

User Guide to the Nurse Visit Datasets Waves 2,4,6,8,9

CONTENTS

1 - Overview of survey	4
1.1 Background and aims	4
1.2 Ethical clearance.....	5
1.3 Contact details	5
2 - Sample design	6
2.1 The ELSA interview sample	6
2.2 The ELSA nurse visit sample	6
3 - Nurse visit	8
3.1 Data collection methods.....	8
3.2 Informed consent	8
3.3 Modules, measurements and samples.....	9
3.4 Available data	10
3.4.1 Physical measures.....	10
3.4.2 Saliva sample.....	10
3.4.3 Hair sample.....	10
3.4.4 Drug coding.....	10
3.4.5 Blood sample	11
➤ Genetic data.....	11
3.4.6 Fluid intelligence (number series) – cognitive function.....	11
3.5 Nurse questionnaire documentation.....	11
4 - Nurse protocols	12
4.1 Blood pressure.....	12
4.2 Grip strength	12
4.3 Height and weight measurement.....	12
4.4 Waist and hip circumference.....	13
4.5 Lung function	13
4.5.1 Waves 2 and 4	13
4.5.2 Wave 6.....	14
4.6 Balance.....	14
4.7 Leg rise.....	14
4.8 Chair rise.....	15
4.9 Saliva sample	15
4.10 Hair sample.....	15
4.11 Polypharmacy: prescribed medications.....	16
4.12 Blood sample	16

5 - Dataset information..... 19

5.1	Datafiles available at the UKDS	19
5.2	Order and content of the nurse dataset.....	19
5.2.1	Dropped variables.....	19
5.2.2	Geographic variables.....	20
5.2.3	Access to non-archived data	20
5.3	Serial numbering.....	20
5.4	Missing values	21
5.5	Questionnaire errors	22
5.6	Coding and editing	22

6 - Weighting 23

6.1	Weights available.....	23
6.1.1	Wave 8 weights.....	24
6.2	Use of weights	24
6.3	Clustering and stratification.....	24

1 - Overview of survey

This User Guide provides detailed information about the data collected in the nurse visit of ELSA which was part of the survey at Waves 2, 4, 6, and 8 and 9. The data for each of these waves are available to download from the [UK Data Service](#). There is a separate User Guide for the interviewer data which can also be downloaded from the UK Data Service.

You may also wish to consult the nurse materials included in the relevant waves' documentation available on the UK Data Service website, e.g. the Wave 8 ELSA documentation includes the nurse questionnaire specification and the 'Coding prescribed medications' booklet including information on the code frame for the drug coding data.

In addition, you can also refer to the other materials used by nurses during ELSA fieldwork (such as nurse project instructions, consent booklet and respondent information leaflets) which are all available at: <http://www.elsa-project.ac.uk/documentation>.

1.1 Background and aims

The English Longitudinal Study of Ageing (ELSA) began in 2002. It is a large scale longitudinal panel study of people aged 50 and over and their partners, living in private households in England. The original sample was drawn from households that had previously responded to the Health Survey for England (HSE) between 1998 and 2001. The sample has been refreshed at several waves (Waves 3, 4, 6, 7 and 9) with new respondents from later years of HSE joining the ELSA sample.¹

The ELSA sample is invited for an interview at two-yearly interviews, known as 'waves', to measure changes in their health, economic and social circumstances. ELSA can complete the picture of what it means to grow older in the 21st century, and help us understand what accounts for the variety of patterns that are seen. There have been 9 waves of data collection so far.

Up to and including Wave 6, a nurse visit has been carried out at alternate waves in addition to the main interview. Towards the end of the main interview, respondents are asked whether they would be willing to take part in a nurse visit within the following few weeks. All core sample members have been offered a nurse visit at Waves 2, 4 and 6. In Waves 8 and 9, two mutually exclusive subsets of the sample were offered a nurse visit. The nurse visit includes the collection of biological samples and anthropometric measurements.

Many of the measures adopted in ELSA are comparable with measures used in the US Health Retirement Study (HRS) and the Survey of Health, Ageing and Retirement in Europe (SHARE).

ELSA is the result of collaboration between University College London (UCL), the Institute for Fiscal Studies (IFS), the University of Manchester and NatGen Social Research. Other academic collaborators based at the Universities of Cambridge, Exeter and East Anglia provided expert advice on specific modules. Funding for the first eight waves of ELSA has been

¹ Further detailed information about the refreshment samples at each wave can be found in the User Guide to the interviewer data available to download from the [UK Data Service](#).



provided by the US National Institute on Aging, and a consortium of British Government departments².

1.2 Ethical clearance

Ethical approval for all ELSA waves was obtained from NHS Research Ethics Committees under the National Research and Ethics Service (NRES). More information is available on the website: <http://www.nres.nhs.uk/>

1.3 Contact details

Any queries related to this study or the datasets should be sent to: elsadata@natcen.ac.uk

² More detailed information about funding can be found in the User Guide to the interviewer data.

2 - Sample design

2.1 The ELSA interview sample

The ELSA sample has been designed to represent people aged 50 and over, who were living in private households in England in the first wave of ELSA (2002/2003). Three years of the Health Survey for England (HSE) were selected as the sampling frame: 1998, 1999 and 2001. These years were chosen because they were recent and could provide a sufficiently large sample size. ELSA used the core samples for these years, all of which were nationally representative.

