

English Longitudinal Study of Ageing (ELSA)

Wave One to Wave Five

User Guide to the datasets

Prepared by Natcen Social Research
August 2012

CONTENTS

1	Overview of the Survey	1
1.1	Background and Aims	1
1.2	Ethical Clearance	2
1.3	Contact Details	2
2	Sample design.....	3
2.1	Sample Design	3
2.2	Refreshment samples.....	3
2.3	Sample Types.....	3
2.4	Proxies.....	4
2.5	Institution Interviews	5
3	Survey content.....	6
3.1	Development and piloting	6
3.2	Details of survey content - Overview	6
3.3	Details of survey content - Wave one.....	7
3.4	Details of survey content – Wave two	7
3.5	Details of survey content – Wave three.....	8
3.6	Details of survey content – Wave four.....	10
3.7	Details of survey content – Wave five	11
4	Fieldwork and response rate	13
4.1	Fieldwork - Overview	13
4.2	Fieldwork - Tracing Movers	13
4.3	Response rates at wave one	14
4.4	Response rates at wave two	14
4.5	Response rates at wave three.....	15
4.6	Response rates at wave four.....	15
4.7	Response rates at wave five	15
5	Weighting.....	17
5.1	Weighting for waves one and two.....	17
5.2	Weighting for waves three, four and five	17
5.3	Longitudinal weights waves three - five.....	18
5.4	Cross-sectional weights waves three – five.....	18
5.5	Self-completion weight	18
6	Dataset information	20
6.1	Updated dataset information – Waves three, four.....	20
6.2	Multi-coded and Dichotomous Variables – Wave three and beyond.....	21
6.3	Serial Numbering.....	22
6.4	Interview Outcome.....	23
6.5	Survey Module Identifiers.....	23
6.6	Sample Type	24
6.7	Age	24
6.8	Gender.....	24
6.9	Financial Variables	24
6.10	Dropped Variables.....	24
6.11	Missing Values.....	25
6.12	Coding and Editing.....	26
7	Appendices.....	27
7.1	Notes about particular variables in the Health module.....	27
7.2	Notes about particular variables in the Risk module (wave 5).....	31

1 Overview of the Survey

This User Guide provides detailed information pertaining to each of the first five waves of the ELSA data, all of which are available to download from the UK Data Archive [here](#).

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 sample was drawn from households that had previously responded to the Health Survey for England (HSE) between 1998 and 2004 and 2006.

The same group of respondents have been interviewed 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.

Although new topics can be introduced at different waves, every module has been reviewed to ensure that it will provide data that can measure change over time. This was achieved by repeating some measures exactly, by asking directly about change and by adopting questions to allow people to update or amend past responses. The information collected provides data about:

- Household and individual demographics
- Health – physical and psychosocial
- Work and pensions
- Income and assets
- Housing
- Cognitive function
- Social participation
- Expectations
- Walking speed

Certain waves contain one-off modules and questions and specific details of these are outlined in section 3.

ELSA is the result of collaboration between University College London (UCL), the Institute for Fiscal Studies (IFS), and NatCen 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 five waves of ELSA has been provided by the US National Institute on Aging, and a consortium of British Government departments, which are listed below:

- Department of Environment, Food and Rural Affairs;
- Department of Health
- Department for Transport
- Department for Work and Pensions;
- HM Treasury;
- HMRC (formerly Inland Revenue and HM Customs and Excise);
- Communities and Local Government (formerly Office of the Deputy Prime Minister);
- Office for National Statistics.

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).

1.2 Ethical Clearance

Ethical approval for all the ELSA waves was granted from the National Research and Ethics Committee. For further information see here: <http://www.nres.npsa.nhs.uk/>

1.3 Contact Details

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

2 Sample design

2.1 Sample Design

The ELSA sample has been designed to represent people aged 50 and over, living in private households in England. The sample is based on respondents who participated in the Health Survey for England (HSE). The original sample was selected from three years of HSE: 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. The HSE 1999 sample design also included a boost sample that represented ethnic minorities. Because of funding constraints, it was not possible to follow-up the boost sample and it was discarded. Together these three HSE years contained 23,132 responding households.

Households were removed from the HSE sampling frame for ELSA Wave 1 if it was known that there was no adult of 50 years or older in the household who had agreed to be re-contacted at some time in the future. Individuals in the remaining households provided the basis for the ELSA Wave 1 sample (11,578 households containing 18,813 eligible individuals). The Wave 1 Technical Report, available from the Data Archive, provides more details.

2.2 Refreshment samples

The aim of ELSA is to study a sample of people aged 50 and over. As the study progresses, the ELSA respondents get older and the sample effectively ages, therefore the youngest people need to be replaced as they are no longer represented.

At Wave 3 the ELSA sample was refreshed to make the sample representative of the youngest people, aged between 50 and 53^{Error! Bookmark not defined.}. The refreshment sample included new people from HSE 2001 - 2004 who were previously too young to join ELSA (or become an ELSA core member)¹ in 2002, but who were now aged 50 or over (i.e. people aged 50 to 53 and their partners). At Wave 4 the ELSA sample was further refreshed across a wider age range of 50 to 74 years. This refreshment sample included new people from HSE 2006 and their partners.

Households from these HSE survey years were selected for the study if at least one HSE interview was conducted with an eligible respondent in the household and they agreed to be re-contacted.

