10survey_mode'** variable in the survey dataset.

Table 10: Mode of response

Mode	Frequency	Percent
In-person interviews	5528	65.8%
Video call interviews	2281	27.1%
Web interviews	596	7.1%
Total	8405	100%

7.4 Response by country of issue

Survey response (including the mop-up survey but excluding emigrants) varied by country. Scotland had the highest response rate (75%), then England (73%) and Wales had the lowest response rate (71%). See table 11.

Table 11: Response by country (including pilot, main stage and Mop-up but excluding emigrant cases)

	ENGLAND		SCOTLAND		WALES		JERSEY/ GUERNSEY/ ISLE OF MAN		TOTAL	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Total issued	9,756	100	1,020	100	664	100	48	100	11,488	100
Productive	6,973	71	746	73	454	68	42	88	8,215	72
Non-contact	339	3	50	5	28	4	0	0	417	4
Refusal	1,563	16	133	13	123	19	4	8	1,823	3
Other unproductive	297	3	29	3	21	3	1	2	348	3
Unknown Eligibility (no-contact)	344	4	40	4	14	2	1	2	399	
Ineligible	240	2	21	2	24	4	0	0	285	2
Response rate	73%		75%		71%		88%		73%	
Base: All issued sample (excluding emigrants, those without address and data deletion cases) 11,488										

7.5 Response rates for each element

Further detailed information on the response to each element is available in the NCDS - Age 62 Survey: Technical Report in section 12.8.

7.5.1 Paper self-completion: Your Life Now

A total of 7,074 paper self-completion questionnaires are included in the data. Those taking part in the mop-up stage were not asked to complete a paper-self completion questionnaire, but 29 cases did receive and complete this. These cases were initially invited to take part in an earlier stage (and were provided with a paper questionnaire then) but did not take part until the mop up. These cases are included in the shared data but are excluded from table 12 along with proxy interviews who were not asked to complete a paper self-completion questionnaire.

Excluding mop-up cases and proxy interviews we have 7,042 questionnaires returned from the main stages, a response of 90%.

Table 12: Completion of paper self-completion questionnaire (by interview mode)

	In-person		*Video*		*Total*	
	N	*%*	*N*	*%*	*N*	*%*
Completed PSC	4,910	88.9	2,132	93.5	7,042	90.3
Not Completed PSC	610	11.1	147	6.5	757	9.7
TOTAL	5,520	100	2,279	100	7,799	100
Base: all 7,799 fully productive and partially productive cases from main stage, pilot and video pilot (proxy and mop-up cases excluded from this table)						

7.5.2 Paper self-completion: Childhood (Life History Questionnaire)

A total of 6,753 paper self-completion questionnaires are included in the shared data. Those taking part in the mop-up stage were not asked to complete a paper-self completion questionnaire, but 30 cases did receive and complete this. These cases were initially invited to take part in an earlier stage (and were provided with a paper questionnaire then) but did not take part until the mop up. These cases are included in the shared data but are excluded from table 13 along with proxy interviews who were not asked to complete a paper self-completion questionnaire.

Table 13: Completion of paper self-completion questionnaire by interview mode

	In-person		*Video*		*Total*	
	N	*%*	*N*	*%*	*N*	*%*
Completed PSC	4,605	83.4	2,117	92.9	6,722	86.2
Not Completed PSC	915	16.6	162	7.1	1,077	13.8
TOTAL	5520	100	2,279	100	7,799	100
Base: all 7,799 fully productive and partially productive cases from main stage, pilot and video pilot (proxy and mop-up cases excluded from this table)						

7.5.3 Completion of sensitive questions

In total, 95% respondents completed the section with sensitive questions. Rate of completion varied by mode, with 96% of those interviewed in-person completing the self-completion module compared to 93% interviewed by video.

Table 14 provides further detail on the completion of this section.

Self-completion mode is denoted by variable: '**nd10sc_mode'**.

Table 14: Completion of sensitive questions

	In-person		*Video*		*Total*	
	N	%	N	%	N	%
Web completed during interview	-	-	1,294	56.7	1,294	16.6
Web completed after interview	-	-	426	18.7	426	5.5
Completed with interviewer during video interview	-	-	391	17.1	391	5.0
Self-completion during in-person/telephone interview	5,002	90.5			5,002	64.1
With interviewer during in-person interview	294	5.3			294	3.8
Mode of completion unknown/not asked			10	0.4	10	0.1
Not completed	232	4.2	161	7.1	393	5.0
TOTAL	5,528	100	2,282	100	7,810	100

7.5.4 Cognitive assessments completion

The cognitive assessments include immediate word-list recall, animal naming, letter cancellation and delayed word recall, which were completed during an in-person or video interview.

The agreement rate for all four of the cognitive assessments was high, and with the exception of the letter cancellation test, rates of completion were similar across both in person and video call modes.

The letter cancellation test had a slightly lower response of 84%, with a lower completion rate in the video mode compared to the in-person mode (79% vs. 87% respectively).

The letter cancellation task had to be posted to the cohort member before the video interview which may explain the lower agreement rates in this mode – some

participants may not have received the task or were unable to locate it when required.[7]

Table 15: Completion of cognitive tests

	In-person		Video		Total	
	N	%	N	%	N	%
Immediate word recall test						
Completed	5,399	98	2,250	99	7,649	98
Not completed	116	2	25	1	141	2
Animal naming						
Completed	5,400	98	2,251	99	7,651	98
Not completed	115	2	24	1	139	2
Letter cancellation						
Completed	4,781	87	1,790	79	6,571	84
Not completed	734	13	485	21	1,219	16
Delayed word recall test						
Completed	5,399	98	2,250	99	7,649	98
Not completed	116	2	25	1	141	2
Base_ 7,790 cases asked to complete cognitive tests, excludes mop-up cases, proxy cases and some partial interviews						

7.6 Response to health visit

In total 7,774 NCDS cohort members were asked during their interview if they would be willing to be contacted by a Health Professional8 and 7,166 cohort members agreed to this (92%).

Of those who agreed to be visited, 6,309 took part in the Health Visit (88%). Table 16 outlines the response to each of the individual biomeasures.

[7] A small number of pilot letter cancellation forms were destroyed in error prior to being coded, as a result, these cases are missing from the data.

[8] Proxy cases, partial cases and mop-up cases were not asked permission to be contacted by a Health Professional.

Table 16: Response to each biomeasure

	At least 1 measure obtained	Base (eligible for measure)	Response (%)
Blood pressure – seated	6,246	6,309	99
Blood pressure – standing	6,175	6,309	98
Grip strength	6,177	6,241	99
Blood sample	5,149	5,966	86
Consent to blood storage for future analysis, and blood sample obtained	5,071	5,966	85
Consent to DNA extraction and storage, and blood sample obtained	4,982	5,966	84
Weight	6,125	6,237	98
Body fat	5,906	6,187	95
Waist	6,231	6,309	99
Hip	6,228	6,309	99
Timed Walk	5,998	6,009	100
Leg raise	5,970	5,990	100
Base: productive health visits, for which cohort member was eligible to complete the particular measurement			

8. Survey Research Data

8.1 Licensing and data access

The NCDS Age 62 survey research have been processed by CLS and supplied to the UK Data Service (UKDS).

All users of the data need to be registered with the UKDS and to sign the UKDS End User Licence. Details of how to do this are available at www.ukdataservice.ac.uk/get-data/how-to-access/registration.

Please refer to section 8.10 for information on how these data have been de-identified for sharing.

8.1.1 Safeguarded data (EUL)

The majority of the NCDS Age 62 survey data are available from the UK Data Service as safeguarded data, which can be downloaded once the End User Licence (EUL) has been signed by the user.

The safeguarded data exclude detailed information that presents a potential risk for disclosure or is of sensitive nature, which is instead shared as controlled data.

8.1.2 Controlled data (Secure Access, SA)

Some NCDS survey must be accessed as controlled data from the UK Data Service SecureLab due to their potentially disclosive and/or sensitive nature. This applies to:

1. uncommon health conditions (including age at diagnosis)
2. full employment codes and income/finance details
3. specific life circumstances (e.g. pregnancy details, year/age of emigration from GB).

Applicants wishing to access this data need to abide by the terms and conditions of the UK Data Service Secure Access licence. Before gaining access, researchers must make an application 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/.

Data access will be granted once the form has been reviewed by UK Data Service and approved by the CLS Data Access Committee.

8.2 Datasets

The Age 62 survey research data consists of two wide format datasets, which contain the main survey and the biomeasures data, and a number of long format (hierarchical) datasets. Datasets are organised by topic. Dietary diary data, additional derived variables and full derived geographical identifiers will be made available separately in the future (see section 8.2.4).

8.2.1 Wide format datasets

The data from main interview and from the nurse interview are presented in wide (flat) format, that is, with one row per cohort member (CM). The identifier is ncdsid, which is used for merging with other NCDS datasets from this and other sweeps.

Table 17: List of wide format datasets

Dataset name	Contents	Data access
ncds10_age62_main_interview	MainStage interview Your Life Now self-completion Life History self-completion MopUp interview Derived variables Weights	EUL + SA
ncds10_age62_biomeasures	Nurse/Health visit interview General blood test results Heart health markers Derived variables	EUL + SA
ncds10_age62_main_interview_unfolding_brackets	Questions that were asked as unfolding brackets once (not in a loop, the looped brackets are in the longf files)	EUL

8.2.2 Long format files

The long format (hierarchical) datasets contain information from interview questions that require multiple responses, thus containing one row for each response given by

a cohort member. These questions are common in several sections of the questionnaires and for a variety of topics, for example relationship to each of their children (one row per child mentioned), employment history (one row per employment incidence mentioned), benefits (one row per type of benefit mentioned), covid vaccinations (one row per vaccination mentioned), debt (one row per debt mentioned).

These datasets are identified with suffix _**longf** in the file name. Each of them contain 2 identifiers:

- ncdsid, which is used for merging with other NCDS datasets from this and other sweeps
- an additional ID variable, either named **longid** or **n10gridid.**

The **longid** identifier refers to the row number of the CM responses. This identifier is exclusive to each long format dataset and therefore the _**longf** datasets are neither mergeable to each other, nor to other datasets from other sweeps.

The **n10gridid** refers to the HH member within an NCDS family for sweep 10.

Table 18: List of long format datasets

Dataset name	Contents	Data access
ncds10_age62_persons_grid_longf	Household members	EUL+SA
ncds10_age62_medicin_longf	Medication	EUL+SA
ncds10_age62_long_acting_medicin_longf	Long acting medication	EUL+SA
ncds10_age62_benefits_longf	Benefits	EUL+SA
ncds10_age62_casichrel_longf	Relationship with child	EUL
ncds10_age62_child16ab_longf	Children above 16	EUL+SA
ncds10_age62_covidbens_longf	Benefits – Covid	EUL+SA
ncds10_age62_covidvax_longf	Covid vaccinations	EUL+SA
ncds10_age62_debt_longf	Debt	EUL+SA
ncds10_age62_econshock_longf	Economic shock	EUL
ncds10_age62_econshockp_longf	Economic shock - partner	EUL
ncds10_age62_emphist_longf	Employment history	EUL+SA
ncds10_age62_giftgiv_longf	Gifts given	EUL+SA

Dataset name	Contents	Data access
ncds10_age62_giftrec_longf	Gifts received	EUL+SA
ncds10_age62_housinghist_longf	Housing history	
ncds10_age62_inheritance_longf	Inheritance	EUL+SA
ncds10_age62_parent_longf	Parents	
ncds10_age62_pension_longf	Pensions	EUL+SA
ncds10_age62_qual_academic_longf	CM's qualifications: academic	EUL+SA
ncds10_age62_qual_vocational_longf	CM's qualifications: vocational	EUL
ncds10_age62_qual_other_longf	CM's qualifications: other	EUL
ncds10_age62_qual_partner_vocational_longf	Partner's qualifications: vocational	SA
ncds10_age62_regular_financial_help_longf	Regular financial help	EUL+SA
ncds10_age62_regular_income_longf	Regular income	EUL+SA
ncds10_age62_savings_longf	Savings	EUL+SA

Case counts in the long format datasets may differ slightly from those reported in the wide datasets. These discrepancies can arise due to changes made to responses during the interview— for example, if a respondent starts answering a question that triggers a loop but later goes back to change their response, this can affect whether and how entries appear in the looped data.

8.2.3 Additional datasets

Additional Age 62 Survey data available at the UKDS are shown in table 19.

Table 19: List of additional datasets

Dataset name	Contents	Data access
ncds_response	Outcomes across sweeps, with the addition of ncds10 and the information on whether issued/deceased	EUL
ncds10_metabolomics_v1	A panel of 250 metabolites measurements and 21 QC flags from Nightingale Health metabolomics	Special Licence

8.2.4 Future datasets

Due to the large-scale nature of this survey, two further data deposits have been planned to provide additional data which was not possible to deposit with the initial deposit. Data will be deposited over two phases.

Phase 1

The following data will be added as part of the first data update:

- additional cases: data collected from 37 participants (proxy and partial interviews) has not been included in some of the long format files and will be added at a later stage. This does not apply to medications data, employment history and qualifications long format datasets, which contain all relevant data.
- height and weight variables from the Life History Questionnaire
- updated geography variables to distinguish moving between stages of participation, as the current geography variables only refer to the address at the last point of interview (main stage or mop up).

Phase 2

Dietary diary data, additional derived variables and full derived geographical identifiers will be made available later in 2025.

In addition, the variables included in the long format datasets (all apart from the persons_grid) will be added to the wide format files that contain main interview data. The aim is for these variables to be available in both formats in the future.

8.3 Data documentation

In addition to this User Guide, the following documentation accompanies the data shared via the UKDS:

Table 20: NCDS survey documents

Name of the document	Content summary
NCDS_sweep10_age62_questionnaire.pdf	Questions asked in the Age 62 Survey. It includes details on routing as well as both mode and stage specific adjustments. Questions asked as part of the 'mop-up' stage are flagged.
NCDS_sweep10_age62_Your Life Now self-completion questionnaire.pdf	Self-completion questionnaire – 'Your Life Now'
NCDS_sweep10_age62_Childhood self-completion questionnaire.pdf	'Childhood self-completion questionnaire' (also known as the Life History Questionnaire – LHQ)
NCDS_sweep10_age62_technical_report.pdf	Technical details regarding the design and implementation of the survey. This document has been produced by the leading fieldwork agency – NatCen.
NCDS _sweep10 - age62: technical report – appendix A – fieldwork documents.pdf	All fieldwork documents provided to respondents during the survey. This document has been produced by the leading fieldwork agency - NatCen
NCDS_sweep10_age62_technical report appendix B_biomedical protocols.pdf	All protocols regarding the design and implementation of the Health Visit. This document has been produced by the leading fieldwork agency – NatCen

8.4 Main Stage and Mop Up combined data

As described in section 4.3.8, the main stage fieldwork was followed by a mop-up web survey. This survey targeted cohort members who had not responded to the main interview, including emigrants (who were not originally included in the main survey). Additionally, participants who had completed the main interview before the COVID-19 pandemic—specifically, those interviewed between January and March 2020, as well as pilot and dress rehearsal cases - were also invited to take part in the mop-up survey. For these participants, data is available at two time points.

The variable **n10mu_samptype** identifies the type of mop-up respondent, indicating whether the case was a non-responder to the main survey, an emigrant, or a pre-COVID participant completing the survey a second time.