As the sample has aged, it has been refreshed at various waves to ensure that at each wave it is representative of the population in England aged 50 years and over. The refreshment samples have been selected from later years of the HSE. For more detailed information about sampling at each wave of ELSA, please see the User Guide to the interviewer data.

2.2 The ELSA nurse visit sample

Only people originally sampled into ELSA – so-called ‘core sample members’³ – who had an interview in person (i.e. not by proxy) at the relevant ELSA wave are eligible for a nurse visit at that wave. The nurse datasets also include data from some respondents who are not core sample members. These respondents are cohabiting partners of the core members who were offered a discretionary nurse visit because they specifically requested one from the nurse.

At Waves 8 and 9, nurse eligibility was further determined by sub-sampling. The selection was done in two stages. The first stage, prior to Wave 8 fieldwork, used purposive sampling (within cohort) and prioritised those who had responded to all previous nurse visits from cohorts 1 through to 6 to be issued for a nurse visit at Wave 8⁴. The remaining cohort members were flagged for a nurse visit in Wave 9, thus ensuring that all cohort members were eligible for a nurse visit in Wave 8 or Wave 9, conditional on completing the mainstage interview at the Wave to which they were eligible for a visit. All respondents from Cohort 9 were flagged as eligible for a nurse visit in Wave 9.

³ There are a number of different individual level sample types within the main ELSA sample. Individual sample type is based on age, participation at each wave and which cohort of HSE respondent was sampled from. More detailed information on individual level sample types can be found in the User Guide to the interviewer data.

⁴ Cohort 7 was excluded from this first group (presumably) on the basis that the non-response patterns across waves was unknown for this group which had only completed one prior wave

Table 1 shows the number of eligible respondents for a nurse visit at each wave, followed by the number of participating respondents. The final column shows the number of partners in the nurse dataset (who were not eligible for a visit, but received one if they requested it).

Table 1 Response rates and productive nurse interviews by wave

	Eligible*	Productive**	Response Rate	Partner Interviews
Wave 2	8,780	7,666	87.3%	0
Wave 4	9,592	8,218	85.7%	425
Wave 6	9,169	7,731	84.3%	323
Wave 8	3,714	3,479	93.7%	46
Wave 9	3,640	3,047	83.8%	22

* CAPI personal interview, eligible core member

** Core Member

3 - Nurse visit

3.1 Data collection methods

The ELSA nurse visits on Waves 2, 4, 6 and 8 have comprised a personal face-to-face CAPI interview and the collection of various biological samples and measurements.

The importance of reading out the questions exactly as specified was emphasised to the nurses during face-to-face briefings for the study to ensure consistency across the sample. Each respondent was offered a copy of their results for suitable measures included in that wave's nurse visit (blood pressure, height, weight, waist, hip and lung function). The nurse was asked not to give any interpretation of the results except for blood pressure, for which the nurse was asked to indicate only whether the measurement was normal or high and, where necessary, whether the respondent should contact their GP.

With the respondent's consent, a letter was also sent to those who gave a blood sample showing whether the result of the analyses conducted on the blood sample they gave was within or outside the normal range (for results where there is an agreed "normal range"). If any results were out of range, respondents were told that they should contact their GP in the near future.

Again, with the respondent's consent, we sent their blood pressure, lung function and blood sample results (as relevant for each wave) to their GP. The exact results for the blood analyses were included, and GPs were informed of the normal range used for each analysis.

The results to respondents and their GPs were sent within three months of the nurse visit, unless there was a clinical indication to do so more urgently.

3.2 Informed consent

Each respondent who took part in a nurse visit was asked to complete a consent booklet for samples to be taken and for any results to be sent out to their GP. This involved respondents reading and signing consent statements, giving permission for each sample to be taken and for each test result to be sent to their GP. Written consent was obtained for the following:

- blood pressure readings to be sent to their GP
- lung function readings to be sent to their GP
- blood samples to be taken
- blood sample test results to be sent to their GP
- blood sample for storage for future analysis
- blood sample for DNA extraction and storage
- saliva samples to be collected (Waves 2 and 4)
- hair sample to be collected (Wave 6)
- PAXGene sample to be collected and stored (Wave 8 and Wave 9)

Respondents were given a copy of the consents they had signed to keep for their records. Nurse visit consent booklets for each wave can be viewed at the ELSA project website [here](#).

3.3 Modules, measurements and samples

The nurse visit has been a feature of HSE since the survey was first carried out in 1991. When the nurse visit was first introduced into ELSA at Wave 2, most modules from the HSE nurse visit were included and a number of new ones were added in. The modules that were taken from HSE were:

- blood pressure,
- blood sample,
- standing and sitting height,
- weight,
- waist and hip measurement
- lung function
- drug coding.
- The modules that were added were
- balance,
- leg raise,
- chair rise,
- grip strength,
- saliva log (Waves 2 and 4)
- hair sample (Wave 6).

While the majority of the nurse visit content is the same across all waves, some additional modules and analytes have only been included at particular waves, and a narrower set of measures were included on the Wave 8 nurse visit. The balance, leg raise and chair rise, taken alongside the walking speed measurement carried out in the main ELSA interview, form a battery of tests that have been shown to be highly predictive of level of disability, future use of health care and mortality. These measures were adapted from the EPESE (Established Populations for Epidemiologic Studies of the Elderly) protocol, which looks at older cohorts and the development of disability ^{5,6}.