The Technical Report for Wave 3 contains further information about the first Refreshment Sample detailed above, and is available from the UK Data Archive. Please refer to the Methodology chapter of the Wave 4 Main Report or the Wave 4 Technical Report for more details about the Wave 4 refreshment sample.

2.3 Sample Types

Across all waves within households there were different types of individual who were eligible to be invited to take part in the study. The variable **SAMPSTAT** gives an indication of the sample type of

¹ A small number of respondents who were previously classified in the original sample taken from HSE 2001 as a 'younger partner' became eligible under the Wave 3 Refreshment sample rules to become an ELSA Core Member. Please see the Wave 3 Technical Report for further details.

the majority of respondents. This variable is supplemented by **FINSTAT**², which gives the final status of all respondents in the file.

The wave 1 interview provided the baseline for the study, with sample members who took part at this stage termed '**core members**'. Members recruited in the refreshment sample at waves 3 and 4 are also termed '**core members**'. Partners of core members were also eligible for the ELSA study. In waves 1 - 5 partners were divided into categories (summarised in Box 1 below).

Box 1 Summary of eligibility criteria for the ELSA partner interview waves 1- 5

Core partners (original sample only) were individuals who had been living within the household at the time of the HSE interview in 1998, 1999 or 2001 and were born on or before 29 February 1952. However they were *not* interviewed as part of Wave 1, so missing the baseline survey. Consequently, they were only approached by virtue of their being the partner of a core member.

Young partners were the cohabiting spouses or partners of core members, who were living within the household at the time of the HSE interview. Young partners from the original sample were still cohabiting with the core member at the time of the Wave 1 interview and were born *after* 29 February 1952. Young partners from the Wave 3 refreshment sample were still cohabiting with the core member at the time of the Wave 3 interview and were aged under 50 at 1st March 2006. Young partners from the Wave 4 refreshment sample were still cohabiting with the core member at the time of the Wave 4 interview and were aged under 50 at 1st March 2008.

Older partners (refreshment samples only) were the cohabiting spouses or partners of core members, who were living within the household at the time of the HSE interview and were still cohabiting with the core member at the time of the Wave 3 (or Wave 4) interview. They were aged over 53 at 1st March 2006 (for the Wave 3 refreshment sample) or aged 75 or above at 1st March 2008 (for the Wave 4 refreshment sample).

New partners were the cohabiting spouses or partners of core members at the time of the ELSA interview at Wave 1, 2, 3, 4 or 5 who had joined the household since the original HSE interview.

From wave five, all types of partner fell under the consolidated term '**partner**'.

All core members were eligible for subsequent waves of the study unless they had since died, had explicitly asked to not be re-contacted, or had moved out of Britain. Respondents remained eligible if they had moved to Wales or Scotland. If a partner was no longer living with a core member, through divorce, separation or widowhood, the same eligibility criteria applied until the partner was successfully interviewed for **one** subsequent wave of the study³; in order to understand their circumstances after this event had occurred. The only exception to this was if the partner had moved into an institution. In this case a subsequent interview was **not** pursued.

2.4 Proxies

A proxy interview was pursued if an eligible respondent was physically or cognitively impaired; in hospital; or temporarily in care for the whole of the fieldwork period. These criteria exclude members from wave 1 or any refreshment sample, as a first interview cannot be a proxy interview.

Interviewers were asked to identify a proxy informant as any responsible adult (aged 16 years or over) who knew enough about the respondent's circumstances to be able to provide information about them. Where possible, a close family member such as a partner or child fulfilled this role.

² Finstat may have a wave identifier in some files such as Finstatw4

³ However partners were only followed up to two consecutive waves, if after that time they had still not participated then they were removed from the eligible sample.

The CAPI program guided the interviewer through the proxy interview automatically. However, only a subset of questions were asked during a proxy interview. The proxy interview contained the following modules (the asterisked modules were not asked for all respondents – see below):

HD*	Household grid
ID	Individual demographics
HE	Health
WP	Work and Pensions
IA*	Income and Assets
HO*	Housing
FQ	Final questions and consents

All proxy informants completed ID, HE, WP and FQ. Some specific questions for proxies were included in the Health module, HE.

The modules on household demographics and housing were done as part of the proxy interview only when no-one else in the household was eligible for interview.

In cases where no-one else in the same financial unit was eligible for interview, the proxy interview included the Income and Assets module. If two proxy interviews were needed for a couple, the Income and Assets module would only appear in one of the interviews (asking about both their finances). For couples comprising of one person who was interviewed in person and another who needed a proxy interview, the former would automatically be asked the Income and Assets module on behalf of the couple. The question about whether they keep finances together or separate would not be asked.

2.5 Institution Interviews

Excluding those eligible for the wave 1 sample, an 'institution interview' was sought with core members who had moved from a private household at the first ELSA interview into a residential care home or similar establishment. Institution interviews could be carried out in person or by proxy. Routing has been developed within the interview specifically for people who have moved into institutions. Institution interviews were not pursued for partners of core members.

The content of these interviews is broadly the same as the main interview. People in institutions were asked different questions in the Housing and consumption module and were not asked some questions in other modules. There are also certain questions which are *only* asked of/about institutional respondents such as **MIREL** to **MIFIN** (within Individual Demographics).

People in institutions who were interviewed by proxy were also not asked questions that non-institution proxy informants were not asked (see Section 2.4).