The variable **n10part** shows whether a participant has taken part in the main stage, the mop-up, or both, as follows:

(1) Once – Main stage

(2) Once – Mop up

(3) Twice – Main stage & Mop up

Most of the questions in the mop-up survey were also included in the main interview. Where questions were common to both surveys, data from the mop-up survey (for main stage non responders and emigrants) has been merged with the main stage data. A small number of questions in the mop-up survey were unique and were not asked in the main interview. The suffix **mup** is used for variables that exist only in the mop-up (these variables will include all mop-up cases who responded).

The mop-up data for respondents who participated twice, both in the main stage (before the COVID-19 pandemic) and the mop-up survey has been added as additional variables to the main stage data. These mop-up variables names are suffixed with _**r2**. This indicates data from a second time-point and includes only those cases interviewed twice.

Two sets of derived variables are also available for mop-up participants who participated twice. Derived variables from the second interview are available with the suffix _r2. For example – the derived variable nd10loneliness provides the UCLA loneliness scale based on their first interview and nd10loneliness_r2 provides their UCLA loneliness scale from their second interview. This then allows comparisons to be made between two time points.

8.5 Variable description

8.5.1 Variable order

The order in which variables appear in the datasets broadly follows the order of sections, and of questions within sections, of the survey instruments. However, due to the repeating loop nature of some sections of the survey instruments the order is determined by the structure of the CAI program, which does not necessarily hold each question in the order in which they are put to the respondent. As a result, several variables have been re-ordered so similar variables are together.

The survey included several stages with questions occasionally asked in a different order, or in one stage and not another. Variable order primarily follows the main stage questionnaire structure, and mop-up survey only variables moved after the rest of the main survey questions rather than their original order.

8.5.2 Variable names

The variable names are all prefixed by ‘n10’, denoting the wave/sweep of the cohort study. For ease of tracking variables longitudinally, other than the prefix, variable names are consistent with those used the prior sweep (where the prefix was ‘n9’).

The variable names are based on those used in the CAI program and are documented in the questionnaire and self-completion questionnaire documentation, but do not match exactly in all cases.

The mop-up has several survey specific variables which are identified in the following ways (See section 8.4 for a detailed description):

- _**r2** is used for variables that exist on both the main stage and the mop-up and indicates data from a second time point (only includes data from cases who have taken part twice).
- _**mup** is used for variables that exist only in the mop-up.

For multi-coded variables, where a single question produces more than one response, a suffix has been used to identify the iteration. 01, 02, 03…..been used to denote the 1st, 2nd, 3rd,...iterations respectively. Any new codes which were added following the interview, during the coding process, have been allocated a suffix starting at 50 e.g. ….n10wpreas50, .n10wpreas51 are codes which were added following a review of what the cohort members detailed in their ‘other specify’ response.

Variables from the paper self-completion questionnaire have names derived from the question numbers as they appear on the printed questionnaire (e.g. Q1A from YLN = ‘n10ylq1a’, Q1 from LHQ = ‘n10lhq1’).

8.5.3 Variable labels

The variable labels included in the dataset are based on the question wording that can be found in the core interview and self-completion questionnaire documentation. Where necessary, labels have been modified in an effort to ensure they are comprehensible and accurate. Certain mode/stage information is added to the start of labels for ease of identification. These are as follows:

- ‘(SC)’- Variables in the self-completion section of the main survey

- ‘(YLN PSC)’- Variables from the ‘Your Life Now’ paper self-completion
- ‘(LHQ PSC)’- Variables from the ‘Life History Questionnaire’ also known as the ‘Childhood’ paper self-completion.

8.5.4 Value labels

The value labels for valid responses are based on the question responses used in the CAI program as documented in the questionnaire documentation. Value labels have been individually reviewed and amended, where necessary.

8.6 Derived variables

Several derived variables have been produced based on the questionnaire data. Detailed documentation on their derivation can be found in Appendix 1.

Derived variables in the dataset are given the prefix “nd10”.

8.7 Income and payment unfolding brackets

A feature of income or payment questions is the use of unfolding brackets for those cases where a respondent refuses or is unable to provide an exact answer. The unfolding brackets questions are designed to elicit a minimum and maximum value that define a range or “closed band” within which the actual value lies. On entering the unfolding brackets, respondents are asked to say whether they have more, less or about the same as a particular value. This question is repeated using different values (which will be a lower or higher value depending on the answer to the preceding question). The procedure stops at the point when either: an upper and lower bound is provided; the respondent refuses or says “don’t know”; or the respondent places themselves in the top or bottom bracket. The unfolding bracket questions are randomly ordered for each respondent. This will average any possible 'anchoring' effects (i.e. where people use the suggested figure as a reference point and adjust it to reach their answer) from the procedure across the distribution. The bracket values are selected based on the density of the underlying financial variable.

8.8 Missing values

Missing values are consistently labelled as follows (unless otherwise stated):

-9 = Refusal

-8 = Don't Know

-3 = Not asked at case fieldwork stage

-2 = Not asked due to scripting/routing error /measurement not taken

-1 = Item not applicable

In the Age 62 survey –3 has been reserved for questions not asked at a particular stage e.g. not asked in the pilots, soft launch or mop-up.

Code -2 has been used in a small number of cases where a response was missing, caused by a respondent giving a particular answer at one question and going back and changing their answer, resulting in a subsequent question being missed. It has also been used has to flag questions unanswered as a result of a routing error, or in the nurse dataset for measurements not taken/aborted after initial attempt. Routing errors are described below.

The value -1 is used for missing responses to questions which study members would not have been asked if they only partially completed the survey (or they were not routed to this question). For derived variables -8 is typically reserved for 'Not codeable' values, where there is insufficient data for the variable to be derived.

Missing data due to routing errors

During the data editing and cleaning process, routing errors, resulting in missing data was discovered in three variables, noted in the table below.

In addition to this, data has not been deposited by stroke type due to a questionnaire error in the description of 'Ischemic stroke' and 'Haemorrhagic stroke'.

Table 21: Missing data due to routing error

Variable name	Variable label	Cases missing	% of total base
n10lolw	Whether illness/conditions reduce ability to work	80	3%

Variable name	Variable label	Cases missing	% of total base
n10finlit3	Financial literacy – interest one year	150	2%
n10finlit4	Financial literacy – interest after 5 years	32	0.4%

8.9 Data cleaning of back-coded variables ('other')

Where possible, 'Other' variables have been back-coded to provide categorical data from these open-text responses. Questions that include 'Other (please specify)' categories allow the respondent to give open text responses that are back coded after the interview is completed. Some of these variables are used in filtering cases to subsequent questions. Where backcoding has occurred after the interview, the value will not be used for filtering.

8.10 Data de-identification

In addition to the pseudonymisation (use of IDs), the data have been examined for sensitive topics and disclosive information, as well as for rare responses (low counts), and the data have been distributed to safeguarded (EUL) and controlled (SA) datasets.

Sometimes information is too disclosive or sensitive to be shared in the EUL version. In these cases, the complete response in those variables is made available under Secure Access (SA), and a de-identified version is created and released as safeguarded data (EUL version).

The following methods have been used to de-identify variables that can be shared under EUL. Certain suffixes have been used in variable names to highlight that they have been de-identified:

- Truncation: the truncated variables in the EUL version are named with suffix _**tr.** This has been applied to Socio Economic Codes (SOC/SIC), ICD and BNF medications.
- Recoding: the recoded variables in the EUL version are named with suffix _**rec**. This has been applied in different manners depending on the variable contents:

    - Grouped in the upper and/or lower end of the distribution where the values have low counts (e.g. height, weight).
    - Recoded certain values that contain sensitive or disclosive information.
- Creation of a flag: a variable with the suffix _**flag** indicates whether any of a group of variables contains a response. In the meantime, the flag provides the information on whether at least one of the conditions in that group is mentioned (e.g. any type of cancer, or any type of eye sight condition, or any type of hearing condition) and it is available on the safeguarded data (EUL). All the variables from that group are available as controlled data under Secure Access (SA).

For certain potentially disclosive multi-coded data, including health conditions, low-count responses have been combined into a new variable labelled 'Other answer(s) from code frame (Derived)'. This variable combines all coded and back-coded categories not shared as safeguarded data under EUL, while the full breakdown is available as controlled data under Secure Access.

In addition, all text variables that contain verbatim information provided by the respondents have been removed from both the EUL and SA research datasets. This includes job titles, job descriptions, exact names of education institutions, town name, postcodes and the final open-ended question. These potentially identifiable CLS data can be accessed securely by applying directly to the CLS Data Access Committee.

8.11 Weights variables

The variables containing the calculated weights are as follows:

Variable name	Dataset	Variable Description
nd10weight_mainmopup	ncds10_age62_main_interview	(Derived) Non-response weight for main survey and mopup (1st response and 2nd response)
nd10weight_yln	ncds10_age62_main_interview	(Derived) Non-response weight for Your Life Now paper self-completion
nd10weight_lhq	ncds10_age62_main_interview	(Derived) Non-response weight for the Life History Questionnaire paper self-completion

Variable name	Dataset	Variable Description
nd10weight_nurse	ncds10_age62_biomeasures	(Derived) Non-response weight for biomeasures: nurse visit
nd10weight_blood	ncds10_age62_biomeasures	(Derived) Non response weight for biomeasures: blood markers

8.12 Output Disclosure Control (for controlled data)

The two UK Data Service Secure Lab rules of thumb that will be applied to all research outputs (summary tables, graphs, etc) are:

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

The controlled data included in Table 17 is only available via the UKDS Secure Lab. 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: www.securedatagroup.org/sdc-handbook/

9. Derivation and Implementation of Non-Response Weights

9.1 Introduction

Non-response is common in longitudinal surveys. Missing values mean less efficient estimates because of the reduced size of the analysis sample. It can also introduce the potential for bias since respondents are often systematically different from non-respondents. To support researchers in producing robust analyses, we have developed comprehensive advice on how to deal with missing data (www.cls.ucl.ac.uk/data-access-training/handling-missing-data/).

The approaches we recommend to researchers capitalise on the rich data cohort members provided over the years before their nonresponse. These approaches include well known methods such as Multiple Imputation (MI), Inverse Probability Weighting (IPW), and Full Information Maximum Likelihood (FIML).

To correct for non-response in the Age 62 NCDS survey, non-response weights are provided with the data, so that IPW analysis can be undertaken, either in isolation or in combination with MI or FIML. This section of the User Guide describes the derivation and implementation of these non-response weights. The weights were created and documented by Liam Wright and Richard Silverwood and closely followed the procedures used to derive non-response weights for CLS' COVID-19 Surveys (www.cls.ucl.ac.uk/wp-content/uploads/2017/02/UCL-Cohorts-COVID-19-Survey-user-guide.pdf).

9.2 Target population and response

We created five weights for response to specific elements of the Age 62 NCDS Survey:

1. Completed Mainstage Survey or Mop-Up Questionnaire
2. Completed 'Your Life Now' Self-Completion Questionnaire
3. Completed Childhood Self-Completion Questionnaire

4. Completed Health Visit
5. Provided a Valid Blood Sample

For the purposes of weighting, we defined two target populations.

- For #1 (Completed Mainstage Survey or Mop-Up Questionnaire), the target population was individuals born in 1958 in Great Britain and who were alive at the start of fieldwork (n = 16,000); emigrants were technically eligible for the mop-up questionnaire, so were included in the target population for this weight.
- For #2-5, the target population was individuals born in 1958 in Great Britain and who, at the start of fieldwork, were alive and residing in GB (n = 14,591); this corresponded to cohort members who were eligible for these respective elements of the Age 62 NCDS Survey.[9] Respondents to the self-completion questionnaires, the health visit, and those who provided a valid blood sample were subsamples of those who responded to the main interview. Emigrant responders and web-mop-up participants were not invited to complete these components of the Age 62 NCDS survey.

Response rates based on the respective target populations were:

1. Completed Mainstage Survey or Mop-Up Questionnaire (n = 8,405, 52.5%)
2. Completed 'Your Life Now' Self-Completion Questionnaire (n = 7,045, 48.4%)
3. Completed Childhood Self-Completion Questionnaire (n = 6,723, 46.2%)
4. Completed Health Visit (n = 6,309, 43.4%)
5. Provided a Valid Blood Sample (n = 5,149, 35.5%)

[9] We created a weight for 'Completed Mainstage Survey' among the non-emigrant sample used in weights #2-5. As this correlated > 0.99 with the 'Completed Mainstage Survey or Mop-Up Questionnaire' (#1) weight, we have not released this in order to simplify the dataset.

9.3 Derivation of non-response weights

The derivation of each Age 62 Survey non-response weight proceeded as follows:

1. Within the sample corresponding to the target population (e.g., those alive and living in Britain for #2-5), for each survey element, model Age 62 survey response conditional on a set of covariates using logistic regression. The selection of covariates – further discussed below – was informed from results of the CLS Missing Data Strategy (www.cls.ucl.ac.uk/data-access-training/handling-missing-data).
2. Among respondents to the relevant Age 62 survey element, calculate the probability of response implied by each model.
3. Calculate each Age 62 non-response weight as the inverse of the probability of response.
4. Examine the distribution of derived non-response weights to decide whether truncation may be desirable, applying truncation if so.
5. Finally, calibrate the Age 62 non-response weights so that they sum to the number of respondents for each response type.

The variables included in the response model in stage 1 are listed in Table 22. We used the same set of variables as used to derive weights for the NCDS in CLS's COVID-19 Surveys, with the exception that we also included a variable for the number of COVID-19 sweeps a participant responded to (continuous; 0-3). We additionally included a variable for Age 42-44 biomedical response (categorical: non-respondent; completed nurse visit; gave blood sample; allowed DNA extraction).[10] The code used to clean the data and create the weights can be viewed at www.osf.io/eha23/.

[10] These categories are progressive: blood samples were taken as part of the nurse visit, and DNA from extracted from blood samples (i.e., all people who allowed DNA extraction gave a blood sample as part of the nurse visit).

Table 22: Variables used as predictors in models used to create non-response weights

Variable	
Sex	Categories: Male, Female.
Social Class @ Birth	Categories: Professional or Managerial, Intermediate, Partly-Skilled or Unskilled.
Rooms per Person @ Age 0y	Mean: 1.5, Range: 1 - 6
Cognitive Ability @ Age 7y	Mean: 0, Range: -4.2 - 2.8
Externalising @ Age 16y	Mean: 0, Range: -1.2 - 4.6
Internalising @ Age 16y	Mean: 0, Range: -1.4 - 4.4
Educational Qualifications @ Age 42y	Categories: None, NVQ Level 1-3, NVQ Level 4-5.
Organisation Membership @ Age 42y	Categories: No, Yes.
Union Membership @ Age 42y	Categories: No, Yes.
Voted in 1997 General Election	Categories: No, Yes.
Consent to Biological Samples @ Age 44y	Categories: No Consent, Bio Sweep, Blood Sample, DNA Extraction.
Malaise @ Age 50y	Mean: 7.5, Range: 0 - 9
Self-rated Health @ Age 50y	Categories: Excellent or Very Good, Good, Fair or Poor.
BMI @ Age 50y	Mean: 27.4, Range: 14 - 74.7
Smoking Status @ Age 50y	Categories: Never, Former, Current.
Income Quintile @ Age 50y	Categories: 1, 2, 3, 4, 5.
Employed @ Age 50y	Categories: Currently employed, Not currently employed.
Marital Status @ Age 50y	Categories: Single & Never Married, Married or Civil Partnered, Separated, Divorced or Widowed.
Internet Access @ Age 50y	Categories: Yes, No.
Social Listen @ Age 50y	Categories: A Little or Not at All, Somewhat, A Great Deal.
Social Visits @ Age 50y	Categories: Never, Fairly Frequently, Very Frequently.
Social Trust @ Age 50y	Categories: Most people can be trusted, Can't be too careful, Other/depends.
# Previous Nonresponses	Mean: 3.2, Range: 0 - 10
# Nonresponses (COVID)	Mean: 1.9, Range: 0 - 3

Missing values in the above variables were handled using multiple imputation (MI). The imputation procedures were carried out in the eligible sample and the imputation models included the variables above plus response for the relevant element of the Age 62 Survey. Separate imputation models were conducted for each survey weight being derived (twenty imputed datasets created using chained equations). Twenty

was deemed sufficient as only a point estimate (probability of Age 62 survey response) was to be estimated from the MI analysis; more imputations may be required for inference.