⁵ Studenski S, Perera S, Wallace D, Chandler JM, Duncan PW, Rooney E, Fox M, Guralnik JM. 2003, 'Physical performance measures in the clinical setting', *J Am Geriatr Soc.*, 51, pp. 314-22.

⁶ Kuh D, Hardy R, Butterworth S, Okell L, Richards M, Wadsworth M, Cooper C, Sayer AA. 2006, 'Developmental Origins of Midlife Physical Performance: Evidence from a British Birth Cohort', *Am J Epidemiol.*

The grip strength measure was taken from the Survey of Health, Ageing and Retirement in Europe (SHARE).⁷

The changes between HSE and ELSA were made because ELSA focuses on an older population. The collection of saliva and the accompanying questionnaire (in Waves 2 and 4) and hair (in Wave 6), in order to measure cortisol, was added because preliminary data from the Whitehall II study showed that cortisol levels are linked to social environments and ageing.⁸ A saliva log was included at Waves 2 and 4. At Wave 6, a hair sample was taken to measure cortisol and so a saliva log was not included at this wave.

Additionally, at Wave 8 the fluid intelligence section of the cognitive function module was moved from the interviewer schedule to the nurse schedule. This change was made in order to making the Wave 8 interviewer questionnaire too long.

3.4 Available data

3.4.1 Physical measures

The results of a full set of physical measures (blood pressure, grip strength, height and weight, waist and hip, lung function, balance, leg rise, chair rise) are included in the main nurse datasets of Waves 2, 4 and 6. The Wave 8 main nurse dataset includes the results of blood pressure and grip strength physical measures. The weight measurement was moved to the interviewer visit at Wave 8 and the data can be access in the ELSA main interviewer dataset.

3.4.2 Saliva sample

Due to funding restrictions, not all saliva samples have yet been analysed. Wave 2 salivary cortisol data is available in the UK Data Service.

3.4.3 Hair sample

Data from the analyses of hair samples are not yet available in the archived datasets. Information about who gave a hair sample is available in the main nurse data.

3.4.4 Drug coding

Coded data of the prescribed medication is available in the Wave 6 and 8 nurse data. Up to 27 prescribed medications were recorded and coded for each respondent at Wave 6 (with information on why the drug was taken coded for up to 22 drugs), and up to 22 drugs at Wave 8. Please see the separate 'Coding prescribed medications' manual, available as part of the UKDS ELSA documentation, for more details on the specific drug codes.

In the Wave 8 data, the more disclosive full drug code and type variables are archived under a nurse Special Licence dataset. Please see the Wave 8 Questionnaire & Data Documentation for more details.

⁷ <http://www.share-project.org/>, retrieved 12/7/2010

⁸ Cohen S, Schwartz JE, Epel E, Kirschbaum C, Sidney S, Seeman T. 2006, 'Socio-economic status, race, and diurnal cortisol decline in the Coronary Artery Risk Development in Young Adults (CARDIA) Study', *Psychosom Med.*, 68, pp.41-50.

3.4.5 Blood sample

The analytes from blood sample analysis are included in each wave's main nurse dataset.

➤ *Genetic data*

Further Genome-wide genotyping has been performed on blood samples collected on ELSA Waves 2 and 4. For more details and to enquire about data access please see: <http://www.elsa-project.ac.uk/gwas> or contact elsa@ucl.ac.uk.

3.4.6 Fluid intelligence (number series) – cognitive function

Data from the fluid intelligence number series questions are included in the nurse data for Wave 8 and 9, and the ELSA main interviewer data for earlier waves.

3.5 Nurse questionnaire documentation

Detailed documentation for each of the waves can be downloaded from the UK Data Service. The documentation includes:

- the nurse questionnaire, i.e. details of the data variables where nurses have captured information on prescribed medication, physical measurements and the administration of the physical samples (saliva, hair and blood)
- information about nurse derived variables
- information archived data files and their contents
- the 'Coding prescribed medications' booklet containing detailed drug codes (used to code prescribed medication) for Waves 6 and 8.

4 - Nurse protocols

Full documentation of the protocols for all measurements and samples detailed in this section are included in the nurse project instructions for each wave which can be downloaded at <http://www.elsa-project.ac.uk/documentation>.

In addition to any specific exclusion criteria detailed below, if a participant did not wish to take part in a measure or sample or the nurse felt that a procedure was not safe for a given individual, the measure or sample should not have been done.

4.1 Blood pressure

All respondents were eligible to have their blood pressure measured except those who were pregnant.

Three measurements were taken of systolic and diastolic pressure as well as pulse rate on the respondent's right arm while they were seated. The respondent was given advice if their results indicated a higher than normal reading. The nurses were instructed to give this advice based on the higher of the last two blood pressure readings – the first reading can be high, as people are sometimes nervous about having their blood pressure taken.

If you wish to compare the blood pressure results to earlier HSE ones, please note that Omron machines were used to take the readings in the ELSA Waves 2, 4 and 6 nurse visits and in HSE from 2003 onwards. In HSE prior to 2003, Dinamap machines were used to take the readings. A conversion factor will need to be applied to the results, as the machines are not comparable. Please contact the data team at elsadata@natcen.ac.uk for further information or help with this.