The household identifiers for those who have moved out of a household and into an institution have been kept the same as for the household they have left, rather than being assigned a new one. This is only visible in the data where, for example, one respondent remains in the original household but another has moved into an institution. This decision was taken as some modules ask about the household as a whole and it would be confusing, especially in the Household Demographics module, if a respondent was being referred to who was now in a different household. Additionally, these modules may be asked of the respondent in the institution or the respondent still remaining in the previous household, depending on the circumstances.

3 Survey content

3.1 Development and piloting

For each wave, at least one pilot of the computer-assisted personal interview (CAPI) instrument, the self-completion questionnaire and associated documents were conducted around a year before the beginning of the mainstage fieldwork. All pilot respondents were drawn from a separate sample of households that participated in the ELSA Wave 1 pilots or were selected to become the refreshment pilot sample (from HSE 2000). The pilots tested the fieldwork procedure and interview content. Some new questions underwent cognitive testing.

For further information on specific waves' survey development please see the technical report.

3.2 Details of survey content - Overview

Overall, the intention of ELSA was to collect data about the same topics in each wave. However, at some waves there were some additions to the content of the interview to respond to new areas of enquiry. Some questions at certain waves were omitted, as it was decided that they did not need to be asked at every wave (see sections below for more specific details). Furthermore, several elements of the questionnaire were amended to take account of responses given at the previous waves. At each wave all interviews comprised of a personal face-to-face CAPI interview and a self-completion questionnaire.

The ELSA program allowed flexibility in administering the interview. Respondents could be interviewed individually, or in houses with more than one eligible respondent, interviewed at the same time (in a single session) using concurrent interviewing techniques. In a concurrent session the same block of questions was asked alternately of each person. Concurrent interviews tended to be quicker than two separate individual interview sessions, and were generally more convenient for respondents.

ELSA uses a technique called 'feeding forward' data. It is a technique that feeds responses individuals made at earlier interviews to either aid recall and/or improve consistency of responses across interviews. ELSA uses proactive dependent interviewing: i.e. earlier responses are provided to the individual before they respond to a question (rather than reactive dependent interviewing when the earlier response is used after the individual has responded as a form of validation). Dependent interviewing was used to check some information collected during previous waves (or HSE), to determine changes in status, and to control routing within the questionnaire.

There were various modules each covering a different area of interest. The content and major routing of each module at waves 1 - 5 is described below. Although interviews tended to follow the same module order, some flexibility was given to the interviewer. For example, the timed walk could be administered at any time after the Health module, and it was possible for interviewers to skip the Income and Assets or Housing modules if it was more convenient to do them at another time.

Five of the modules (Cognitive Function, Expectations, Effort and Reward, Psychosocial Health and Final Questions) form the 'private modules' block. Wherever possible, these modules were administered with no other household members present. If two respondents were being interviewed concurrently, whilst the first respondent was being asked the private block, the second responding

individual was asked to fill in the self-completion in a separate room. The two respondents then switched places.

3.3 Details of survey content - Wave one

The wave one questionnaire modules are listed in the table below in the order programmed in the CAPI interview. In general, all subsequent waves contain these modules; any major changes are listed in sections 3.4 – 3.7.

Module	Description
Household Demographics (“DH”)	Answered by one person on behalf of the household with answers copied onto the data for other household members. Collects basic demographic information on everyone in the household. Uses feed forward data from HSE.
Individual Demographics (“DI”)	Answered by all. Included details on: legal marital status; living children (including adopted, fostered or stepchildren), grandchildren & great-grandchildren; number of siblings; childhood circumstances
Health (“HE”)	Answer by all. Included: self-reported general health; long-standing illness or disability; eyesight and hearing; specific diagnoses and symptoms; pain; difficulties with activities and instruments of daily living; health behaviours. Those aged 60+ were asked about falls and fractures.
Social Participation (“SP”)	Answered by all. Covered frequency respondents participated in social activities, limitations to these activities, their care giving and public transport usage.
Work & Pensions (“WP”)	Answered by all. Included: current work activities; current or past pensions. If retired and receiving a pension details collected on the pension, including amounts.
Income & Assets (“IA”)	If a couple’s finances are kept together one of them answers. If separate, both answer. Details included: wages; state & private pensions; annuity income; state benefits; financial & non-financial assets held, income from these assets; regular & one-off transfers from non-household members.
Housing (“HO”)	Answered by one person on behalf of the household. Included: current housing situation (e.g. size, quality); housing-related expenses; ownership of durables and cars; expenditure on food. House owners were asked about value of property, mortgage details etc
Cognitive Function (“CF”)	Answered by all, start of ‘private modules’ block. Measured: cognitive function including memory, speed, mental flexibility & numeracy.
Expectations (“EX”)	Answered privately by all. Measured expectations regarding certainty of future events and financial decision-making.
Psychosocial Health (“PS”)	Answered privately by all. Measured respondents’ views on life including a mental health scale and attitudes towards ageing.
Final Questions (“FQ”)	Answered privately by all. Included: demographic information; stable address contact & consent to obtain health and economic data from administrative sources.
Walking Speed (“MM”)	Completed by those aged 60+ when safe to do so. Respondents walked distance of 8 feet (2.4m) twice, timed using a stopwatch.
Core Self-completion (“SC”)	Completed either during other members private blocks or after and returned by post.

3.4 Details of survey content – Wave two

Generally, the modules and questions remained largely similar as wave one (see 3.3 above). An overview of wave 2 and any major changes are listed in the table below.