Models for Age 62 survey response were fitted in each imputed dataset. Pooled models are reported in Appendix 2. From these models, the probability of Age 62 survey response was predicted for each respondent, with the non-response weight calculated as the inverse of the response probability. These weights were then averaged to get a single weight per individual.

Test analyses were conducted in each cohort at different levels of weight truncation (max = 10, 20, and 50). This suggested that truncation at a maximum of 20 could provide some improvement in precision without undue introduction of bias. The non-response weights were therefore truncated to 20 and were then calibrated so that they summed to the number of Age 62 survey respondents for the relevant survey element by multiplying them by the ratio of the number of responses to the total of the uncalibrated nonresponse weight. The distributions of the resultant calibrated non-response weights are presented in Table 23 along with the distribution prior to truncation.

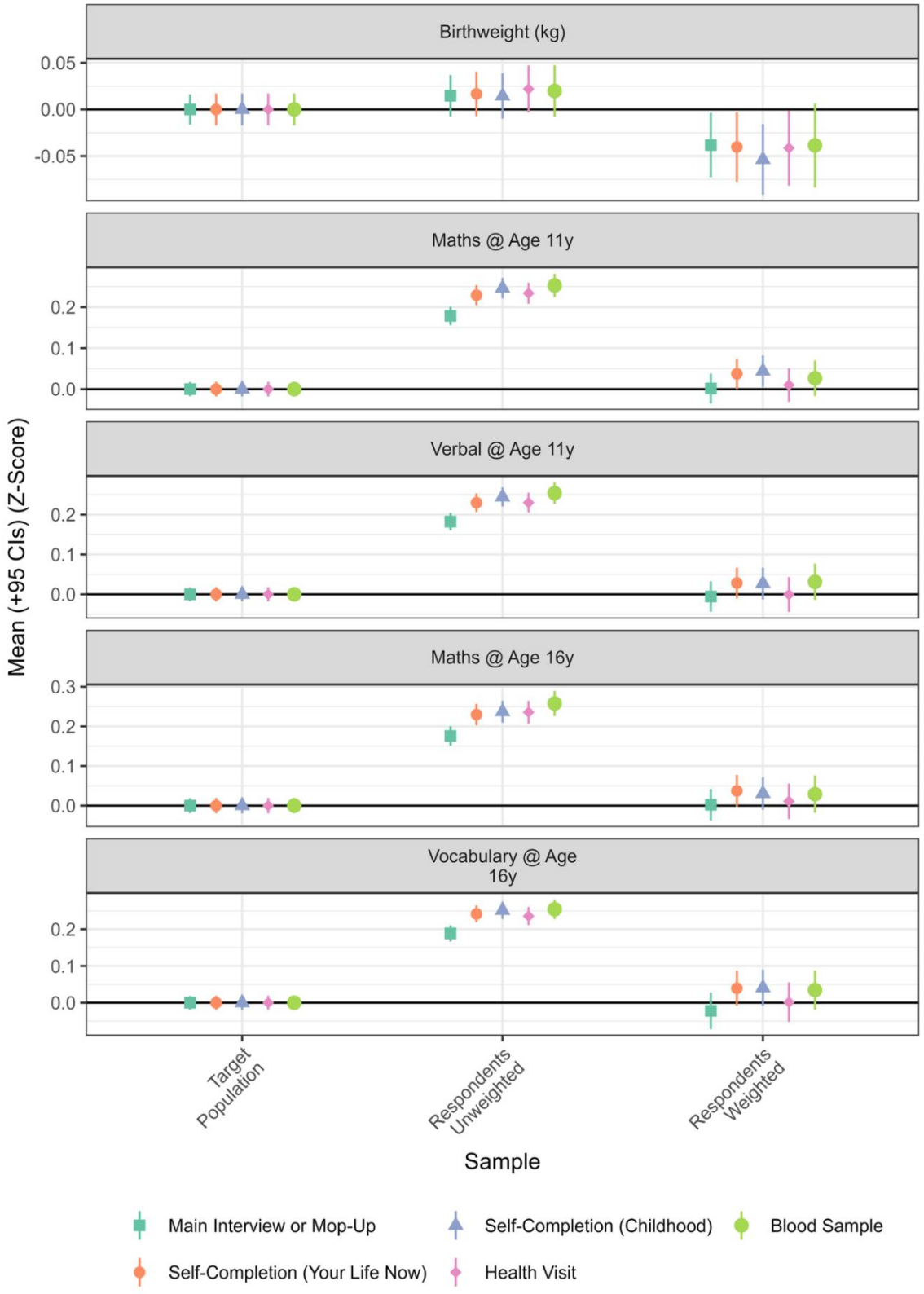
Table 23: Distributions of the non-response weight (prior and post truncation and calibration)

	Main Interview or Mop-Up		**Self-Completion (Your Life Now)**		**Self-Completion (Childhood)**		**Health Visit**		**Blood Sample**	
	Prior	Post	Prior	Post	Prior	Post	Prior	Post	Prior	Post
Mean	1.73	1.00	1.87	1.00	1.94	1.00	2.11	1.00	2.50	1.00
SD	2.88	1.29	2.72	1.21	2.85	1.19	3.57	1.26	3.70	1.16
0%	1.01	0.60	1.04	0.57	1.05	0.56	1.06	0.53	1.14	0.47
5%	1.02	0.61	1.06	0.58	1.08	0.57	1.10	0.55	1.21	0.50
25%	1.03	0.61	1.09	0.60	1.11	0.59	1.15	0.57	1.31	0.54
50%	1.07	0.64	1.18	0.65	1.22	0.65	1.27	0.63	1.51	0.63
75%	1.38	0.83	1.59	0.88	1.68	0.89	1.76	0.88	2.20	0.91
95%	3.84	2.30	4.19	2.30	4.47	2.37	4.78	2.38	6.15	2.56
100%	81.32	11.95	69.30	10.99	66.07	10.60	87.88	9.94	80.82	8.33

9.4 Weight effectiveness

To examine the effectiveness of the derived non-response weights in restoring sample representativeness we conducted several analyses, some of which are presented here (with several more in Appendix 3). We considered the sample means of five variables – birthweight and four measures of cognitive ability at ages 11y and 16y – in the full sample and in the (weighted and unweighted) sample of Age 62 respondents. These variables were collected early in cohort members lives so were observed for a very high proportion of participants. Further, the variables were 'unseen' in the sense that they did not appear as predictors in the response models, so are not balanced mechanistically. A good performing weight is indicated by having a similar weighted sample mean to that in the full sample. The results of this exercise are presented in Figure 2. We observed that there is considerable bias when using the unweighted Age 62 survey data but this is markedly reduced – though not completely – when applying the non-response weights.

Figure 2: Means for five variables that were not used to create non-response weights in the target population and in weighted and unweighted Age 62 responding sample. Variables standardised (mean = 0, SD = 1) using the target population mean and SD



Appendix 3 shows the results of a similar exercise, this time looking at the ability of the weights to recover the distribution (means for continuous variables, proportions for categorical variables) of variables that appeared in the response predictor models (e.g., sex). Weighting reduces bias in almost all cases, in several recovering sample means and proportions close to those in the full sample. Bias remains in several variables however, indicating that either observed information in NCDS does not fully account for attrition or that to fully restore representativeness, it may be necessary to adjust the weighting of these variables or apply a targeted multiple imputation approach, both of which would incorporate additional information from NCDS.

8.5 Implementation of non-response weights

The choice of non-response weight to use depends on the analysis being conducted. As the paper self-completion questionnaire, health visit and blood samples were completed by subsamples of the main interview respondents, users of NCDS Age 62 sweep data who want to use variables from these survey elements should consider using the relevant non-response weight. Note, however, that different non-response patterns may warrant the creation of extra weights *de novo*. For instance, data users who want to use data from both the self-completion questionnaires and the nurse visit should consider creating their own weights for this specific sample or apply analysis specific MI or FIML. Further, data users wanting to use variables from other sweeps should consider creating weights to reflect the particular response pattern their analysis implies – non-response patterns in the NCDS are not uniformly monotonic. In this instance a combination of IPW and MI, or just MI may be more suitable.[11]

[11] Note, some analyses using variables from the mainstage survey may drop data from the mop-up questionnaire as this was shorter. As noted above, we created (but have not released) a weight for 'Completed Mainstage Survey' among the non-emigrant sample used in weights #2-5. This correlated > 0.99 with the 'Completed Mainstage Survey or Mop-Up Questionnaire' (#1) weight, so it may not be necessary to create a new weight when using the mainstage survey data only.

References

F. Babor, T., C. Higgins-Biddle, J., B. Saunders, J., & G. Monteiro, M. (2001). The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care. W. H. Organisation.

Brown, M., Fitzsimons, E., Goodman, A., Peters, A., Ploubidis, G.B., Sanchez, A., Silverwood, R., Smith, K. (2021) COVID-19 Survey in Five National Longitudinal Studies: Waves 1, 2 and 3 User Guide (Version 4). London: UCL Centre for Longitudinal Studies and MRC Unit for Lifelong Health and Ageing. www.cls.ucl.ac.uk/wp-content/uploads/2017/02/UCL-Cohorts-COVID-19-Survey-user-guide.pdf.

Brown, M., Peters, A. (2023) 1970 British Cohort Study: Age 46 User Guide (Version 2). UCL Centre for Longitudinal Studies Cutrona CE, Russell DW. The provisions of social support and adaptation to stress. Advance in Personal Relationships. 1987;1:37–67

Cutrona CE, Russell DW. The provisions of social support and adaptation to stress. Advance in Personal Relationships. 1987;1:37–67

Daniel W. Russell (1996) UCLA Loneliness Scale (Version 3): Reliability, Validity, and Factor Structure, Journal of Personality Assessment, 66:1, 20-40, DOI: 10.1207/s15327752jpa6601_2

Donnellan MB, Oswald FL, Baird BM, Lucas RE. The mini-IPIP scales: tiny-yet effective measures of the Big Five factors of personality. Psychol Assess. 2006 Jun;18(2):192-203. PubMed PMID: 16768595.

Ng Fat, L., Scholes, S., Boniface, S., Mindell J., & Stewart-Brown S. (2017) Evaluating and establishing the national norms for mental well-being using the short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS): findings from the Health Survey for England. Quality of Life Research, 26(5), 1129-1144.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. Psychological Assessment, 4, 26-42.

Hanzlová, R. and Lynn, P., 2023. Item response theory-based psychometric analysis of the Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS) among adolescents in the UK. Health and Quality of Life Outcomes, 21(1), p.108.

Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A Short Scale for Measuring Loneliness in Large Surveys: Results From Two Population-Based Studies. Res Aging. 2004;26(6):655-672. doi: 10.1177/0164027504268574. PMID: 18504506; PMCID: PMC2394670.

Johnson, J. A. (2014). Measuring thirty facets of the Five Factor Model with a 120-item public domain inventory: Development of the IPIP-NEO-120. Journal of Research in Personality, 51, 78-89.

Katsoulis, M., Narayanan, M., Dodgeon, B., Ploubidis, G., & Silverwood, R. (2024). A data driven approach to address missing data in the 1970 British birth cohort. *medRxiv*, 2024.2002. 2001.24302101.

Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med. 2007;146:317-25.

Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: Validity of a Two-Item Depression Screener. Medical Care. 2003;41:1284-92.

Michael T. McKay, James R. Andretta, Evidence for the Psychometric Validity, Internal Consistency and Measurement Invariance of Warwick Edinburgh Mental Well-being Scale Scores in Scottish and Irish Adolescents, Psychiatry Research, Volume 255, 2017,Pages 382-386, ISSN 0165-1781, https://doi.org/10.1016/j.psychres.2017.06.071.

Mostafa, T., Narayanan, M., Pongiglione, B., Dodgeon, B., Goodman, A., Silverwood, R. J., & Ploubidis, G. B. (2021). Missing at random assumption made more plausible: evidence from the 1958 British birth cohort. J Clin Epidemiol, 136, 44-54.

Ringdal, R., Eilertsen, M.-E., Bjørnsen, H., Espnes, G., & Moksnes, U. (2018). Validation of two versions of the Warwick-Edinburgh Mental Well-Being Scale

among Norwegian adolescents. Scandinavian Journal of Public Health, 46, 140349481773539. https://doi.org/10.1177/1403494817735391

Rubin, D. B. (1987). Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470316696

Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health, and behaviour.* London: Longman.McGee, R., Williams, S., and Silva, P. A. (1986) 'An evaluation of the Malaise Inventory', Journal of Psychosomatic Research, 30(2), pp.147-152.

Silverwood, R., Narayanan, M., Dodgeon, B., Katsoulis, M., Ploubidis, G. (2024) Handling missing data in the CLS cohort studies: User guide. London: UCL Centre for Longitudinal Studies.

Ware J.E., Snow K.K., Kosinsk,i M., & Gandek, B. (1993), SF-36 Health Survey Manual and Interpretation Guide. Boston, MA: New England Medical Center, The Health Institute.

Wiggins, R.D., Brown, M., Ploubidis G.B. (2017). A measurement evaluation of a six item measure of quality of life (CASP6) across different modes of data collection in the 1958 National Child Development Survey (NCDS) Age 55 years: Working Paper 2017/2: CLS-WP-20172.pdf

Wiggins, R. et al (2004) Quality of life in the third age: key predictors of the CASP-19 measure, Ageing & Society, 24, pp. 693–708.

Wiggins, R. et al (2008) 'The Evaluation of a Self-enumerated Scale of Quality of Life (CASP-19) in the Context of Research on Ageing: A Combination of Exploratory and Confirmatory Approaches', Social Indicators Research, 89, pp 61-77.

Wright, L., Ploubidis, G., Silverwood, R. (2024) Handling Mode Effects in the CLS Cohort Studies: Executive Summary. London: UCL Centre for Longitudinal Studies.

Appendix 1: Derived Variables

A1.1 Paradata

The following variables are derived from data surrounding the interview circumstances rather than data derived from within the interview itself.