4.2 Grip strength

All respondents were eligible to have their grip strength measured.

Three measurements of grip strength were taken on both the dominant and non-dominant hand. The respondent was asked which hand was their dominant one. The precise measure carried out was the isometric handgrip strength measure.

4.3 Height and weight measurement

Height was measured both standing and sitting in Waves 2 and 4, and just standing in Wave 6. Weight was measured by nurses on Waves 2, 4 and 6. On Wave 8 and 9 weight measurement was carried out by interviewers in order to collect weight data from the full sample and is included in these interviewer datasets.

Sitting height is a measure of pre-pubertal growth. All respondents were eligible to have their height and weight measured. If height or weight could not be measured (because the respondent was chair-bound or too unsteady) then an estimate was obtained from the respondent instead. If the nurse thought the measurement was likely to be more than 2 cm (3/4 inch) from the true figure for height or more than 1 kg (2 lbs.) from the true figure for weight, it was considered unreliable and they were asked to code it as such.



The maximum weight capacity of the scales was 130kg (20½ stone). If the nurse thought the respondent exceeded this limit then they were instructed to code “Weight not attempted” and ask the respondent for an estimate instead.

Users of the data are reminded to consider the variables RELHITE and RELWAIT when looking at the measurements in this module, as they show whether the height measurements and weight measurements respectively are likely to be reliable.

Using the height and weight measurements obtained, body mass index (BMI) was calculated for each respondent in Waves 2, 4 and 6. This is a measure of relative weight based on an individual's height and weight that applies to both men and women. BMI values were then grouped according to World Health Organisation definitions of obesity.

4.4 Waist and hip circumference

All respondents were eligible to have their waist and hip measurements taken, unless they were chair-bound or had a colostomy or ileostomy.

Both of these measurements were taken twice, however, if the second measurement differed from the first by 3cm or more, the nurse received an error message in the CAPI program and was prompted to either amend one of the previous responses if a mistake had been made entering a measurement, or to take a third measurement. If the nurse believed that the measurements they took were 0.5cm more or less than the true measurement because of problems encountered (e.g. clothing the respondent was wearing), this was considered unreliable.

4.5 Lung function

At Wave 6, due to major technological advances, a different model of spirometer was used to measure lung function. The model differed significantly from the model used at Waves 2 and 4 and so results across waves should be interpreted separately.

4.5.1 Waves 2 and 4

All respondents were eligible to have their lung function measured, except those who:

- had had abdominal or chest surgery in the preceding 3 weeks;
- had been admitted to hospital with a heart complaint in the preceding 6 weeks;
- had had eye surgery in the preceding 4 weeks;
- were pregnant;
- had a tracheotomy.

Three measurements each were taken of forced vital capacity (FVC), forced expiratory volume (FEV) and peak flow (PF) using a spirometer.

It should be noted that the variables HTFVC and HTFEV (highest technically satisfactory values of FVC and FEV respectively) should not be combined to give a FEV/FVC ratio without checking that they are from the same blow.

4.5.2 Wave 6

All respondents were eligible to have their lung function measured, except those who:

- were pregnant;
- had had abdominal or chest surgery in the last three months;
- had had a heart attack in the last three months;
- had had a detached retina or eye or ear surgery in the last three months;
- had been admitted to hospital with a heart complaint in the preceding month;
- had a resting pulse rate more than 120 beats/minute (after sitting for at least five minutes prior to the pulse rate being taken);
- were currently taking medications for the treatment of tuberculosis.

The aim was to collect three acceptable blows from each eligible respondent. After each attempt, the program advised the nurse whether the blow was acceptable. If it wasn't, the program instructed the nurse to ask the respondent to try again. At least 3 and up to 8 measurements were taken. As in Waves 2 and 4, measurements taken using a spirometer were of FVC, FEV and PF. The results were output automatically from the spirometer through the computer (rather than being entered by hand by the nurse). The output for FEV/FVC ratio was automatically generated.

4.6 Balance

The eligibility for the balance module depends on age of respondent and performance during the stands. All respondents start with the side-by-side, if they held this for 10 seconds they attempted the semi-tandem stand for 10 seconds. Respondents who completed this were then asked to do the full tandem stand. If the respondent was aged 69 and under they were asked to attempt the full tandem stand for 30 seconds; if they were 70 or over they were asked to do the full tandem stand for 10 seconds.

This module involved the respondent completing up to three stands:

- a side-by-side: stand with feet together, side by side
- a semi-tandem: stand with the side of the heel of one foot touching the big toe of the other foot
- a full-tandem: stand with the heel of one foot in front of and touching the toes of the other foot.

Each of these was demonstrated by the nurse to the respondent beforehand.

4.7 Leg rise

Only respondents aged 69 and under who successfully passed the side-by-side stand were eligible and therefore asked to complete this module. They were asked to stand on one leg with

their eyes open for 30 seconds and then, if they did this, they were asked to complete the same movement with their eyes closed for 30 seconds.

4.8 Chair rise

All respondents were eligible for the chair rise.