Module	Description including changes from wave one
Household Demographics (“DH”)	<i>No changes.</i> Answered by one person on behalf of the household with answers copied onto the data for other household members. Collects basic demographic information on everyone in the household. Uses feed forward data from wave one.
Individual Demographics (“DI”)	<i>No changes.</i> Answered by all. Included details on: legal marital status; living children (including adopted, fostered or stepchildren), grandchildren & great-grandchildren; number of siblings; childhood circumstances.
Health (“HE”)	<i>Limited changes.</i> Answer by all. Included: self-reported general health; long-standing illness or disability; eyesight and hearing; specific diagnoses and symptoms; pain; difficulties with activities and instruments of daily living; health behaviours. Aged 60+ were asked about falls and fractures. <u>New questions</u> added regarding quality of healthcare received for particular health conditions.
Social Participation (“SP”)	<i>Some changes.</i> Answer by all. Shorter than wave one. Only asked about use of public transport. Questions on ‘cultural capital’ were asked in the self-completion instead. Questions on caring moved to a new module “Effort and Reward” (see below).
Work & Pensions (“WP”)	<i>No changes.</i> Answered by all. Included: current work activities; current or past pensions. If retired and receiving a pension details collected on the pension, including amounts.
Income & Assets (“IA”)	<i>No changes.</i> If a couple’s finances are kept together one of them answers. If separate, both answer. Questions included: wages; state & private pensions; annuity income; state benefits; financial & non-financial assets held, income from these assets; regular & one-off transfers from non-household members.
Housing (“HO”)	<i>Limited changes.</i> Answered by one person on behalf of the household. Included: current housing situation; housing-related expenses; ownership of durables and cars; expenditure on food. House owners asked about value of property, mortgage details. <u>New questions</u> added on durable purchases in last 2 years, money spent on clothes in past 4 weeks, gifts, leisure activities and fuel expenditure.
Cognitive Function (“CF”)	<i>Limited changes.</i> Answered by all, start of ‘private modules’ block. Measured: cognitive function including memory, speed & mental flexibility. Questions to assess respondents’ literacy skills replaced wave one numeracy questions.
Expectations (“EX”)	<i>Limited changes.</i> Answered privately by all. Expectations about certainty of future events and financial decision-making. <u>New questions</u> added on expectations of moving home in the future and those aged 60+ expectations of moving to a nursing home in next 5 years.
Effort and Reward (“ER”)	**NEW W2 MODULE** Answered privately by all. Assesses respondents’ motivations for voluntary work, caring for others & the relationship between effort and reward.
Psychosocial Health (“PS”)	<i>Some changes.</i> In place of wave one’s questions this module assesses symptoms of depression, based on CES-D depression scale. Respondents reporting depression symptoms asked about quality of healthcare received for depression. Additional psychosocial questions in W2 self-completion.
Final Questions (“FQ”)	<i>No changes.</i> Answered privately by all. Included: demographic information & stable address contact. Verbal reminder if consent already given, otherwise asked for consent to obtain health and economic data from administrative sources.
Walking Speed (“MM”)	<i>No changes.</i> Completed by those aged 60+ when safe to do so. Respondents walked distance of 8 feet (2.4m) twice, timed using a stopwatch.
Core Self-completion (“SC”)	In addition to wave one’s questions, new questions added on social isolation, life satisfaction and alcohol consumption.

3.5 Details of survey content – Wave three

Again the modules and topic covered in wave 3 remained largely the same as wave 1 and 2 (see sections 3.3 and 3.4). An overview of wave 3 and any major changes are listed in the table below.

Module	Description including changes from wave two
Household Demographics (“DH”)	<i>No changes.</i> Answered by one person on behalf of the household with answers copied onto the data for other household members. Collects basic demographic information on all in the household. Uses feed forward data from wave one and HSE for refreshment.
Individual Demographics (“DI”)	<i>No changes.</i> Answered by all. Included details on: legal marital status; living children (including adopted, fostered or stepchildren), grandchildren & great-grandchildren; number of siblings; childhood circumstances.
Health (“HE”)	<i>Limited changes.</i> Answer by all. Included: self-reported general health; long-standing illness or disability; eyesight and hearing; specific diagnoses and symptoms; pain; difficulties with activities and instruments of daily living; health behaviours. Aged 60+ were asked about falls and fractures. <u>New questions</u> added regarding dental health.
Social Participation (SP)	<i>Some changes.</i> Answer by all. Questions on use of public transport. <u>New questions</u> on use of services e.g. transport provided by hospitals/day centres or meals on wheels.
Work & Pensions (“WP”)	<i>No changes.</i> Answered by all. Included: current work activities; current or past pensions. If retired and receiving a pension details collected on the pension, including amounts.
Income & Assets (“IA”)	<i>No changes.</i> If a couple’s finances are kept together one of them answers. If separate, both answer. Questions included: wages; state & private pensions; annuity income; state benefits; financial & non-financial assets held, income from these assets; regular & one-off transfers from non-household members.
Housing (“HO”)	<i>No changes.</i> Answered by one person on behalf of the household. Included: current housing situation; housing-related expenses; ownership of durables and cars; expenditure on food, clothes, gifts and leisure activities. House owners asked about value of property, mortgage details.
Cognitive Function (“CF”)	<i>Limited changes.</i> Answered by all, start of ‘private modules’ block. Measured: cognitive function including memory, speed & mental flexibility. Questions to assess respondents’ literacy skills have been dropped for this wave.
Expectations (“EX”)	<i>Limited changes.</i> Answered privately by all. Expectations about certainty of future events and financial decision-making including moving home in future and for those aged 60+, likelihood of moving to a nursing home in next 5 years.
Effort and Reward (“ER”)	**MODULE INTRODUCED AT WAVE TWO** <i>No changes.</i> Answered privately by all. Assesses respondents’ motivations for voluntary work, caring for others & the relationship between effort and reward.
Psychosocial Health (“PS”)	<i>Some changes.</i> Questions how respondent views their life across a variety of dimensions. Questions about when perceived middle age ends and old age starts, in wave one, have been added back in.
Final Questions (“FQ”)	<i>Limited changes.</i> Answered privately by all. Included: demographic information & stable address contact. Verbal reminder of consent if already given, otherwise asked for consent to obtain health and economic data from administrative sources.
Walking Speed (“MM”)	<i>No changes.</i> Completed by those aged 60+ when safe to do so. Respondents walked distance of 8 feet (2.4m) twice, timed using a stopwatch.
Core Self-completion (“SC”)	<i>Limited changes.</i> Included: respondents’ quality of life, social participation, control at work, life satisfaction, social networks and alcohol, fruit and vegetable consumption.
Health and Work Self-completion	**NEW W3 MODULE** Given to sub-sample of W3 respondents (third sample received work, third sample received health, third only received core). Health – rating situation on 5-point scale including mobility, pain, sleep, cognition and depression. Work – rate hypothetical people in various scenarios on 5-point scale