ND10AGEINT/ ND10AGEINT_R2 - "(Derived) Age in months at interview"

Description: Cohort member's age in months at interview. Both first and second interviews if applicable

Population: All cohort members

Value Labels:

(-1) "Not applicable"

(-3) "Not asked in fieldwork stage"

(-8) "Not enough information"

Derivation description: (Interview year (n10intdate_year/n10intdate_year_r2) – birth year (1958)) * 12 + (Interview month (n10intdate_month/n10intdate_month_r2) – birth month (3))

ND10AGEINTY/ ND10AGEINTY_R2 – "(Derived) Age at interview (year part)" and **ND10AGEINTM/ ND10AGEINTM_R2 – "(Derived) Age at interview (month part)"**

Description: This is a more readable format of the above variables separated into years and months (to be used in combination).

Value labels:

(-1) "Not applicable"

(-3) "Not asked in fieldwork stage"

(-8) "Not enough information"

ND10SC_MODE – "(Derived) Mode for self-completion"

Description: How cohort members completed the section with sensitive questions

Value labels:

(-8) "Not enough information"

(-1) "Not applicable

(1) "Web – during video interview"

(2) "Web – after video interview"

(3) "Web – with interviewer during video interview"

(4) "Self-completion – during in person interview"

(5) "With interviewer – during in person interview"

(6) "Not completed"

Derivation description: Mode sensitive questions completed in, as recorded in paradata.

A1.2 Geographical variables

A series of geographical variables have been derived from the addresses at which participants were interviewed (see below)[12]. Country of interview, Region, Index of Multiple Deprivation Rank Decile and Urban/Rural indicator are available in the main file.

The remaining variables are provided in the dataset: ncds_age62_geographical_identifiers_2011 - available via secure access due to the potentially disclosive nature of the information.

[12] For participants who took part in both the main and the mop-up stages, the geographical variables have been derived using address provided during the mop-up interview. A small number will have moved home between taking part in the main stage and the mop-up and for these individuals the geographical variables will not relate to their address when they took part in the main stage. A future deposit will include separate geographical variables for each interview.

Variable name	Derived variable label	Dataset
ND10CNTRY	Country of interview	ncds10_age62_main_interview
ND10RGN	2020 Region of interview	ncds10_age62_main_interview
ND10IMDD	2019/2020 Index of Multiple Deprivation rank decile	ncds10_age62_main_interview
ND10RU11IND	2011 Urban/rural indicator	ncds10_age62_main_interview
n10osward	May 2023 Ward	ncds_age62_geographical_identifiers_2011
N10CASWARD	Census Area Statistics Ward	ncds_age62_geographical_identifiers_2011
N10OSLAUA	April 2023 Local Authority	ncds_age62_geographical_identifiers_2011
N10PCON	July 2024 Westminster Parliamentary Constituency	ncds_age62_geographical_identifiers_2011
N10IMD	IMD Overall Rank - Eng 2019, Sco 2020, Wal 2019, NI 2017	ncds_age62_geographical_identifiers_2011
N10WZ11	2011 Census Workplace Zone	ncds_age62_geographical_identifiers_2011
N10OA11	2011 Output Area	ncds_age62_geographical_identifiers_2011
N10LSOA11	2011 Lower Super Output Area	ncds_age62_geographical_identifiers_2011
N10MSOA11	2011 Middle Super Output Area	ncds_age62_geographical_identifiers_2011
N10OAC11	2011 Output Area Classification	ncds_age62_geographical_identifiers_2011
N10OA21	2021 Output Area	ncds_age62_geographical_identifiers_2021
N10LSOA21	2021 Lower Super Output Area	ncds_age62_geographical_identifiers_2021
N10MSOA21	2021 Middle Super Output Area	ncds_age62_geographical_identifiers_2021

A1.3 Household and family variables

Household derived variables were taken from the five person loops in the CAPI household grid: partner grid, two child grids (children reported at last sweep and additional children not previously mentioned), two other household members grids (other household members reported at last sweep and additional household members not previously mentioned). These cover all possible household members

at the time of interview. The five loops have subsequently been combined into the person grid dataset.

ND10HSIZE/ ND10HSIZE_R2 – "(Derived) HH Size"

Description: Number of people currently living in same household as cohort member (includes cohort member)

Population: All cohort members

Value Labels:

Greater than zero

Derivation description: Household size includes: the cohort member, partner, own children in household and all other household members (n10gslive=1).

ND10NUMCH/ ND10NUMCH_R2 – "(Derived) Number of children in HH"

Description: The total number of children in the household according to relationship and not age, including children of other household members.

Population: All cohort members

Value labels:

(-8) "No information"

Derivation description: Number of children (n10grtok=4,5,6,7,8,28) in household (n10gslive=1).

ND10NOCHH/ ND10NOCHH_R2 – "(Derived) Number of own children in HH"

Description: How many of the respondent's natural children live in the household at the time of interview.

Population: All cohort members

Value labels:

(-8) "No information"

Derivation description: Number of own (biological) children (n10grtok=4) in household (n10gslive=1).

ND10NPCHH/ ND10NPCHH_R2 – "(Derived) Number of children of current or previous partner in HH"

Description: How many children of the respondent's current or previous partner live in the household at the time of interview.

Population: All cohort members

Value labels:

(-8) "No information"

Derivation description: Number of current partner's (n10grtok=6), or previous partner's (n10grtok=7) children in household (n10gslive=1).

Number of natural children by age

Description: There are 7 variables that provide the counts of natural children of the respondent according to age group.

ND10NOC2A/ ND10NOC2A_R2 – "(Derived) Number of own children aged 0 to 2 in HH or absent"

ND10NOC4A/ ND10NOC4A_R2 – "(Derived) Number of own children aged 3 to 4 in HH or absent"

ND10NOC11A/ ND10NOC11A_R2 – (Derived) Number of own children aged 5 to 11 in HH or absent

ND10NOC15A/ ND10NOC15A_R2 – (Derived) Number of own children aged 12 to 15 in HH or absent

ND10NOC20A/ ND10NOC20A_R2 – (Derived) Number of own children aged 16 to 20 in HH or absent

ND10NOC30A/ N10NOC30A_R2 – (Derived) Number of own children aged 21 to 30 in HH or absent

ND10NOC31A/ ND10NOC31A_R2 – (Derived) Number of own children aged 31 and over in HH or absent

Population: All cohort members

Value labels:

Zero and above

Derivation description: The number of own children (n10grtok=4) reported within each age category (n10gage).

ND10NACAB/ N10NACAB_R2 – "(Derived) Number of absent children (including step-children etc)"

Description: Reported children of the respondent or other hh members who are not living in the household at the time of interview.

Population: All cohort members

Value labels:

Zero and above

Derivation description: The number of children (n10grtok=4,5,6,7,8,28) absent from household at interview (n10gslive=0).

ND10GCHLD – "(Derived) Whether has grandchildren (own or consider themselves to be)"

Description: Reported grandchildren of the respondent.

Population: All cohort members

Value labels:

Zero and above

Derivation description: Any grandchildren in household (n10grtok=20) or absent (n10gcnum > 0).

ND10NGCHLD – "(Derived) Total number of grandchildren (own or consider themselves to be)"

Description: Number of reported grandchildren of the respondent

Population: All cohort members

Value labels:

Zero and above

Derivation description: The number of grandchildren in household (n10grtok=20) or absent (n10gcnum).

A1.4 Relationships

ND10MS/ ND10MS_R2 – "(Derived) CM's legal marital status"

Description: The legal marital status of the respondent, both first and second interviews if applicable

Population: All cohort members

Value labels:

(-8) "No information"

(1) “Legally separated”

(2) “Married”

(3) “Divorced”

(4) “Widowed”

(5) “A Civil Partner”

(6) “A former Civil Partner”

(7) “A surviving Civil Partner”

(8) “Never married or in a Civil Partnership”

Derivation description: Combined updated marital status (n10hms/n10hms_r2 or status specified in relationship histories) and existing marital status if unchanged from previous sweep/response.

ND10COHAB/ ND10COHAB_R2 – “(Derived) Whether CM cohabiting as a couple”

Description: Whether the cohort member has a co-resident spouse or partner

Population: All cohort members

Value labels:

(-8) “No information”

(-1) “Not applicable”

(0) “No”

(1) “Yes”

Derivation description: Partners in relationship histories grid recorded as continuously living with CM.

A1.5 Housing

ND10TENURE – "(Derived) Housing Tenure"

Description: Tenure of cohort member's accommodation

Population: All cohort members

Value labels

(-8) "No information"

(-3) "Not asked in fieldwork stage"

(1) "Own outright"

(2) "Own, buying with help of mortgage/loan"

(3) "Part rent, part mortgage (shared equity)"

(4) "Rent it"

(5) "Live rent-free, incl. relatives/friends"

(6) "Squatting"

(7) "Other"

Derivation description: n10tenure (Current tenure) if tenure has changed or previous reported tenure if same as last sweep (n10tenck=1).

ND10RENTFROM – "(Derived) Who rents from"

Description: Owner of cohort member's rented housing

Population: Cohort members in rented housing

Value labels:

(-8) "No information"

(-3) "Not asked in fieldwork stage"

(-1) "Not applicable"

(1) “A Local Authority”

(2) “A Housing Association”

(3) “A Private landlord”

(4) “A Parent”

(5) “Someone else”

Derivation description: n10rentfrom (current rental provider) if tenure has changed, or previous reported rental provider if tenure is same as last sweep.

ND10WHOTEN – “(Derived) Whose name is accommodation held in”

Description: Name cohort member’s housing held in

Population: Homeowners and/or rent payers

Value labels:

(-8) “No information”

(-3) “Not asked in fieldwork stage”

(-1) “Not applicable”

(1) “CM's name only”

(2) “Both CM and partner's names”

(3) “Partner's name only”

(4) “CM and someone else's name's”

(5) “Someone else's name (including parents)”

Derivation description: Based on n10whoten, provided nd10tenure is between 1-5 so that the current accommodation can be in someone’s name.

ND10TIMAD - "(Derived) Time at current address (months)"

Description: Number of months cohort member has lived at current address

Population: All cohort members

Value Labels:

(-9) "Refused"

(-8) "Insufficient information"

(-3) "Not asked in fieldwork stage"

(-1) "Not applicable"

Derivation description: The date the cohort member moved into their current address (n10movinm, n10moviny) subtracted from the date of interview (n10intdate_month, n10intdate_year).

A1.6 Education

ND10ACHQ1 – "(Derived) Highest Academic Qualification CM reported in Age 62 Survey"

Description: This is highest academic qualification obtained since the last completed survey.

Population: All main-stage cohort members

Value labels:

(-8) "Not enough information"

(-3) "Not asked in fieldwork stage"

(-1) "Not applicable"

(0) "no academic qualification"

(1) "gcse d-e"

(2) "cses2-5, other Scottish quals"

(3) "gcse a-c, good o levels scot standards"

(4) “as levels or 1 a level”

(5) “2+ a levels, scot higher/6th”

(6) “diploma”

(7) “degree level”

(8) “higher degree”

Derivation description: If any qualifications reported (n10anyqual), the highest-level academic qualification reported in n10whatqual.

ND10ANVQ1 – “(Derived) Highest NVQ level from an academic qualification reported in Age 62 Survey”

Description: Highest academic qualification at the current sweep (as defined in nd10achq1) but categorised by their equivalent NVQ level rather than qualification type.

NVQ format (levels 1-5) for consistency with previous NCDS data.

Population: All main-stage cohort members

Value labels: (-3) “Not asked in fieldwork stage”

(-1) “not codeable”

(0) “none”

(1) “nvq1 level”

(2) “nvq2 level”

(3) “nvq3 level”

(4) “nvq4 level”

(5) “nvq5 level”

Derivation description: The academic qualification from n10whatqual attributed to the highest value of those listed in value labels

ND10ALVL1 - "(Derived): Highest qualification level from an academic qualification reported in the Age 62 Survey– 8 level version"

Description: Highest academic qualification at the current sweep (as defined in nd10achq1) but categorised by level rather than qualification type (9 categories – 0-8).

Population: All main-stage cohort members

Value Labels:

(-8) "Insufficient information"

(-3) "Not asked in fieldwork stage"

(-1) "Not applicable"

(0) "Entry Level"

(1) "Level 1"

(2) "Level 2"

(3) "Level 3"

(4) Level 4"

(5) "Level 5"

(6) "Level 6"

(7) "Level 7"

(8) "Level 8"

(95) "Other academic qualification"

(96) "No academic qualification"

Derivation description: The academic qualification from n10whatqual attributed to the highest value of those listed in value labels.

Nine category qualification levels for England, Wales and Northern Ireland What qualification levels mean: Overview - GOV.UK with equivalent levels for Scottish qualifications www.sqa.org.uk/sqa/64561.html

ND10VNVQ1 – "(Derived) Highest NVQ level from a Vocational Qualification reported in Age 62 Survey"

Description: Vocational qualification with the highest associated NVQ level obtained since last completed survey.

Used NVQ format (levels 1-5) for consistency with previous NCDS data.

Population: All main-stage cohort members

Value labels:

(-3) "Not asked in fieldwork stage"

(-1) "not codeable"

(0) "none"

(1) "nvq1 level"

(2) "nvq2 level"

(3) "nvq3 level"

(4) "nvq4 level"

(5) "nvq5 level"

Derivation description: All loops of n10voclev in the vocational qualifications grid. The qualification attributed to the highest value of those listed in value labels (n10voclev)

ND10VLVL1 - "(Derived): Highest qualification level from a vocational qualification reported in the Age 62 Survey– 8 level version"

Description: Vocational qualification with the highest associated education level obtained since last completed survey. (9 categories, 0-8)

Population: All main-stage cohort members

Value Labels:

(-8) "Insufficient information"

(-3) “Not asked in fieldwork stage”

(-1) “Not applicable”

(0) “Entry Level”

(1) “Level 1”

(2) “Level 2”

(3) “Level 3”

(4) Level 4”

(5) “Level 5”

(6) “Level 6”

(7) “Level 7”

(8) “Level 8”

(95) “Other vocational qualification”

(96) “No vocational qualification”

Derivation description: All loops of n10voclev in the vocational qualifications grid. The qualification attributed to the highest value of those listed in value labels (n10voclev)

Nine category qualification levels for England, Wales and Northern Ireland What qualification levels mean: Overview - GOV.UK with equivalent levels for Scottish qualifications www.sqa.org.uk/sqa/64561.html

ND10NVQ1 – “(Derived) Highest NVQ level from an Academic or Vocational Qual reported in Age 62 Survey”

Description: The associated NVQ level of the highest overall qualification obtained since last completed survey.

NVQ format (levels 1-5) for consistency with previous NCDS data.

Population: All cohort members

Value labels:

(-3) “Not asked in fieldwork stage”

(-1) “not codeable”

(0) “none”

(1) “nvq1 level”

(2) “nvq2 level”

(3) “nvq3 level”

(4) “nvq4 level”

(5) “nvq5 level”

Derivation description: The NVQ level of either the highest academic qualification (nd10anvq1) and vocational qualification (nd10vnvq1) depending on which is higher.

ND10LVL1- “(Derived): Highest qualification level from an academic or vocational qualification reported in the Age 62 Survey– 8 level version”

Description: Cohort member’s highest qualification level achieved from an academic or vocational qualification since last completed survey – self reported (9 categories, 0-8)

Population: All main-stage cohort members

Value Labels:

(-8) “Insufficient information”

(-3) “Not asked in fieldwork stage”

(-1) “Not applicable”

(0) “Entry Level”

(1) “Level 1”

(2) “Level 2”

(3) “Level 3”

(4) Level 4”

(5) “Level 5”

(6) “Level 6”

(7) “Level 7”

(8) “Level 8”

(95) “Other qualification”

(96) “No qualifications”

Derivation description: Level of either the highest academic qualification (nd10anvq1) and vocational qualification (nd10vnvq1) depending on which is higher.