This is a measure of lower body strength, during which respondents were asked to stand up from a firm chair without using their arms. If they succeeded, they were asked to stand up and down as quickly as they could for either five rises if they were aged 70 and over, or up to ten rises if aged 69 and under. The nurse recorded the time that respondents took to do the number of rises required. For respondents who did ten rises, the nurse recorded the times taken to do both five and ten rises (in the same attempt) so that all respondents had a time for five rises which could be compared.

4.9 Saliva sample

Selected eligible respondents at Wave 2 and 4 were asked to give a saliva sample. All respondents aged 79 years and under at Wave 2 and all respondents from the Wave 4 refreshment sample were asked to give a saliva sample. We also selected 10% of respondents who gave a saliva sample at Wave 2 to give a further sample at Wave 4.

Respondents who had been pre-selected to give a saliva sample were asked to collect four samples of their saliva at certain times during a 24-hour period. The purpose of collecting saliva was to measure respondents' cortisol levels, which are related to stress. Respondents were asked to fill in a log book each time they collected a saliva sample that asked how they were feeling at that time. The saliva and log book data have not been archived as part of the main ELSA data release.

4.10 Hair sample

In Waves 6 a sample of hair was collected with the purpose to measure respondents' cortisol levels, which are related to stress. All respondents were eligible to have the hair sample taken, except those in the following circumstances:

- Pregnancy (in 'younger partner' respondents)
- Breastfeeding (in 'younger partner' respondents)
- Current scalp condition rendering the hair sample soiled or at risk of transmission of an unknown / known blood borne virus (e.g. active bleeding or infection)
- Respondent unable to sit with head remaining still (e.g. continual tremor, head shaking)
- Has less than 2cms of hair length in the posterior vertex scalp area (not including any hair extensions)

The sample needed to be a minimum of 2cms in length and weigh a minimum of 10mg (i.e. the smaller the sample length, the greater the number of strands required for a 10mg sample). The sample was collected in the posterior vertex area of the scalp which is the area of hair growth that shows the most consistent levels of Cortisol. Respondents were asked to sit in a suitable chair (i.e. where the nurse could access the posterior vertex area of the respondent's head and have the equipment close to hand) and tip their chin down towards their chest so that the

posterior vertex of the head rotates upwards. Nurses gently combed the respondent's hair over the posterior vertex area to align the hair strands together. Then using the hair sectioning grips, they gently secured the hair off on each side of the identified sample site ensuring the hair strands left in the middle constitute the width of around 1.5mm when twisted. The sample of hair was taken as close as possible to the scalp.

4.11 Polypharmacy: prescribed medications

In Waves 6 and 8, respondents were asked whether they were currently taking any medication that had been prescribed for them by a doctor. The name of each medication was recorded by the nurse and a code was attributed to the medication according to the British National Formulary (BNF) (version 61 at Wave 6, version 69 at Wave 8). Codes are recorded in a six-digit format reflecting three levels of classification in the BNF, using a leading zero where appropriate. For example a drug code 100101 is a drug listed in Section 10.1.1 in the BNF, and so on.

Please see the 'Coding prescribed medications' booklet – included in the ELSA Wave 8 documentation on the UKDS – for details of the most commonly used drug codes and the associated drug names.

4.12 Blood sample

All sample members who gave consent were eligible for a blood sample to be taken. The only exceptions to this were people with clotting or bleeding disorders, people with a history of fits or convulsions, or people who were on anticoagulant drugs (e.g. Warfarin, protamine, acenocoumarol).

Respondents under 80 years old were asked to fast before their nurse visit so a fasting blood sample could be taken. Respondents were not asked to fast if they had diabetes and were on treatment or if they were considered to be malnourished or otherwise unfit to fast (this information was obtained from the interviewer). Respondents who were asked to fast were given guidelines about when and what they could eat based on their appointment time.

In the nurse visit, respondents were asked when they had last eaten and, if this was on the day of the nurse visit, what they had eaten. The CAPI program used their responses to work out if they had fasted adequately. A respondent was considered to have fasted and therefore be eligible for a fasting blood sample if (see variable FASTELI in Wave 2 and FASTELIG in Waves 4, 6 and 8):

- They had last eaten the day before the nurse visit OR
- They hadn't eaten or drunk anything (apart from water) on the day of their nurse visit OR
- They had eaten or drunk on the day of their nurse visit but at minimum of 5 hours earlier, and had only had a light meal or a piece of fruit or drink the last time they ate.

Blood was only taken from respondents on one occasion; so if they had fasted adequately (i.e. met one of the conditions above) then all the analytes for that person should be considered as a fasting sample, otherwise they were non-fasting samples. All the blood analytes (except blood glucose) were measured for all the blood samples (i.e. both fasting and non-fasting samples). Therefore, for some cases the lipids measures were on fasting samples and for others it was on non-fasting samples. If you are doing analyses that are dependent on the blood being a fasting

sample, e.g. fasting lipids for metabolic syndrome or cardiac risk, please ensure that you only use the sub-sample of respondents who actually fasted (i.e. FASTELI or FASTELIG=1).

Blood glucose was only measured for people who had fasted.

Respondents were asked if they consented to DNA being extracted from their blood sample and stored for future analysis⁹.