3.6 Details of survey content – Wave four

Once more, the modules and topic covered in wave 4 remained largely the same as wave’s 1-3 (see sections 3.3, 3.4 and 3.5 respectively). An overview of wave 4 and any major changes are listed in the table below.

Module	Description – changes from previous waves
Household Demographics (“DH”)	<i>No changes.</i> Answered by one person on behalf of the household with answers copied onto the data for other household members. Collects basic demographic information on all in the household. Uses feed forward data from wave three and HSE for refreshment.
Individual Demographics (“DI”)	<i>No changes.</i> Answered by all. Included details on: legal marital status; living children (including adopted, fostered or stepchildren), grandchildren & great-grandchildren; number of siblings; childhood circumstances.
Health (“HE”)	<i>Limited changes.</i> Answer by all. Included: self-reported general health; long-standing illness or disability; eyesight and hearing; specific diagnoses and symptoms; pain; dental health; difficulties with activities and instruments of daily living; health behaviours. Aged 60+ were asked about falls and fractures. <u>New questions</u> for females about women’s health and the menopause.
Social Participation (SP)	<i>Limited changes.</i> Answer by all. Questions on use of public transport; use of services e.g. transport provided by hospitals/day centres and meals on wheels.
Work & Pensions (“WP”)	<i>No changes.</i> Answered by all. Included: current work activities; current or past pensions. If retired and receiving a pension details collected on the pension, including amounts.
Income & Assets (“IA”)	<i>No changes.</i> If a couple’s finances are kept together one of them answers. If separate, both answer. Questions included: wages; state & private pensions; annuity income; state benefits; financial & non-financial assets held, income from these assets; regular & one-off transfers from non-household members.
Housing (“HO”)	<i>No changes.</i> Answered by one person on behalf of the household. Included: current housing situation; housing-related expenses; ownership of durables and cars; expenditure on food, clothes, gifts and leisure activities. House owners asked about value of property, mortgage details.
Cognitive Function (“CF”)	<i>Limited changes.</i> Answered by all, start of ‘private modules’ block. Measured: cognitive function including memory & speed. Questions on numeracy re-introduced.
Expectations (“EX”)	<i>Limited changes.</i> Answered privately by all. Expectations about certainty of future events and financial decision-making including moving home in future and for those aged 60+, likelihood of moving to a nursing home in next 5 years.
Effort and Reward (“ER”)	**MODULE INTRODUCED AT WAVE TWO** <i>No changes.</i> Answered privately by all. Assesses respondents’ motivations for voluntary work, caring for others & the relationship between effort and reward. <u>New questions</u> at wave four on provision of care and use of respite care services.
Psychosocial Health (“PS”)	<i>Limited changes.</i> Asks how respondent views their life across a variety of dimensions.
Final Questions (“FQ”)	<i>Limited changes.</i> Answered privately by all. Included: demographic information & stable address contact. Verbal reminder of consent if already given; otherwise asked for consent to obtain health and economic data from administrative sources.
Walking Speed (“MM”)	<i>No changes.</i> Completed by those aged 60+ when safe to do so. Respondents walked distance of 8 feet (2.4m) twice, timed using a stopwatch.
Core Self-completion (“SC”)	<i>Limited changes.</i> Included: respondents’ quality of life, social participation, control at work, life satisfaction, social networks and alcohol, fruit and vegetable consumption.
Health and Work Self-completion	REMOVED FOR WAVE FOUR. WAVE THREE ONLY.

3.7 Details of survey content – Wave five

The modules and topics covered in wave 5 remained largely the same as wave's 1-4 (see sections 3.3 - 3.6). An overview of wave 5 and any major changes are listed in the table below.