Nine category qualification levels for England, Wales and Northern Ireland What qualification levels mean: Overview - GOV.UK with equivalent levels for Scottish qualifications www.sqa.org.uk/sqa/64561.html

ND10HACHQ – “(Derived) Highest Academic Qualification up to Age 62”

Description: The highest overall academic qualification of the cohort member.

Population: All cohort members

Value labels:

(-8) “Not enough information”

(-3) “Not asked in fieldwork stage”

(-1) “Not applicable”

(0) “no academic qualification”

(1) “gcse d-e”

(2) “cses2-5, other scottish quals”

(3) “gcse a-c, good o levels scot standards”

(4) “as levels or 1 a level”

(5) “2+ a levels, scot higher/6th”

(6) “diploma”

(7) “degree level”

(8) “higher degree”

Derivation description: The highest academic qualification at the current sweep (nd10achq1) compared to the highest level recorded from all previous sweeps

ND10HANVQ – “(Derived) Highest NVQ level from an academic qualification up to Age 62”

Description: The highest overall NVQ level from an academic qualification of the cohort member.

Population: All cohort members

Value labels:

(-3) “Not asked in fieldwork stage”

(-1) “not codeable”

(0) “none”

(1) “nvq1 level”

(2) “nvq2 level”

(3) “nvq3 level”

(4) “nvq4 level”

(5) “nvq5 level”

Derivation description: The highest NVQ level from an academic qualification (nd10anvq1) was compared to the highest academic qualification from all previous sweeps

ND10HVNVQ – "(Derived) Highest NVQ Level from a Vocational Qualification up to Age 62"

Description: The highest overall NVQ level from a vocational qualification of the cohort member.

Population: All cohort members

Value labels:

(-3) "Not asked in fieldwork stage"

(-1) "not codeable"

(0) "none"

(1) "nvq1 level"

(2) "nvq2 level"

(3) "nvq3 level"

(4) "nvq4 level"

(5) "nvq5 level"

Derivation description: The highest NVQ level from a vocational qualification (nd10vnvq1) compared to the highest vocational qualification from all previous sweeps

ND10HNVQ – "(Derived) Highest NVQ Level from an Academic or Vocational Qual up to Age 62"

Description: The highest qualification of any type the cohort member has ever obtained

Population: All cohort members

Value labels:

(-3) "Not asked in fieldwork stage"

(-1) "not codeable"

(0) “none”

(1) “nvq1 level”

(2) “nvq2 level”

(3) “nvq3 level”

(4) “nvq4 level”

(5) “nvq5 level”

Derivation Description: Either the highest academic qualification (nd10hanvq) or the highest vocational qualification (nd10hvnvq) ever obtained represented as an NVQ level.

ND10DEGP - “(Derived) Whether achieved first degree or higher”

Description: Whether cohort member has achieved a degree level qualification or higher

Population: All cohort members

Value labels:

(-8) “Insufficient information”

(0) “No degree”

(1) “First or higher degree”

Derivation Description: Based on derived academic qualifications recorded, if highest recorded level is degree or higher

A1.7 Health

ND10DISEQ/ND10DISEQ_R2 – “(Derived) Disability classification Equality act (2010)”

Description: The classification of whether or not the respondent is disabled according to the Equality act 2010

Population: All cohort members

Value labels:

(-8) “Not enough information”

(-3) “Not asked in fieldwork stage”

(0) “Not disabled (Equality act)”

(1) “Disabled (Equality Act)”

Derivation description: Classified disabled if any physical/mental health conditions lasting or expected to last 12 months (n10loil/n10loil_r2=1) and illnesses/conditions reduce ability to carry out day to day activities (n10lolm/n10lolm_r2=1,2).

ND10DISLS/ND10DISLS_R2 – “(Derived) Disability classification EU-SILC”

Description: Further classification of disability to ascertain the severity of the respondent’s illness/condition.

Population: All cohort members

Value labels:

(-8) “Not enough information”

(-3) “Not asked in fieldwork stage”

(0) “No EU-SILC long-standing health condition”

(1) “EU-SILC classification to some extent”

(2) “EU-SILC classification severely hampered”

Derivation description: Longstanding illness and ability to carry out day to day activities is reduced a little for 6 months or more (n10loil/n10loil_r2=1 AND (n10lolp/n10lolp_r2=2,3) AND (n10lolm/n10lolm_r2=2)), CM is disabled/hampered to some extent. Longstanding illness and the ability to carry out day to day activities is reduced a lot for 6 months or more (n10loil/n10loil_r2=1 AND (n10lolp/n10lolp_r2=2,3) AND (n10lolm/n10lolm_r2=1)) CM is categorised as being disabled, severely hampered.

For more information on disability classification see ONS document 'Long-lasting Health Conditions and Illnesses; Impairments and Disability': www.ons.gov.uk/ons/guide-method/harmonisation/primary-set-of-harmonised-concepts-and-questions/index.html

ND10HGHTM – "(Derived) Self-reported height in metres"

Description: Most recent self-reported height measurement in metres and centimetres

Population: All cohort members

Value labels:

(-8) "No information"

(-3) "Not asked in fieldwork stage"

Derivation description: Self-reported height provided as metres and centimetres (n10htmetres, n10htcentem) combined with height in feet and inches converted using the formula (n10htfeet*12)+n10htinches)/39.370..

ND10NHGHTM – "(Derived) Nurse height in metres"

Description: Height according to nurse in metres and centimetres

Population: All cohort members who participated in nurse module

Value labels:

(-8) "No information"

(-3) "Not asked in fieldwork stage"

Derivation description: The most recent measured height recording from previous biomedical survey, or self-reported height (nd10hghtm) if correct according to the nurse (n10htchk=1), otherwise uses the nurse estimate (n10htnestmet, n10htnestcm, n10htnestft, n10htnestin) as metres and cm using the same formula as above

ND10WGHTK/ND10WGHTK_R2 – "(Derived) Self-reported weight in kilograms"

Description: Current self-reported weight in kilograms

Population: All cohort members

Value labels:

(-8) "No information"

(-3) "Not asked in fieldwork stage"

Derivation description: Self-reported weight in kilograms (n10wtkilos) combined with weight in stones and pounds, converted to kg with the formula (n10wtstones*14+n10wtpounds)*0.453592.

ND10BMI/ND10BMI_R2 – "(Derived) Body mass index (based on self-reported data)"

Description: Body mass index of cohort member, based on self-reported height and weight

Population: All cohort members

Value labels:

(-8) "Not enough information"

(-3) "Not asked in fieldwork stage"

Derivation description: Calculated with the derived self-reported height and weight measurements using the formula nd10wghtk/(nd10hghtm* nd10hghtm).

ND10BMIC/ND10BMIC_R2 – "(Derived) Body mass index – classification (based on self-reported data)"

Description: Self-reported BMI values (ND10BMI/ND10BMI_R2) separated into 5 weight range categories defined by the same classification as that of the NHS.

Population: All cohort members

Value labels:

(-8) “Not enough information”

(-3) “Not asked in fieldwork stage”

(0) “Underweight (below 18.5)”

(1) “Healthy weight range (18.5-24.9)”

(2) “Overweight (25-29.9)”

(3) “Obese (30-39.9)”

(4) “Morbidly obese (Over 40)”

ND10NBMI – “(Derived) Body mass index (based on nurse measured data)”

Description: Body mass index of cohort member, based on nurse estimated height and measured weight

Population: All cohort members who participated in nurse module

Value labels:

(-8) “Not enough information”

(-3) “Not asked in fieldwork stage”

Derivation description: Calculated with the derived nurse estimated height and weight measurements using the formula (n10mweight or n10wgtonly)/(nd10nhghtm* nd10nhghtm).

ND10NBMIC – “(Derived) Body mass index – classification (based on nurse measured data)”

Description: Nurse measured BMI values (ND10NBMI) separated into 5 weight range categories defined by the same classification as that of the NHS.

Population: All cohort members who participated in nurse module

Value labels:

(-8) “Not enough information”

(-3) “Not asked in fieldwork stage”

(0) “Underweight (below 18.5)”

(1) “Healthy weight range (18.5-24.9)”

(2) “Overweight (25-29.9)”

(3) “Obese (30-39.9)”

(4) “Morbidly obese (Over 40)”

ND10SMOKE – “(Derived) Smoking habits”

Description: Typical number of cigarettes smoked daily

Population: All cohort members

Value labels:

(-8) “Not enough information”

(-3) “Not asked in fieldwork stage”

(-1) “Not applicable”

(0) “Never smoked”

(1) “Ex smoker”

(2) “Occasional smoker”

(3) “Up to 10 a day”

(4) “11 to 20 a day”

(5) “More than 20 a day”

(6) “Daily but frequency not stated”

Derivation description: For current smokers uses the number of cigarettes smoked daily (n10nofcigs) as a categorical range. Non-smokers and ex-smokers derived from n10smoking

ND10AUDIT – "(Derived) Total AUDIT-PC score"

Description: Cohort member's score on the Alcohol Use Disorders Identification Test

Population: Cohort members that drink alcohol who completed the paper self-completion

Value labels:

(-8) "Not enough information"

(-3) "Not asked in fieldwork stage"

(-1) "NA – does not drink alcohol"

Derivation description: Total score of 5 questions (n10ylq28, n10ylq2 9, n10ylq30a, n10ylq30b, n10ylq31), each on a 0-4 point scale. Scores of 5 or above indicate higher risk drinking

ND10AUDG – "(Derived) AUDIT-PC Group"

Description: A grouped version of nd10audit, according to the definition of low and high-risk drinking

Population: Cohort members that drink alcohol who completed the paper self-completion

Value labels:

(-8) "Not enough information"

(-3) "Not asked in fieldwork stage"

(-1) "NA – does not drink alcohol"

(1) "Unproblematic drinking (0-4)"

(2) “Increasing or high risk drinking (5+)”

A1.8 Mental Health/Wellbeing

ND10WEMWB – “(Derived) Warwick Edinburgh Mental Well-Being Scale”

Description: A short (7-item) version of the scale designed to represent the respondent’s mental well-being. Score ranges from 7-35 where higher scores indicate higher positive mental wellbeing

Population: All cohort members who completed the paper self-completion

Value labels:

(-8) “Not enough information”

(-3) “Not asked in fieldwork stage”

Derivation description: The sum of 7 items, each on a 5-point scale (n10ylq18a+ n10ylq18b+ n10ylq18c+ n10ylq18d+ n10ylq18e+ n10ylq18f+ n10ylq18g)

ND10MAL/ND10MAL_R2 – “(Derived) Total Malaise score (9 questions)”

Description: Total malaise score from a 9-item short version of the malaise inventory. Items cover negative emotions and physical response and are coded so high malaise scores always relate to affirmative responses.

Population: All cohort members who completed the self-completion

Value labels:

(-8) “Not enough information”

(-3) “Not asked in fieldwork stage”

Derivation description: The score is the total number of ‘yes’ responses (n10mal02, n10mal03, n10mal05, n10mal09, n10mal12, n10mal14, n10mal16, n10mal20, n10mal21) with higher scores corresponding to higher malaise. Cases were excluded if the number of items without a response accounted for scores being under 4 (the prerequisite score for high malaise).

ND10MALG/ND10MALG_R2 – "(Derived) Total Malaise score – grouped"

Description: A grouped version of nd10mal, according to the definition of low and high malaise. High malaise requires at least four affirmative responses.

Population: All cohort members who completed the self-completion

Value labels:

(-8) "Not enough information"

(-3) "Not asked in fieldwork stage"

(1) "Low malaise (0-3)"

(2) "High malaise (4+)"

Short-form health survey

A subset of three of the scales based on 17 items from the SF-36 have been derived, with all scores ranging between 0 and 100. For these derived variables higher scores indicate better health in the area described. One response for the associated questions is sufficient for a score to be created for these variables.

ND10PHHE – "(Derived) SF-36 Physical functioning score"

Description: Physical health score coded from 10 items. Higher scores represent better functioning across a range of physical activities (e.g. running, dressing, climbing stairs).

Population: All cohort members who completed the paper self-completion

Value labels:

(-8) "Not enough information"

(-3) "Not asked in fieldwork stage"

Derivation description: Recoded 10 items in scale (n10ylq9a, n10ylq9b, n10ylq9c, n10ylq9d, n10ylq9e, n10ylq9f, n10ylq9g, n10ylq9h, n10ylq9i, n10ylq9j) to 1=0, 2=50, 3=100. Physical functioning scores are the mean of the total number of questions answered (between 1 and 10 responses).

ND10EMWB – "(Derived) SF-36 Emotional Well-Being score"

Description: Also known as the MHI-5, a mental health score coded from 5 items. Higher scores result from an overall more positive emotional outlook during the four weeks prior to interview.

Population: All cohort members who completed the paper self-completion

Value labels:

(-8) "Not enough information"

(-3) "Not asked in fieldwork stage"

Derivation description: Three items (n10ylq16a, n10ylq16b, n10ylq16d) are coded so 1=0, 2=20, 3=40, 4=60, 5=80, 6=100, while the other two items (n10ylq16c, n10ylq16e) are coded so 1=100, 2=80, 3=60, 4=40, 5=20, 6=0. Emotional well-being scores are the mean of the total number of questions answered (between 1 and 5 responses).

ND10PAIN – "(Derived) SF-36 Pain score"

Description: Physical health score coded from 2 items. Lower scores show an increase in pain and its impact on normal life.

Population: All cohort members who completed the paper self-completion

Value labels:

(-8) "Not enough information"

(-3) "Not asked in fieldwork stage"

Derivation description: n10ylq10 is a 6-point scale recoded so 1=100, 2=80, 3=60, 4=40, 5=20, 6=0. n10ylq11 is on a 5-point scale recoded so 1=100, 2=75, 3=50, 4=25, 5=0. Pain scores are the mean of the total number of questions answered (1 or 2 responses).

ND10LONELINESS/ND10LONELINESS_R2 – "(Derived) UCLA Loneliness 3-item"

Description: Measurement of loneliness

Population: All cohort members who completed the self-completion

Value labels:

(-8) "Not enough information"

(-3) "Not asked in fieldwork stage"

Derivation description: Addition of the scores of three questions (n10lonelya, n10lonelyb, n10lonelyc). Only computed if all questions have valid responses.

ND10GAD2/ND10GAD2_R2 – "(Derived) Generalised Anxiety Disorder 2-item"

Description: Measurement of the frequency of feeling nervous, anxious or on edge.

Population: All cohort members

Value labels:

(-8) "Not enough information"

(-3) "Not asked in fieldwork stage"

Derivation description: Addition of the scores of two questions (n10gad2phq2a, n10gad2phq2b). Only computed if all questions have valid responses.

ND10PHQ2/ND10PHQ2_R2 – "(Derived) Patient Health Questionnaire 2-item"

Description: Measurement of the frequency of depressed mood over the past 2 weeks

Population: All cohort members

Value labels:

(-8) “Not enough information”

(-3) “Not asked in fieldwork stage”

Derivation description: Addition of the scores of two questions (n10gad2phq2c, n10gad2phq2d). Only computed if all questions have valid responses.