At waves 2, 4 and 6, a maximum of six small tubes of blood (ranging in size from 2ml to 6 ml) were collected from each respondent. Three of these were collected from all respondents. An additional tube was collected if the respondent had fasted, and the final 2 tubes were collected if the respondent consented to have their DNA analysed.

At Wave 8 and 9, a maximum of five small tubes of blood were collected from each respondent. Three of these were collected from all respondents. An additional tube was collected if the respondent had fasted, and a final tube was collected if the respondent consented to a PAXGene extraction and storage. In addition, as a protocol amendment from half way through Wave 8 nurse fieldwork, nurses who were using a butterfly needle filled an additional small (3ml) discard tube prior to filling the first tube (Blue Citrate tube) in the order of draw. This discard tube did not need to be filled completely and could be swapped for the remainder of the draw as soon as blood enters the tube (and safely discarded). The purpose of the discard tube was to ensure there is no dead air space within the butterfly, which was identified as a contributing factor to a higher than expected level of samples with an under fill of the Blue Citrate Tubes.

The blood samples were sent to an external laboratory where a number of analyses were carried out to measure the levels of certain compounds in the blood. As detailed in Table 2, most of the analytes were included in all waves of ELSA nurse visits.

Table 2 Blood assays carried out by wave

Type of blood assay	Wave 0 ¹⁰	Wave 2	Wave 4	Wave 6 ¹¹	Wave 8/Wave 9
Triglycerides	✓	✓	✓	✓	✓
Total and HDL-cholesterol	✓	✓	✓	✓	✓
LDL Cholesterol		✓	✓	✓	✓
C-reactive protein, fibrinogen	✓	✓	✓	✓	✓
Haemoglobin and ferritin	✓	✓	✓	✓	✓
White blood cell count, mean corpuscular haemoglobin			✓	✓	✓
Fasting lipids, glucose, glycated haemoglobin ²		✓	✓	✓	✓

⁹ Respondents who had taken part in a previous wave and given a sample from DNA extraction and storage were not asked to give another at subsequent waves. Only new sample members or those who had refused at a previous wave were asked about this in subsequent waves.

¹⁰ Original HSE nurse visit

¹¹ At waves 6 and 8, Glycated haemoglobin (HbA1c) was analysed in the more modern IFCC units (mmol/mol) instead of the traditional DCCT (%) units. To convert the values in Wave 6 to the old % values (as in W2 and W4), the formula is $(XX/10.929)+2.15$

Table 2 Blood assays carried out by wave

Type of blood assay	Wave 0 ¹⁰	Wave 2	Wave 4	Wave 6 ¹¹	Wave 8/Wave 9
Cortisol (from saliva)		✓	✓		
IgE / DHM IgE	✓				
IGF-1			✓	✓	✓
DHEAS			✓		
Vitamin D				✓	✓
DNA extraction and storage		✓	✓	✓	
PAXGene extraction and storage					✓

Fibrinogen – A protein necessary for blood clotting. High levels are also associated with a higher risk of heart disease.

Total cholesterol – Cholesterol is a type of fat present in the blood, related to diet. Too much cholesterol in the blood increases the risk of heart disease.

HDL cholesterol – This is ‘good’ cholesterol, which is protective for heart disease.

Triglycerides - Together with total and HDL cholesterol, they provide a lipid profile that can give information on the risk of cardiovascular disease.

LDL cholesterol – This is ‘bad’ cholesterol; increased levels are associated with atherosclerosis, and thus myocardial infarctions, strokes and peripheral vascular disease.

Ferritin and Haemoglobin (Hb) – These are measures of iron levels in the body and are related to diet and other factors.

C-reactive protein (CRP) – The level of this protein in the blood gives information on inflammatory activity in the body, and it is also associated with risk of heart disease.

Apolipoprotein E (ApoE) – This is involved in the transport of cholesterol and plays a protective role.

Fasting glucose and non-fasting glycated haemoglobin (HBA1c) – Both indicate the presence or risk of type 2 diabetes, which is associated with an increased risk of heart disease. The fasting glucose result is now archived with this version of the data.

White blood cell count (WBC) and mean corpuscular haemoglobin (MCH) – When looked at in combination with ferritin and haemoglobin can indicate anaemia.

Vitamin D – Obtained from the diet and from sunshine, Vitamin D is needed for healthy bones.

Insulin-like growth factor 1 (IGF-1) and dehydroepiandrosterone sulfate (DHEAS) – These are hormones that help control reactions to stress and regulate various body processes including digestion, the immune system, mood, and energy usage.

The samples were taken in a particular order so that if a situation arose where there was insufficient blood to fill all the tubes, the analyses with the highest priority could still be undertaken. The analyses in order of priority were fibrinogen, full lipids (total cholesterol, HDL cholesterol and triglycerides), ferritin, CRP, IGF-1, DHEAS, Vitamin D, ApoE, fasting glucose (if applicable), haemoglobin, glycated haemoglobin, white cell count, mean corpuscular haemoglobin and finally DNA extraction (if consent was given).

The nurse dataset includes a derived variable BSOUTC, which shows whether taking a blood sample was attempted and, if so, how successful it was.

5 - Dataset information

This User Guide refers to the nurse visit datasets that have been archived for ELSA waves 2, 4, 6 and the combined Wave 8 and 9 dataset.