Module	Description – changes from previous waves
Household Demographics (“DH”)	<i>No changes.</i> Answered by one person on behalf of the household with answers copied onto the data for other household members. Collects basic demographic information on all in the household. Uses feed forward data from wave four.
Individual Demographics (“DI”)	<i>No changes.</i> Answered by all. Included details on: legal marital status; living children (including adopted, fostered or stepchildren), grandchildren & great-grandchildren; number of siblings; childhood circumstances.
Health (“HE”)	<i>Limited changes.</i> Answer by all. Included: self-reported general health; long-standing illness or disability; eyesight and hearing; specific diagnoses and symptoms; pain; dental health; difficulties with activities and instruments of daily living; health behaviours. Aged 60+ were asked about falls and fractures. Removal of questions on sleep and balance.
Social Participation (SP)	<i>Limited changes.</i> Answer by all. Questions on use of public transport; use of services e.g. transport provided by hospitals/day centres and meals on wheels. New questions on views on the local area.
Work & Pensions (“WP”)	<i>Limited changes.</i> Answered by all. Included: current work activities; current or past pensions. If retired and receiving a pension details collected on the pension, including amounts. Some wave five questions omitted e.g. employer provision of accommodation and arrangements for those with a disability at work.
Income & Assets (“IA”)	<i>No changes.</i> If a couple's finances are kept together one of them answers. If separate, both answer. Questions included: wages; state & private pensions; annuity income; state benefits; financial & non-financial assets held, income from these assets; regular & one-off transfers from non-household members.
Housing (“HO”)	<i>Limited changes.</i> Answered by one person on behalf of the household. Included: current housing situation; housing-related expenses; ownership of durables, cars and pets; expenditure on food, clothes, gifts and leisure activities.
Cognitive Function (“CF”)	<i>Limited changes.</i> Answered by all, start of 'private modules' block. Measured: cognitive function including memory & speed. Questions on literacy re-introduced.
Expectations (“EX”)	<i>Limited changes.</i> Answered privately by all. Expectations about certainty of future events and financial decision-making including moving home in future and for those aged 60+, likelihood of moving to a nursing home in next 5 years.
Effort and Reward (“ER”)	**MODULE INTRODUCED AT WAVE TWO** <i>No changes.</i> Answered privately by all. Assesses respondents' motivations for voluntary work, caring for others, relationship between effort and reward and provision of care and use of respite care services.
Psychosocial Health (“PS”)	<i>No changes.</i> Asks how respondent views their life across a variety of dimensions.
Risk Module	**NEW W5 MODULE** Answered by pre-selected sub-sample only. Module comprises 22 games respondents play on a laptop as a self-completion CASI interview. Games are designed to measure respondents' attitudes towards financial risk taking and willingness to accept delay in receiving money. Respondents won actual money from playing the games, which was sent to them as a cheque after the interview.
Final Questions (“FQ”)	<i>Limited changes.</i> Answered privately by all. Included: demographic information & stable address contact. Verbal reminder of consent if already given; otherwise asked for consent to obtain health and economic data from administrative sources.
Walking Speed (“MM”)	<i>No changes.</i> Completed by those aged 60+ when safe to do so. Respondents walked distance of 8 feet (2.4m) twice, timed using a stopwatch.

Core Self-completion (“SC”)	<i>Limited changes.</i> Included: respondents’ quality of life, social participation, control at work, life satisfaction, social networks and alcohol, fruit and vegetable consumption. <u>New questions:</u> measuring well-being and personality, experience of discrimination and the respondent’s religion (if any).
------------------------------------	---

4 Fieldwork and response rate

4.1 Fieldwork - Overview

The fieldwork process for each wave of ELSA has been very similar. Eligible individuals satisfying a number of criteria (see section 2.3) were sent an advance letter inviting them to take part. Interviewers then either telephoned or visited the households to explain the study, and interview willing individuals straight away, or to make appointments to call back at a convenient time. The main survey comprised of a personal face-to-face interview and a self-completion questionnaire. Waves 2 and 4 also included a separate Nurse visit; the data from both these elements of ELSA is available from the UK Data archive.

At waves 1 - 4, before starting work all new interviewers underwent a two-day personal briefing by a researcher. At wave 5 the two-day briefing for new interviewers was reduced slightly by half a day. Those interviewers who had taken part in a previous wave of ELSA underwent a one day refresher briefing. The briefings covered all fieldwork procedures including training on how to administer the assessments (walking speed and cognitive function), fully explained the documents needed for the study and provided an introduction to all questions within the CAPI interview. Interviewers were provided with written study guidelines to reinforce the briefing.

Addresses within the same postcode sectors were clustered and issued to interviewers. Before starting to carry out their visits, all interviewers were instructed to report to the police station local to where they were working and were expected to show a copy of the ELSA advance letter, leave their name and NatCen's contact details and explain how long they would be carrying out interviews in the area.

For further information on the specific fieldwork procedures at each wave please see the technical reports on the Economic and Social Data Service website [here](#) .

4.2 Fieldwork - Tracing Movers

To minimise the sample attrition caused by respondents moving address, procedures were in place to track respondents who move between waves, or since their HSE interview, to ensure that the more mobile sections of the ELSA sample are not lost.

If the whole household had moved since the last interview, or a core member who had consented to be re-contacted in future waves had moved away, interviewers were directed to try the following possible routes to trace movers:

- Attempt telephone contact with the respondent
- Attempt to find a follow-up address
- Approach the present occupants, neighbours, or friends to obtain the new address
- (For waves 3 - 5) approach the person(s) living at the 'stable address' provided previously by the respondent – Since wave 2 respondents had been asked to give the name and contact details of someone who could be contacted if they moved
- Consider phone books, electoral register, local shops, letting agency, estate agent, post office

A 'mover letter' was offered if interviewers identified a member of the public who was aware of the core member's new address but was reluctant to reveal it to the interviewer. This letter, which was

forwarded with a pre-paid envelope by the member of the public who had been identified, asked the core member to contact the office with their new address.