ND10CASP6 - “(Derived) CASP Quality of Life score 6-item"

Description: The 6-item version of the CASP quality of life scale

Population: All cohort members who completed the paper self-completion

Value labels:

(-8) “Not enough information”

(-3) “Not asked in fieldwork stage”

Derivation description: Sum of all 6 items (n10ylq19a+ n10ylq19b+ n10ylq19c+ n10ylq19d+ n10ylq19e+ n10ylq19f). Only computed if all questions have valid responses

A1.9 Finance

ND10BENE - “(Derived) Whether cohort member or partner receives any benefits”

Description: Whether cohort member or partner receives any benefits

Population: All cohort members

Value labels:

(-9) “Refused”

(-8) “Not known”

(-3) “Not asked in fieldwork stage”

(1) “Receives benefits”

(2) “Does not receive benefits”

Derivation description: Cohort members are categorised as receiving benefits if they reported receiving universal credit (n10uncr) or any of the benefits mentioned in n10bent01-n10bent11.

ND10FINLIT3 - “(Derived) Financial Literacy- All three questions correct”

Description: Whether the cohort member correctly answered all financial literacy questions

Population: Main stage respondents

Value labels:

(1) “Yes”

(2) “No”

(-1) “Not applicable”

(-3) “Not asked in fieldwork stage”

(-8) “Not enough information”

Derivation description: Yes if n10finlit2=2, n10finlit3=102 and n10finlit4=1. No if any responses were incorrect or refused

ND10FINLITC - “(Derived) Financial Literacy- Number of questions correct”

Description: Total number of financial literacy questions correctly answered

Population: Main stage respondents

Value labels:

(0) “None”

(1) “1”

(2) “2”

(3) “3”

(-1) “Not applicable”

(-3) “Not asked in fieldwork stage”

(-8) “Not enough information”

Derivation description: Sum of correct values (n10finlit2=2, n10finlit3=102 and n10finlit4=1)

ND10FINLITA - "(Derived) Financial Literacy- Number of questions attempted"

Description: Total number of financial literacy questions attempted

Population: Main stage respondents

Value labels:

(0) "None"

(1) "1"

(2) "2"

(3) "3"

(-1) "Not applicable"

(-3) "Not asked in fieldwork stage"

(-8) "Not enough information"

Derivation description: Sum of non-missing values (n10finlit2, n10finlit3 and n10finlit4). Don't know/refusal do not classify as attempts

ND10FINLITR - "(Derived) Financial Literacy- At least one question refused"

Description: Whether CM refused any financial literacy questions

Population: Main stage respondents

Value labels:

(1) "Yes"

(2) "No"

(-1) "Not applicable"

(-3) "Not asked in fieldwork stage"

(-8) "Not enough information"

Derivation description: Yes if any (n10finlit2=-9, n10finlit3=-9 or n10finlit4=-9)

A1.10 Activities and Employment

ND10NSSEC7 – "(Derived) NS-SEC 7 analytic classes (CMs current job)"

Description: Cohort member's current job in seven category NS-SEC format

Population: All cohort members

Value labels:

(-9) "Not classified"

(-8) "Never worked and long-term unemployed"

(-1) "Not applicable"

(1) "Higher managerial and professional occupations"

(2) "Lower managerial and professional occupations"

(3) "Intermediate occupations"

(4) "Small employers and own account workers"

(5) "Lower supervisory and technical occupations"

(6) "Semi-routine occupations"

(7) "Routine occupations"

Derivation description: Derived from full NS-SEC analytic sub-classes (n10cjnss8) into 7 classes

ND10NSSEC5 – "(Derived) NS-SEC 5 analytic classes (CMs current job)"

Description: Cohort member's current job in five category NS-SEC format

Population: All cohort members

Value labels:

(-9) "Not classified"

(-8) "Never worked and long-term unemployed"

(-1) “Not applicable”

(1) “Managerial and professional occupations”

(2) “Intermediate occupations”

(3) “Small employers and own account workers”

(4) “Lower supervisory and technical occupations”

(5) “Semi-routine and routine occupations”

Derivation description: Further simplified from NS-SEC analytic sub-classes (n10cjnss8) into 5 classes.

ND10WRKCP - “(Derived) Combined labour market status”

Description: Combined employment status for both CM and partner

Population: All cohort members

Value labels:

(-9) “Refused”

(-8) “Insufficient information”

(-1) “Not applicable”

(1) “Both in work”

(2) “CM in work; partner not in work”

(3) “CM not in work; partner in work”

(4) “Both not in work”

(5) “CM in work; no partner”

(6) “CM not in work; no partner”

(7) “CM status unknown; partner in work”

(8) “CM status unknown; partner not in work”

(9) “CM in work; partner status unknown”

(10) “CM not in work; partner status unknown”

(11) “CM status unknown; no partner”

(12) “CM status and partner status unknown”

Appendix 2: Response Models

Table A2: Response Models by Interview Type. Logistic regression models based upon twenty multiply imputed datasets pooled using Rubin's (1987) rules.

	Variable	Main Interview or Mop-Up	Self-Completion (Your Life Now)	Self-Completion (Childhood)	Health Visit	Blood Sample
Sex (Ref: Male)	Female	0.69 (0.62, 0.77)	0.85 (0.77, 0.94)	0.87 (0.79, 0.95)	0.71 (0.64, 0.78)	0.77 (0.71, 0.85)
Social Class @ Birth (Ref: Professional or Managerial)	Intermediate	0.94 (0.82, 1.09)	0.94 (0.83, 1.07)	0.94 (0.82, 1.06)	1 (0.89, 1.13)	0.99 (0.88, 1.11)
	Partly-Skilled or Unskilled	0.89 (0.75, 1.05)	0.84 (0.72, 0.98)	0.87 (0.74, 1.01)	0.91 (0.78, 1.06)	0.92 (0.8, 1.06)
	Rooms per Person @ Age 0y	1.02 (0.96, 1.07)	1.01 (0.96, 1.07)	1.03 (0.98, 1.08)	1 (0.95, 1.06)	1.04 (0.99, 1.09)
	Cognitive Ability @ Age 7y	0.99 (0.93, 1.06)	1 (0.94, 1.06)	1.04 (0.98, 1.1)	0.98 (0.92, 1.05)	0.99 (0.93, 1.04)
	Externalising @ Age 16y	1.05 (0.98, 1.13)	0.97 (0.91, 1.04)	0.96 (0.9, 1.03)	1.03 (0.96, 1.1)	1.02 (0.96, 1.1)

	Variable	Main Interview or Mop-Up	Self-Completion (Your Life Now)	Self-Completion (Childhood)	Health Visit	Blood Sample
	Internalising @ Age 16y	1.08 (1, 1.16)	1.07 (1, 1.13)	1.07 (1, 1.14)	1.05 (0.98, 1.12)	1.04 (0.98, 1.11)
Educational Qualifications @ Age 42y (Ref: None)	NVQ Level 1-3	0.92 (0.79, 1.07)	1.03 (0.88, 1.2)	1 (0.86, 1.17)	1.04 (0.89, 1.22)	1.01 (0.86, 1.19)
	NVQ Level 4-5	1.06 (0.87, 1.28)	1.18 (0.98, 1.42)	1.08 (0.9, 1.3)	1.12 (0.93, 1.35)	1.09 (0.91, 1.32)
Organisation Membership @ Age 42y (Ref: No)	Yes	1.21 (1.06, 1.37)	1.25 (1.11, 1.4)	1.23 (1.1, 1.38)	1.19 (1.07, 1.33)	1.16 (1.05, 1.28)
Union Membership @ Age 42y (Ref: No)	Yes	1.03 (0.91, 1.18)	1.01 (0.89, 1.14)	1.03 (0.92, 1.16)	0.96 (0.86, 1.07)	0.96 (0.86, 1.06)
Voted in 1997 General Election (Ref: No)	Yes	1.02 (0.89, 1.15)	1 (0.88, 1.12)	1 (0.89, 1.13)	1.01 (0.9, 1.13)	0.95 (0.85, 1.06)
Consent to Biological Samples	Bio Sweep	1.1 (0.84, 1.44)	1.31 (1.02, 1.68)	1.36 (1.06, 1.74)	1.26 (0.99, 1.61)	0.74 (0.57, 0.96)

	Variable	Main Interview or Mop-Up	Self-Completion (Your Life Now)	Self-Completion (Childhood)	Health Visit	Blood Sample
@ Age 44y (Ref: No Consent)	Blood Sample	1.41 (0.98, 2.02)	1.86 (1.35, 2.58)	1.72 (1.26, 2.35)	1.77 (1.32, 2.37)	1.74 (1.31, 2.3)
	DNA Extraction	1.5 (1.3, 1.74)	1.71 (1.48, 1.97)	1.64 (1.42, 1.9)	2.4 (2.08, 2.77)	2.38 (2.05, 2.76)
	Malaise @ Age 50y	1.02 (0.98, 1.05)	1.01 (0.98, 1.04)	1.01 (0.98, 1.04)	0.98 (0.95, 1.01)	1 (0.97, 1.02)
Self-rated Health @ Age 50y (Ref: Excellent or Very Good)	Good	1.04 (0.91, 1.2)	1.02 (0.9, 1.15)	1.03 (0.92, 1.16)	0.94 (0.84, 1.05)	0.85 (0.77, 0.95)
	Fair or Poor	1.21 (1.01, 1.45)	1.04 (0.88, 1.22)	1.05 (0.9, 1.24)	0.98 (0.84, 1.14)	0.8 (0.69, 0.93)
	BMI @ Age 50y	1.01 (0.99, 1.02)	1 (0.99, 1.01)	1 (0.99, 1.01)	0.99 (0.98, 1)	0.97 (0.96, 0.98)
Smoking Status @ Age 50y (Ref: Never)	Former	0.97 (0.85, 1.1)	0.98 (0.86, 1.11)	0.95 (0.84, 1.07)	0.9 (0.81, 1.01)	0.93 (0.84, 1.04)
	Current	0.97 (0.83, 1.13)	0.91 (0.79, 1.05)	0.84 (0.73, 0.95)	0.85 (0.75, 0.97)	0.79 (0.69, 0.9)
Income Quintile @ Age 50y (Ref: 1)	2	1.19 (0.95, 1.47)	1.21 (1.01, 1.46)	1.2 (1, 1.44)	1.26 (1.05, 1.53)	1.13 (0.96, 1.33)

	Variable	Main Interview or Mop-Up	Self-Completion (Your Life Now)	Self-Completion (Childhood)	Health Visit	Blood Sample
	3	1.32 (1.05, 1.67)	1.34 (1.08, 1.65)	1.17 (0.98, 1.41)	1.36 (1.14, 1.63)	1.12 (0.95, 1.33)
	4	1.16 (0.93, 1.43)	1.26 (1.03, 1.54)	1.24 (1.05, 1.47)	1.29 (1.08, 1.54)	1.04 (0.87, 1.24)
	5	1.12 (0.89, 1.41)	1.23 (1.01, 1.51)	1.11 (0.92, 1.33)	1.17 (0.95, 1.45)	1.1 (0.92, 1.32)
Employed @ Age 50y (Ref: Currently employed)	Not currently employed	0.79 (0.66, 0.94)	0.83 (0.7, 0.98)	0.84 (0.72, 0.98)	0.91 (0.78, 1.05)	0.82 (0.7, 0.96)
Marital Status @ Age 50y (Ref: Single & Never Married)	Married or Civil Partnered	1 (0.84, 1.19)	1.19 (1, 1.41)	1.22 (1.03, 1.45)	1.06 (0.9, 1.26)	1.07 (0.92, 1.24)
	Separated, Divorced or Widowed	1.09 (0.89, 1.33)	1.17 (0.98, 1.4)	1.19 (0.98, 1.44)	1.1 (0.91, 1.33)	1.08 (0.91, 1.29)
Internet Access @ Age 50y (Ref: Yes)	No	1.09 (0.93, 1.27)	1.15 (1, 1.33)	1.12 (0.98, 1.29)	1.03 (0.89, 1.17)	0.96 (0.84, 1.09)
	Somewhat	1.09 (0.87, 1.37)	1.08 (0.88, 1.33)	1.03 (0.83, 1.27)	1.04 (0.84, 1.28)	1.09 (0.89, 1.33)

	Variable	Main Interview or Mop-Up	Self-Completion (Your Life Now)	Self-Completion (Childhood)	Health Visit	Blood Sample
Social Listen @ Age 50y (Ref: A Little or Not at All)	A Great Deal	1.1 (0.91, 1.33)	1.06 (0.89, 1.26)	0.97 (0.8, 1.17)	0.99 (0.82, 1.19)	1.05 (0.88, 1.25)
Social Visits @ Age 50y (Ref: Never)	Fairly Frequently	1.1 (0.94, 1.28)	0.99 (0.86, 1.13)	1.06 (0.93, 1.22)	1.14 (1, 1.3)	1.09 (0.96, 1.24)
	Very Frequently	1.06 (0.89, 1.26)	0.99 (0.85, 1.16)	1.09 (0.94, 1.27)	1.11 (0.96, 1.28)	1.08 (0.94, 1.24)
Social Trust @ Age 50y (Ref: Most people can be trusted)	Can't be too careful	1 (0.88, 1.13)	0.87 (0.78, 0.97)	0.89 (0.79, 0.99)	0.82 (0.74, 0.91)	0.86 (0.77, 0.95)
	Other/depends	1.19 (0.96, 1.48)	0.95 (0.78, 1.15)	0.95 (0.79, 1.15)	0.95 (0.79, 1.14)	0.97 (0.82, 1.15)
	# Previous Nonresponses	0.64 (0.62, 0.66)	0.69 (0.67, 0.71)	0.7 (0.68, 0.72)	0.7 (0.68, 0.73)	0.74 (0.71, 0.76)
	# Nonresponses (COVID)	0.34 (0.32, 0.35)	0.45 (0.43, 0.46)	0.46 (0.44, 0.48)	0.5 (0.48, 0.52)	0.57 (0.55, 0.59)

Appendix 3: Weights Performance

Figure A3.1: Sample proportions for categorical predictors included in response model, by variable category, population (target or respondents) and whether sample is weighted or unweighted.