5.1 Datafiles available at the UKDS

For each of the nurse waves, an ELSA nurse visit dataset is available via the UK Data Service under End User Licence (EUL). Data for ELSA respondents collected during HSE (referred to as “Wave 0” data) is also available via the UK Data Service.

For Wave 8 and 9, an additional nurse dataset is available for users under the UKDS Special Licence (SL) conditions. This dataset contains detailed and highly disclosive variables about individual drugs taken in the seven days prior to the nurse visit.

5.2 Order and content of the nurse dataset

The nurse interview datasets are individual level files, including all respondents with a productive nurse visit at waves 4, 6, and all Core Members with a productive nurse visit at Wave 2. Since the nurse sample was split between wave 8 and 9, a dataset combining individuals from these waves has been archived. The main group of respondents for analysis is the Core Members. Data on Partners can be used as characteristics of the Core Members (i.e. to provide supplementary information), but these partners should not be analysed as individuals in their own right. The ineligible partners are unrepresentative, and any analysis using them would need to be unweighted.

The nurse datasets contain variables in the following order:

- Key variables not in the questionnaire (e.g. serial number, nurse visit month and year, outcome codes)
- Variables in the nurse questionnaire (in the order they appear in the CAPI interview) unless marked for not to be archived (see next section). A small number of additional computed variables that are associated with particular questionnaire variables are located alongside these variables in the data. These are annotated in the questionnaire.
- Other variables not in the questionnaire including administrative variables, derived variables, and survey weights. Derived variables are denoted with “(D)” at the beginning of the variable label¹².

5.2.1 Dropped variables

In preparing the data for archiving, certain variables were marked as not to be archived. Non-archived variables are denoted in the nurse questionnaire documentation. The following types of variables have been deleted in order to reduce the potential to identify individuals and for other reasons (specified below):

1. Uncoded open text responses

¹² Derived variables were not computed for Wave 7, but will be added following a retrospective review and re-archiving of the core datasets for waves 1-7.

2. Variables containing a personal identifier (e.g. name/address)
3. Other variables considered to be disclosive, such as:
 - Full interview date
 - Full date of birth
4. Timing variables, which give the time at specific points in the interview (used for administration purposes)
5. Administrative variables (e.g. link variables at start and end of sections) and variables that only contain missing values – excluded because they are not useful. Such variables have only been kept if they are integral to the structure of the data.

Single year of age has been capped at 90 (indager) within the EUL version of the dataset.

All efforts have been made to release as much of the ELSA data as possible to researchers while safeguarding the confidentiality of ELSA participants.

5.2.2 Geographic variables

Government Office Region (GOR) is included in the main interview dataset for all waves.

Various other more disclosive geographical variables are available under secure arrangements. For Wave 8 and 9, geographic variables are available in the Special Licence and Secure Access datasets as detailed above. Please see the Wave 8 and 9 Questionnaire & Data Documentation for details. For waves 1-7, please contact elsadata@natcen.ac.uk if you would like to request access to any geographic variables.

5.2.3 Access to non-archived data

Any further variables not included in the archived datasets may be requested via the NatCen Data Release Panel. Please contact the ELSA Data Manager at elsadata@natcen.ac.uk for more details.

5.3 Serial numbering

Constant Individual Serial Number

All the ELSA data files deposited in the archive contain a unique individual analytical serial number (**IDAUNIQ**) to enable users to link the different files – whether nurse and interviewer datasets for a given wave, or different files across waves. Each respondent has a unique value for **IDAUNIQ**, which will remain constant across all datasets at all waves.

Wave-specific household serial number

The five digit household serial number (for example, for Wave 8 **IDAHHW8**) was randomly generated for each Wave, and does not relate to the serial number used during interviewing, or to the household serial numbers of previous waves. The datasets for each wave of ELSA (including "Wave 0", i.e. HSE) contains a different set of household serial numbers (**IDAHHW0**, **IDAHHW1** etc.). It is necessary to have a different household serial number for each wave as respondents can change households between waves. The ELSA Index File, available from the

Data Archive, enables data users to link the household serial numbers in order to compare data for each respondent and household at different waves.

Person number

Each person within the household was given a number, starting from 01, at the time of the HSE interview (**PERID**). The numbering was continued for new people that entered the household after the HSE interview.

5.4 Missing values

The following missing value codes are used in the ELSA nurse datasets:

Table 5.1 Missing value codes in the nurse data

Code	Description	Relevant modules/sections	Waves
-9	Refusal	All	All
-8	Don't know	All	All
-4	Missing by error	All	Wave 8 only
-3	<i>No valid answer</i>	<i>N/A to nurse dataset – interviewer data only</i>	<i>N/A</i>
-2	<i>Self-completion instrument not completed</i>	<i>N/A to nurse dataset – interviewer data only</i>	<i>N/A</i>
-1	Item not applicable	All	All
-9	No analysis (office)	Blood analysis variables	All
-8	No analysis (lab)	Blood analysis variables	All
-7	Sample unsuitable for analysis (clotted)	Blood analysis variables	All
-6	Sample took more than five days to reach the lab so is unsuitable for analysis	Blood analysis variables	All
-3	Possible presence of variant Haemoglobin	Blood analysis variables	All
-2	Sample received but insufficient blood for analysis	Blood analysis variables	All
-1	Sample not received	Blood analysis variables	All



Most questions in the nurse CAPI questionnaire permit 'refusal' and 'don't know' answers. For the small number of questions where a response of 'don't know' or 'refusal' was not permitted in the CAPI instrument, this is indicated in the questionnaire.