At wave 3 the DWP assisted with the tracing of core members using their state pension databases. The respondent's name, date of birth and address were provided to DWP and they matched this to their databases in order to identify the most up-to-date contact details. If a new address was found, an advance letter was sent to the respondent.

At waves 3, 4 and 5 we were able to trace respondents through the National Health Service Central Register (NHSCR). For those respondents who we could not trace and who agreed to have their data linked to the NHSCR we were able to obtain health authority information, which enabled us to write to the health authority and ask them to pass on an advance letter to the respondent via their GP.

4.3 Response rates at wave one

Wave 1 fieldwork ran from March 2002 to March 2003.

The ELSA wave one fieldwork produced 12,100 productive individual interviews. Of these, 11,392 were core members, 636 younger partners and 72 new partners. Of the core sample member interviews, 204 were partial responses and 158 were proxy responses.

Of the 11,578 households that were issued for wave one, the interviewers made contact with 94% of them. Nearly two-thirds of non-contacts had moved since the HSE interview and could not be traced. Separate analyses show that around a tenth of issued households had moved house by the wave one fieldwork period. Two-thirds of the movers were eventually traced.

A responding household is defined as one where at least one eligible person was interviewed. Using this criterion, a household response rate of 70% was achieved. The majority of non responding households refused to participate (22% of the eligible sample of households).

A small percentage of respondents within the responding households were ineligible (3%). Of the remaining sample of individuals within responding households, a response rate of 96% was achieved. Non-response within households was almost always because of refusal to take part.

The overall response rate for individuals (calculated using the total number of eligible individuals within all issued households) was 66%.

More information about the response rates is provided in the Wave 1 Technical Report.

4.4 Response rates at wave two

Wave 2 fieldwork ran from June 2004 - July 2005. Wherever possible, interviewers were assigned to the same households they had interviewed at the previous wave.

The ELSA wave 2 fieldwork produced 9,432 productive interviews. Of these, 8780 of these were with core sample members. Of all core sample member interviews, 38 people only completed a partial interview in person and 92 were interviewed by proxy. In addition 57 productive interviews were conducted with core partners, 501 with younger partners, and 94 with new partners.

The interviewer made contact with 97% of the households that were issued for ELSA Wave 2 (the household contact rate). The largest component (77%) of non-response was a result of refusals.

Though many people who had moved were traced from their Wave 1 residence, 11% of non responders were individuals who could not be found. This is slightly higher than Wave 1, where those who had moved and could not be traced constituted 10% of issued Wave 1 non-respondents. 82% of those who completed a Wave 1 interview and were eligible for a Wave 2 interview as an ELSA 'core member' took part in the survey.

More information about the response rates is provided in the Wave 2 Technical Report.

4.5 Response rates at wave three

Fieldwork for wave 3 of ELSA began in May 2006 and spanned 15 months, finishing in August 2007. Wherever possible, interviewers were assigned to the same households they had interviewed at the previous wave.

The deposit of the wave 3 data contains, from the original sample, 7535 core members, 91 core partners, 312 younger partners, 74 'new' partners (found at previous waves) and 26 new partners (found at this Wave). For the refreshment sample, this deposit contains 1275 core members, 142 older partners, 295 younger partners and 21 new partners (found at this Wave).

More information about the response rates is provided in the wave 3 Technical Report.

4.6 Response rates at wave four

Fieldwork for wave 4 of ELSA began in May 2008 and spanned 14 months, finishing at the end of July 2009. Wherever possible, interviewers were assigned to the same households they had interviewed at the previous wave.

The wave 4 deposited data contains, from the original sample, 6,623 core members, 101 core partners, 276 younger partners, 98 'new' partners (found at previous Waves) and 21 new partners (found at this Wave). From the wave 3 refreshment sample, this deposit contains 972 core members, 12 core partners, 106 older partners, 226 younger partners, 14 'new' partners (found at Wave 3) and 11 new partners (found at Wave 4). For the wave 4 refreshment sample, this deposit contains 2,291 core members, 165 older partners, 119 younger partners and 15 new partners (found at this Wave).

This deposit contains data from a total of 11,050 respondents.

More information about the response rates is provided in the Wave 4 Technical Report.

4.7 Response rates at wave five

Fieldwork for wave 5 of ELSA began in June 2010 and spanned 13 months, finishing at the beginning of July 2011. Wherever possible, interviewers were assigned to the same households they had interviewed at the previous wave.

The Wave 5 deposited data contains, from the original sample 6,242 core members, 125 core partners, 281 younger partners, 116 'new' partners (found at previous Waves) and 9 new partners (found at this Wave). From the Wave 3 refreshment sample, this deposit contains 936 core members, 17 core partners, 102 older partners, 217 younger partners, 26 'new' partners (found at previous waves) and 10 new partners (found at this wave). For the Wave 4 refreshment sample, this deposit contains 1,912 core members, 27 core partners, 127 older partners, 101 younger partners, 14 'new' partners found at Wave 4 and 12 new partners (found at this Wave).

This deposit contains data from a total of 10,274 respondents.

More information about the response rates is provided in the forthcoming Wave 5 Technical Report.

5 Weighting

Where possible we recommend that analysis be conducted on weighted data since this will help to minimise bias from differential non-response amongst key sub groups. With a longitudinal survey the potential for bias increases with each successive wave. Now that the panel has experienced six successive waves of attrition including non-response at HSE, refusals to be re-interviewed post-HSE and non-response at ELSA waves 1 to 5, it is imperative to adjust for differential non-response.