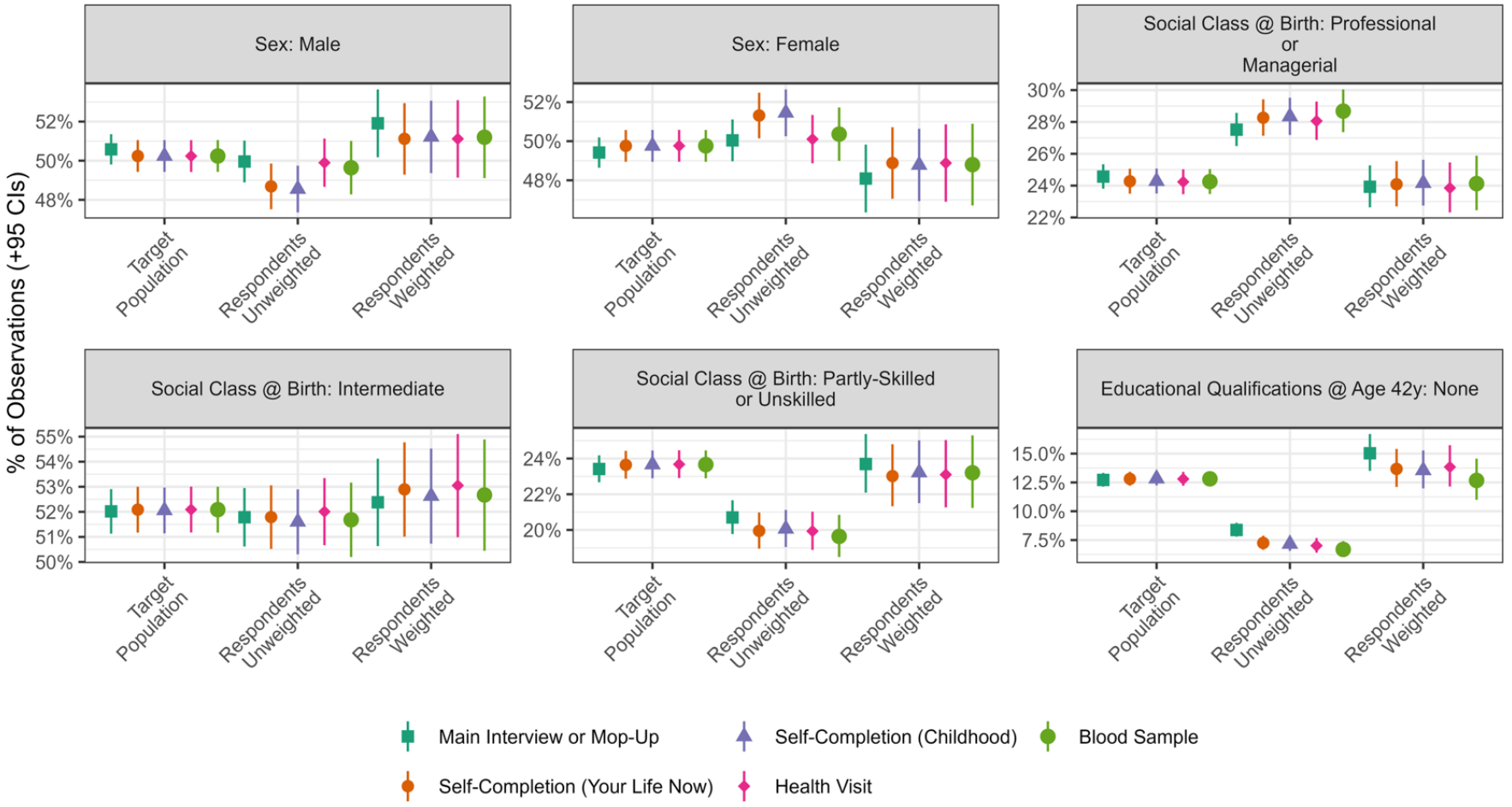


Figure A3.2: Sample proportions for categorical predictors included in response model, by variable category, population (target or respondents) and whether sample is weighted or unweighted.

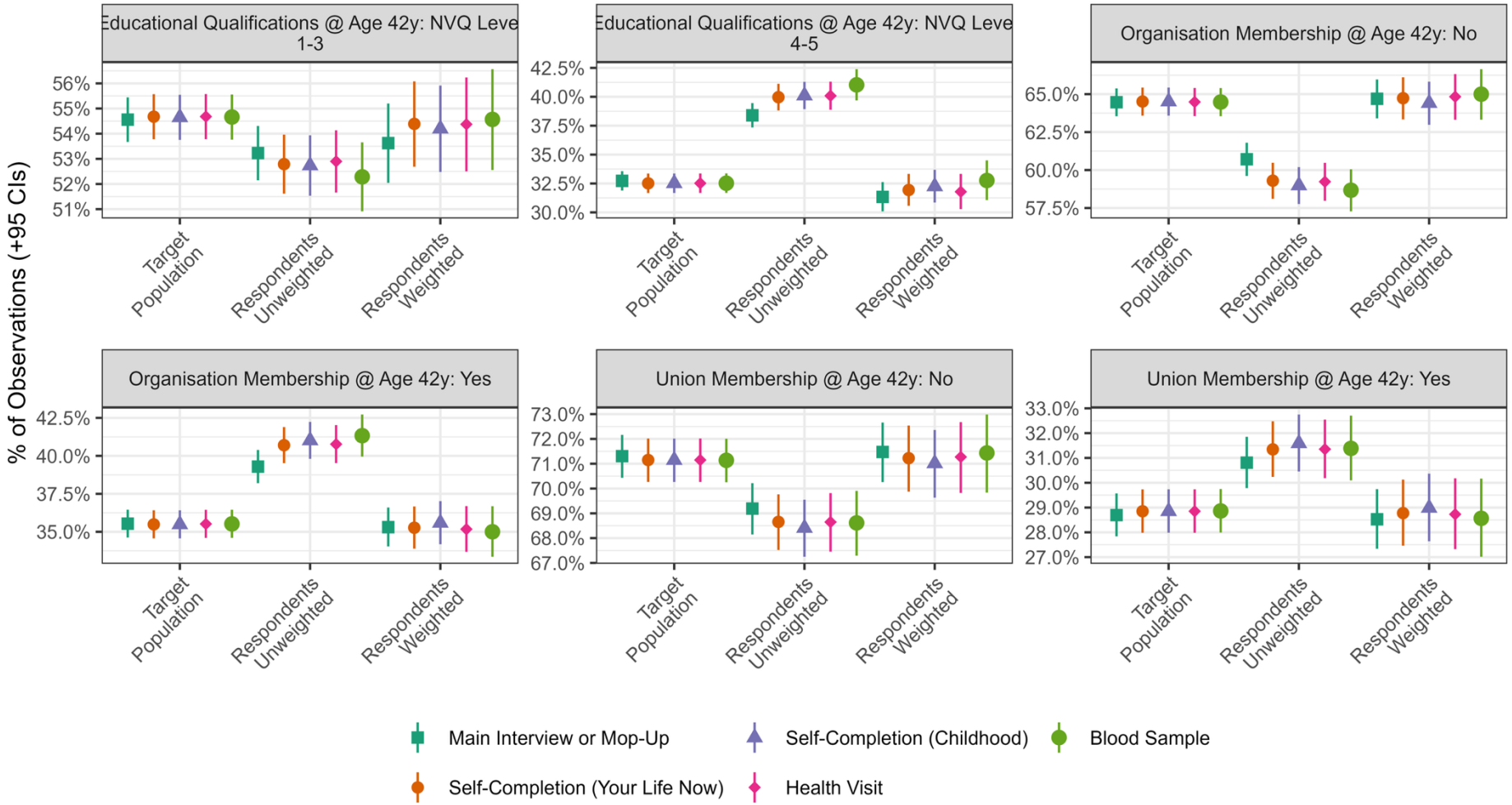


Figure A3.3: Sample proportions for categorical predictors included in response model, by variable category, population (target or respondents) and whether sample is weighted or unweighted.

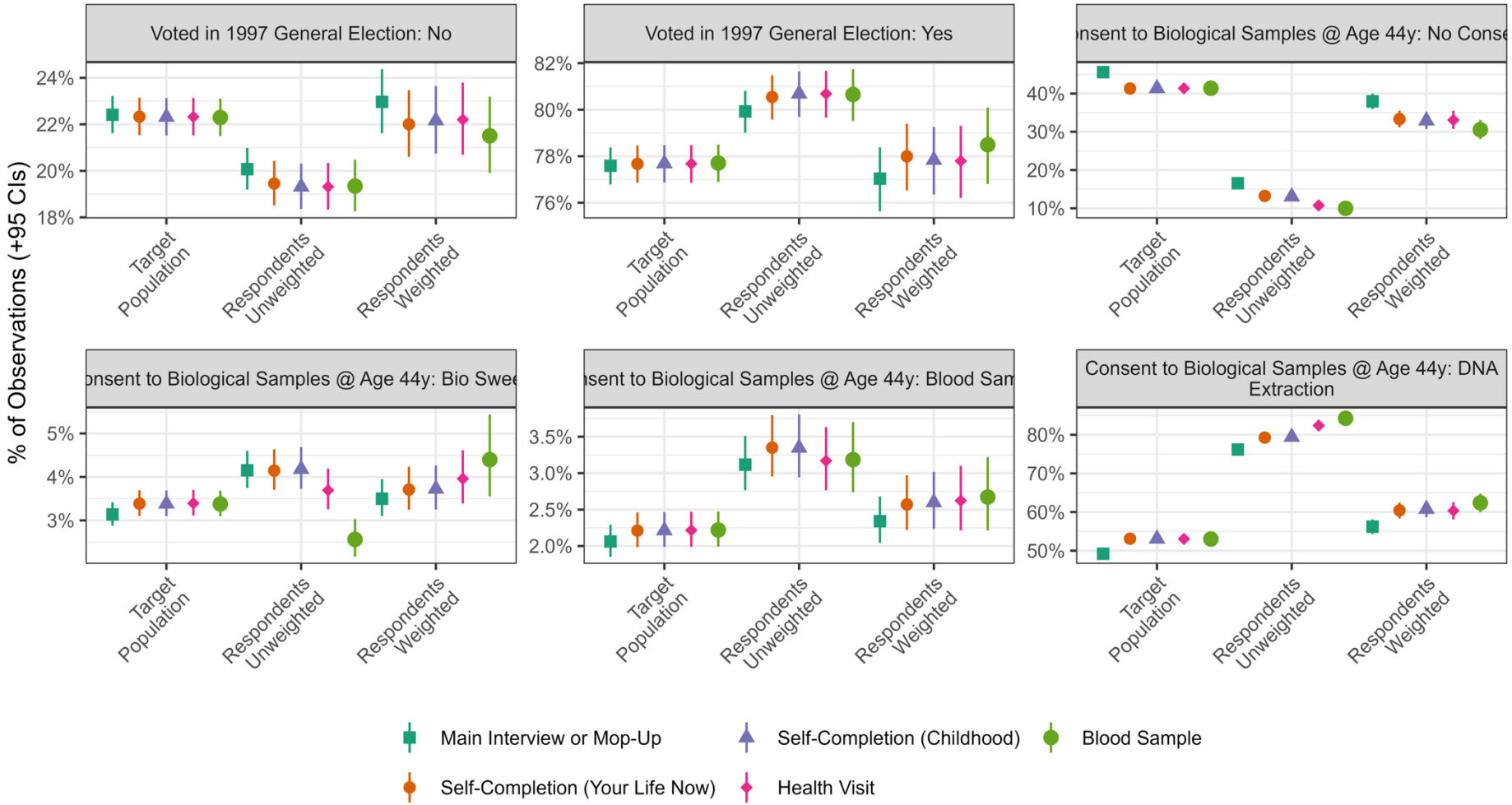


Figure A3.4: Sample proportions for categorical predictors included in response model, by variable category, population (target or respondents) and whether sample is weighted or unweighted.

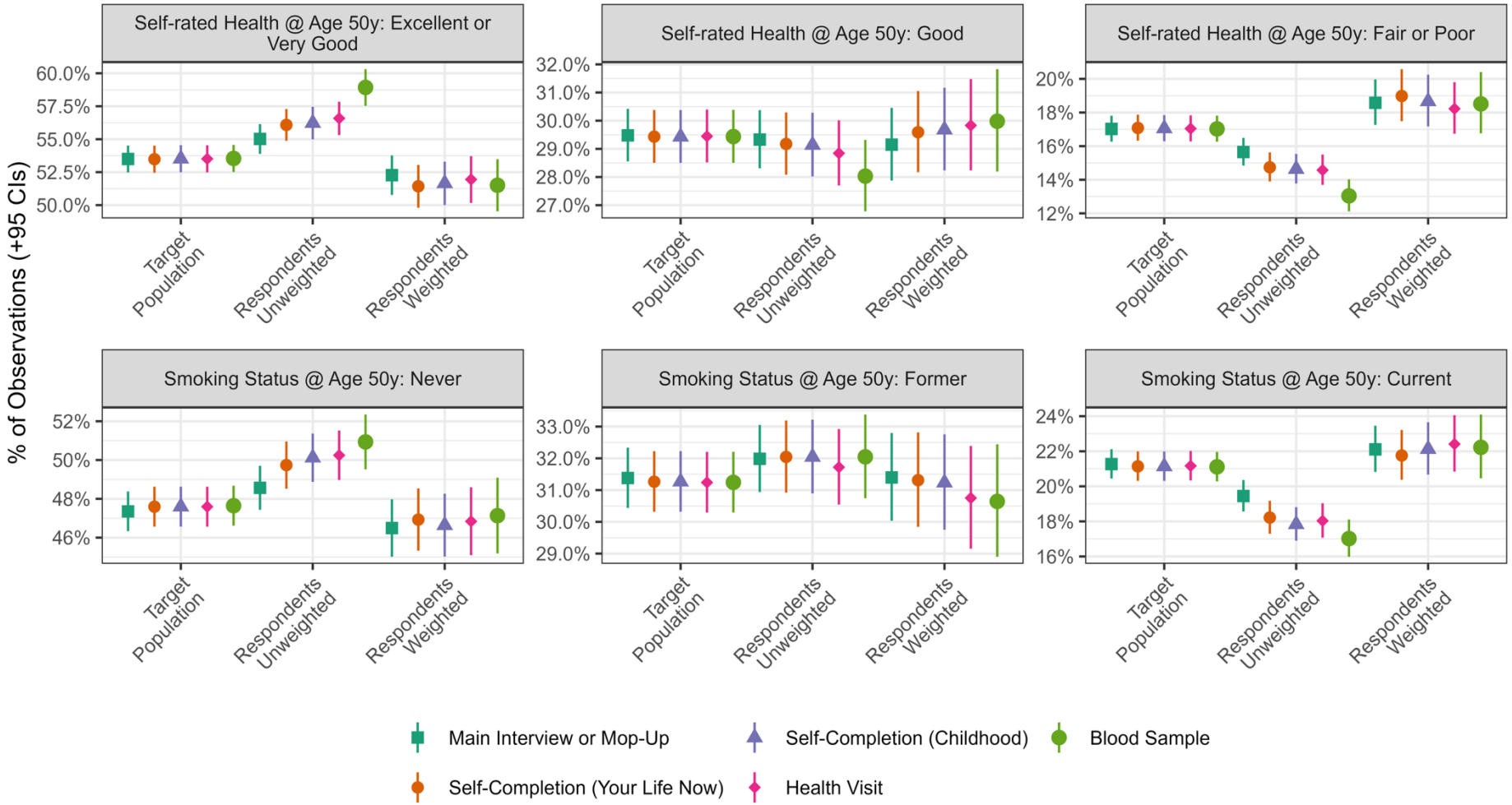


Figure A3.5: Sample proportions for categorical predictors included in response model, by variable category, population (target or respondents) and whether sample is weighted or unweighted.