For various reasons, some respondents did not complete the entire interview. For these 'partial' interviews, the questions towards the end of the questionnaire that were not asked are coded as '-1' (not applicable).

5.5 Questionnaire errors

In Wave 8, three respondents in the Wave 8 nurse dataset are coded to -4 'Missing by error' in the variable FQNURSE from the interviewer stage, indicating agreement for the follow-up nurse visit. This question at the end of the interviewer questions would have been expected to have been completed by these respondents but was left unanswered; they went on to complete a nurse visit regardless.

5.6 Coding and editing

A number of questions in the interview gave the nurse the opportunity to enter an 'other' answer. These answers were coded in the office, to see whether they could fit either the existing code frame or newly assigned codes.

In the nurse dataset, in the majority of cases the archived variables are the original nurse coded variants, rather than office coded variables where the 'other' responses have been back-coded into the original question where possible, because the additional coded comments deal with administrative information about conducting the tests.

6 - Weighting

6.1 Weights available

Nurse weights have been computed for each ELSA wave that included a nurse visit (waves 2, 4, 6, 8 and 9). Wave 8 and 9 samples are combined and weighted together – details for the weighting procedure are found below. Please note, the wave 8 nurse data had previously been archived with a weight which allowed estimates to be adjusted for that wave only. We have decided to remove this dataset because the sample was selected purposively and will be biased, and we would strongly recommend users generate analysis using the combined wave 8 and 9 dataset.

In the wave 2, 4 and 6 two weight variables are included in the nurse data files of each wave. The nurse weight applies to all those who received a nurse visit, while the blood weight applies only to those with blood sample results. The weights should be used when carrying out cross-sectional analysis of the nurse data.

The combined wave 8 and 9 dataset contains three weights. There are two nurse weights – the weight suffixed with “_N” is scaled to the total number of cases in the population, while the weight suffixed with “_G” is scaled to the total grossed population size of over 50s¹³

There is also a blood weight, which should be applied to the blood sample results.

Note that only ELSA Core Members living in private households in England (at the time of interview) have been given a weight. The weights are ‘system missing’ for ELSA Partners and Core Members not eligible to receive a weight (institutional respondents or those living outside of England). When running weighted analyses, researchers should remember to exclude these respondents (with ‘system missing’ weights) from the un-weighted base, if quoted. The data for Partners can be used as supplementary information for Core Members.

The weight variables in each dataset are summarised in the table below.

Table 6.1 Nurse and blood weight variables by wave

	Nurse visit weight variable(s)	Blood sample weight variable
Wave 2	w2nurwt	w2bldwt
Wave 4	w4nurwt	w4bldwt
Wave 6	w6nurwt	w6bldwt
Wave 8/9	w89nurwt_N w89nurwt_G	w89bldwt

Weights are necessary to adjust the composition of the responding sample so that it more accurately reflects the population of interest. If appropriate weights are not applied then the survey estimates will be biased in favour of the types of people who were more likely to participate in the survey and agree to a nurse visit or blood sample.

¹³ The 2017 mid-year estimates were used for population total, as this was the mis points between wave 8 and wave 9

6.1.1 Wave 8 and 9 weights

The Wave 8 and 9 nurse sample is designed to be analysed as a whole rather than split into its constituent waves. As noted above, unlike in previous waves where all core members responding to the main interview were eligible for a nurse visit in that wave, across wave 8 and wave 9, two mutually exclusive subsets of members were pre-selected (prior to fieldwork): one to be offered a nurse visit at wave 8 and the other to be offered a nurse visit in wave 9.

The variable **Wave** indicates which wave the Nurse data was collected in. We have also included **NurGrp**, which defines weighting approach taken for each case. Details of the weighting approach will be in the Wave 9 technical report (Forthcoming)

6.2 Use of weights

Where possible we recommend that nurse analysis be conducted on weighted data since this will help to minimise bias from differential non-response amongst key sub groups.

If co-variables from the ELSA interviewer data are needed for cross-sectional analysis of the combined data, wave 8 interview data should be used for nurse data collected in wave 8 and wave 9 data should be for wave 9 nurse data. The variable **Wave** can be used to select the correct cases for this matching. Please note that, the data dictionaries for wave 8 and wave 9 variables will need to be precisely matched so that the correct version of the variable can be written into one integrated set of columns for use in data analysis.

6.3 Clustering and stratification

With complex sample designs such as ELSA, the effects of clustering and stratification should also be taken into account when conducting weighted analyses.

Analysts should use the following cluster and stratification variables, according to wave. These variables are available on the main interviewer datasets of the relevant.

Table 6.2 Cluster and stratification variables, Waves 1–8

	Cluster variable	Stratification variable
Wave 1	ahsecls2	astratif
Wave 2	Hseclst	astratif
Wave 3 onwards¹⁴	Idahhwn (household serial number: 'idahhw' plus Wave number e.g. for Wave 6, this is idahhw6) (*)	GOR

¹⁴ From Wave 3, attrition means that geographical clustering is negligible; however, clustering within household should be taken into account.