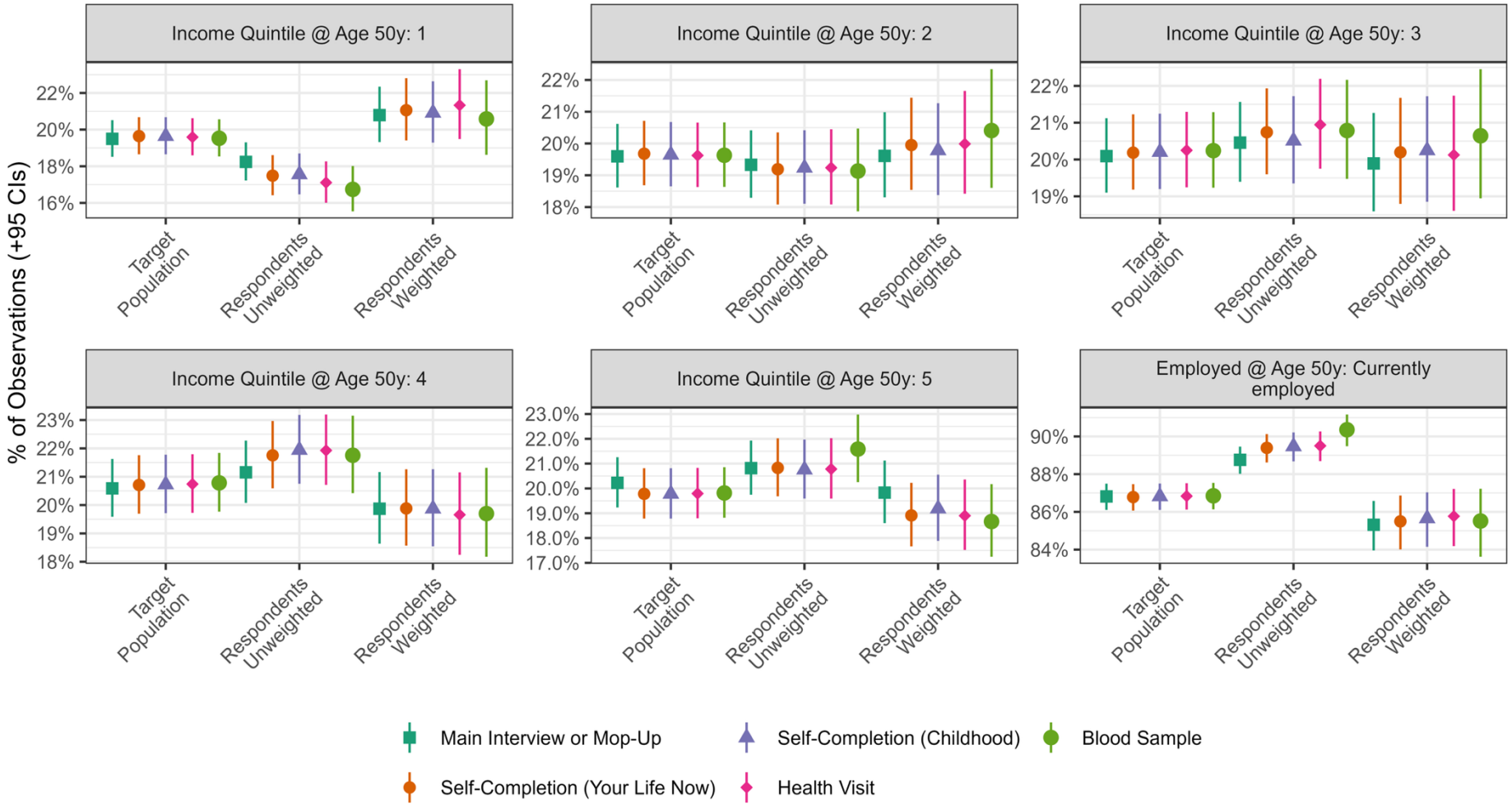


Figure A3.6: Sample proportions for categorical predictors included in response model, by variable category, population (target or respondents) and whether sample is weighted or unweighted.

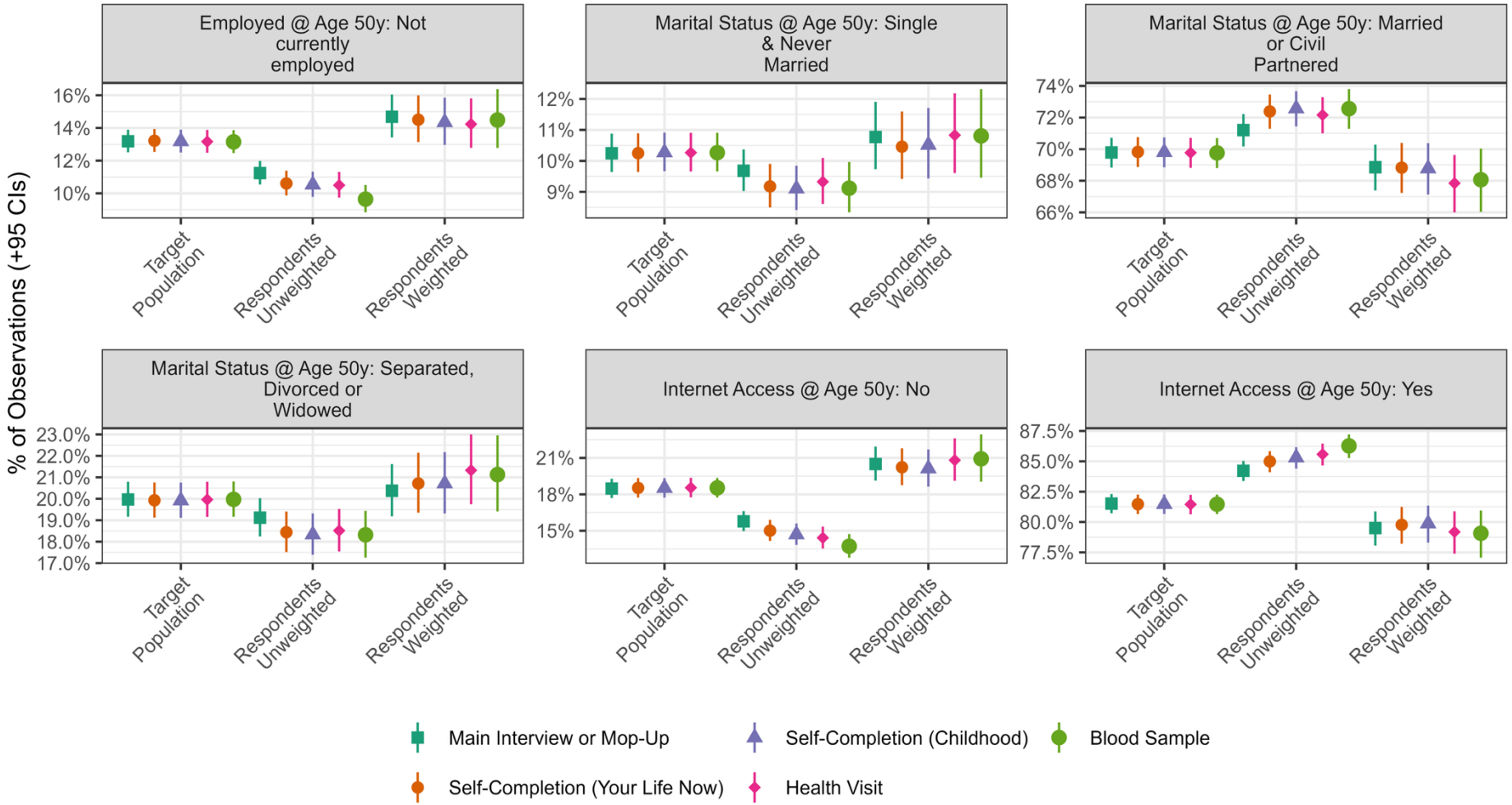


Figure A3.7: Sample proportions for categorical predictors included in response model, by variable category, population (target or respondents) and whether sample is weighted or unweighted.

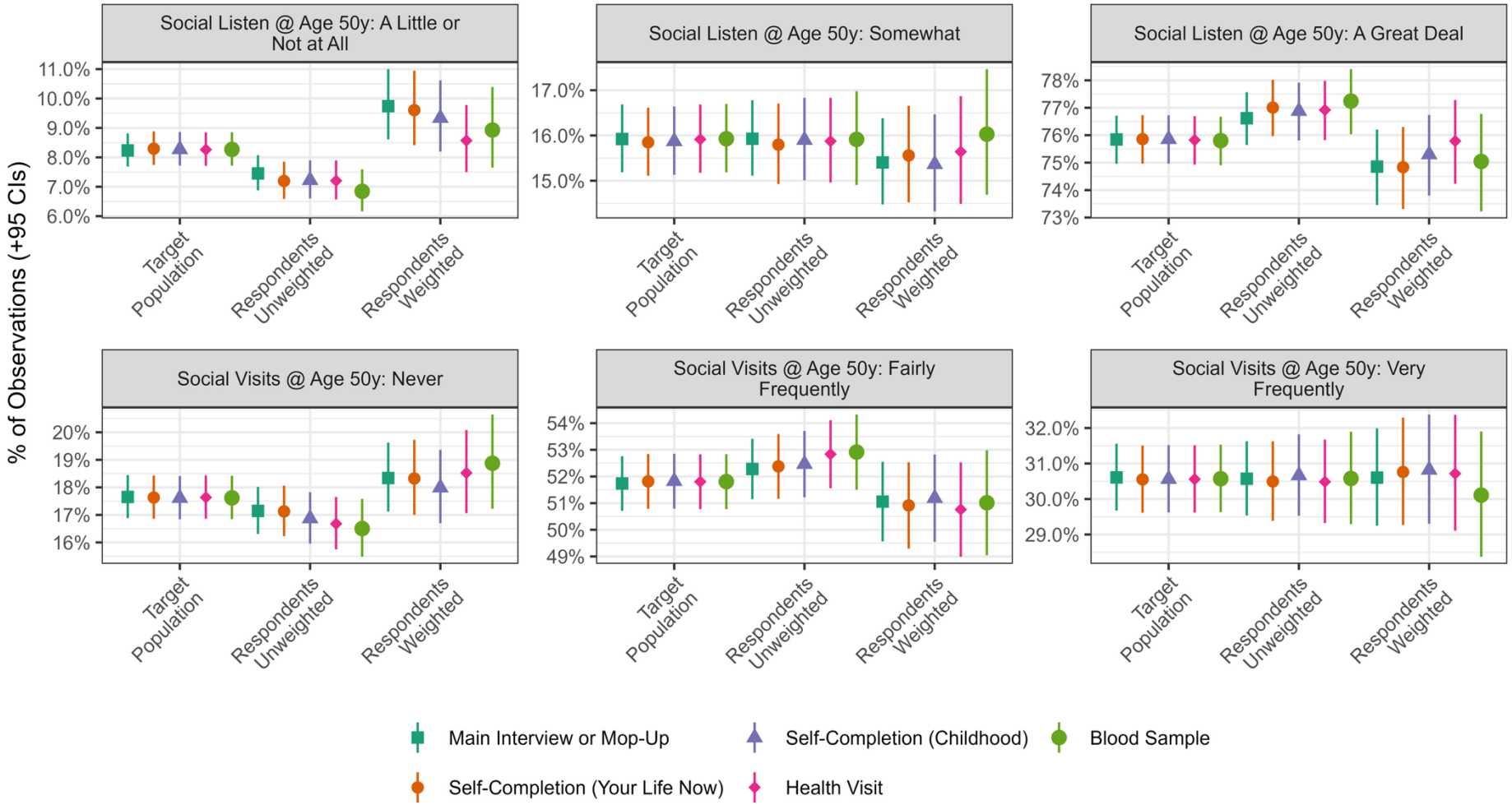


Figure A3.8: Sample proportions for categorical predictors included in response model, by variable category, population (target or respondents) and whether sample is weighted or unweighted.

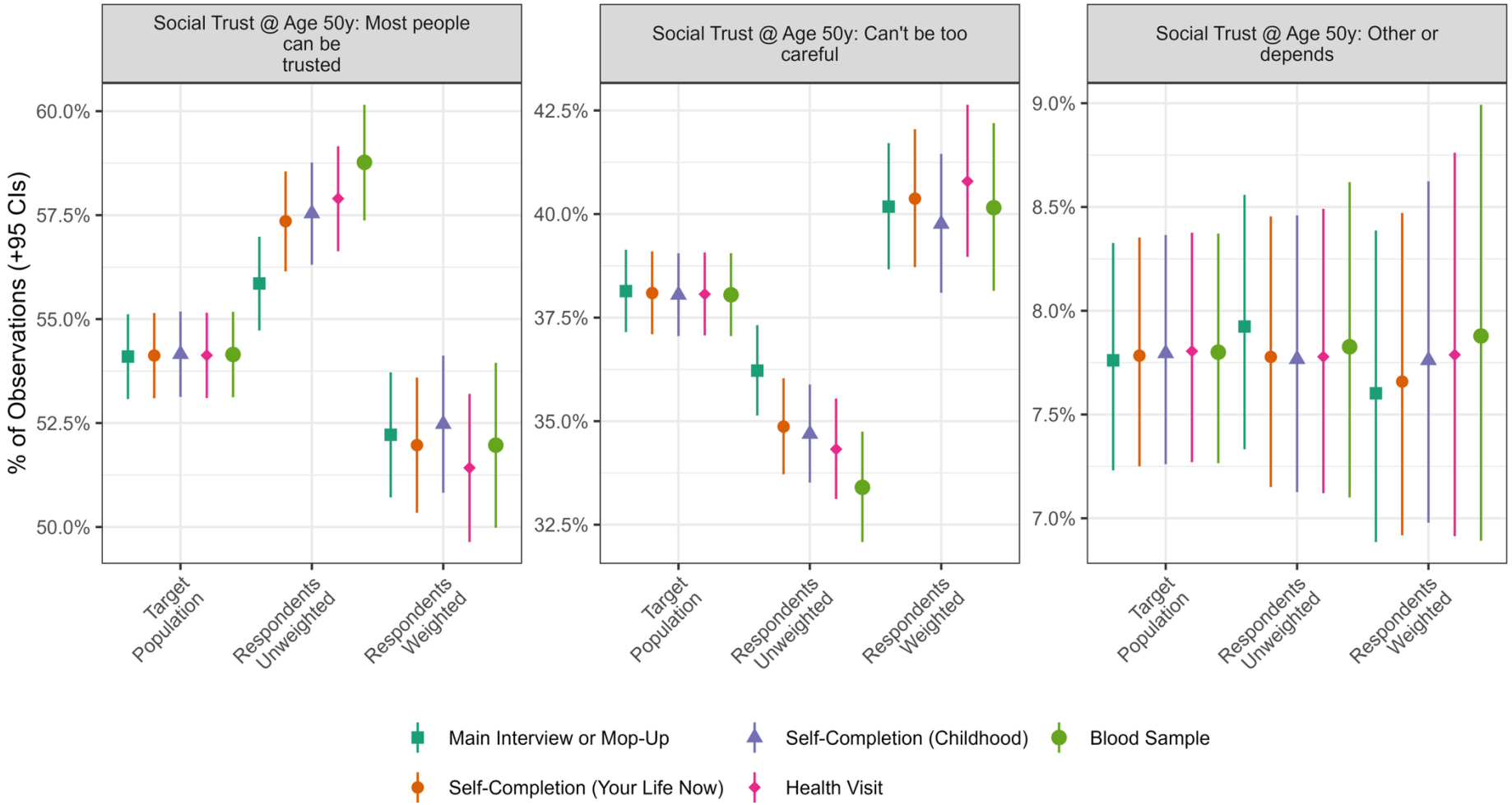


Figure A3.9: Sample means for continuous predictors included in response model, by variable, population (target or respondents) and whether sample is weighted or unweighted.