5.1 Weighting for waves one and two

The variable in the dataset to be used for weighting at wave one is **W1wgt** and the variable for wave two is **W2wgt**. Weights were calculated for core sample members only (including proxy and partial interviews). All other non-sample individuals that were interviewed (new and younger partners) have a weight of zero.

When running weighted analyses, researchers should remember to exclude these respondents (with zero weights) from the un-weighted base, if quoted. The data for partners can be used as supplementary information for core members⁴.

For further information on wave one and two weighting, please see the technical report (available on the data archive website [here](#)).

5.2 Weighting for waves three, four and five

The weighting for waves three, four and five can be divided for analysis cross-sectionally or longitudinally:

- Cross-sectional analysis uses data collected only at a particular wave.
- Longitudinal analysis involves data collected from more than one wave for the purposes of analysing change.

For waves three, four and five cross-sectional and longitudinal weights support these two different objectives. As with waves one and two, weights have been calculated for core members living in private households in England who responded to the survey (including those interviewed by proxy and those giving partial interviews). Cross-sectional weights are zero for all other respondents, including core members living in institutions and partners of core members (i.e. core, new, younger and older partners). To receive a longitudinal weight, respondents must have taken part in all five waves of ELSA; all other respondents are 'system missing'.

When running weighted analyses, researchers should remember to exclude these respondents (with zero or 'system missing' weights) from the un-weighted base, if quoted. The data for partners can be used as supplementary information for core members.

⁴ If non-core sample members are to be analysed they should be analysed unweighted.

5.3 Longitudinal weights waves three - five

These weights are only defined for the subset of cases who have taken part in all waves, up to and including the wave in question.

At each wave, the fully responding core members are re-weighted to take account of respondents at the previous wave that were lost through refusal at the current wave or through some other form of sample attrition. Core members from the original sample who returned to the study having missed a wave do not, therefore, have a non-zero longitudinal weight.

WAVE	Description
WAVE THREE	Variable name: W3LWGT Calculated for the set of 7,168 core members who have responded to all three waves of ELSA, and remain living in private households.
WAVE FOUR	Variable name: W4LWGT Calculated for the set of 5,971 core members who have responded to all four waves of ELSA, and remain living in private households.
WAVE FIVE	Variable name: W5LWGT Calculated for the set of 5,262 core members who have responded to all five waves of ELSA, and remain living in private households.

5.4 Cross-sectional weights waves three – five

The cross sectional weights for wave 3-5 are made up of the following groups:

- Longitudinal sample members (i.e. those given a longitudinal weight)
- Cohort 3 core members (new entrants at wave three)
- Cohort 4 core members (new entrants at wave four)
- Wave non-responders (those core members from a previous cohort who return to the study having missed a previous wave)

The cross-sectional weights were calculated separately for each cohort. In each case the weighting aimed to adjust for differences in the propensity to respond amongst key sub-groups. The final step in the calculation of the cross-sectional weight involved computing a scaling factor to ensure that the original sample (Cohort 1) and refreshment samples (Cohorts 3 and 4) were represented in the same proportions, with respect to age, in which they appear in the population (based on the 2008 household population estimates provided by the Office for National Statistics).

WAVE	VARIABLE NAME
WAVE THREE	Variable name: W3XWGT
WAVE FOUR	Variable name: W4XWGT
WAVE FIVE	Variable name: W5XWGT

5.5 Self-completion weight

Self-completion weights were created at wave 4 and 5 for the core members whose self-completion questionnaire were received with the majority of questions answered. These weights (**W4SCWT** and

W5SCWT respectively) built upon the Wave 4 and 5 cross-sectional weights and aimed to adjust for differential non-response amongst respondents.

6 Dataset information

This User Guide refers to the core datasets that has been archived for ELSA Waves one to five. These datasets lists cases at an individual level.

Data for ELSA respondents collected during HSE (referred to as “wave 0” data) and ELSA waves 1-5 can be used for longitudinal analysis. All this data is available from the Data Archive.

The ELSA core datasets contains data in the following order:

- Key variables not in the questionnaire (e.g. serial number, outcome codes)
- Variables in the questionnaire (in the order they appear in the CAPI interview). A small number of additional 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, other derived variables and variables fed forward from previous waves).

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.

6.1 Updated dataset information – Waves three, four

WAVE THREE DATASET INFORMATION

There are 9771 cases in the Phase 2 deposit. This differs from the number of cases deposited at Phase 1 and, although we have asked that all copies of the Phase 1 data are deleted, we still think it is useful to illustrate the differences in the number of cases between the two datasets.

There are five cases that were in the Phase 1 data but are not in Phase 2. Two of these cases asked that their data be deleted, while the remainder had their outcome codes changed as a result of data cleaning and reconciliation (and no longer fell under the criteria for archiving). There are 58 cases that were not in the Phase 1 data but are in Phase 2. 47 of these cases had an institution interview and the remaining 11 cases had problems with their data which have now been resolved.

There are three cases that have a value in SAMPSTA which may need clarification: 101166, 101375 and 107830. These are partners who had left the household but subsequently moved back in, but they were not issued as part of the Wave 3 sample as they were not eligible for interview in their own right. The ELSA Technical Reports provide more details about eligibility and sampling. Also note that case 107830 was initially thought to be new partner and were assigned an IDAUNIQ of 121362 at Phase 1.

There are 103 more refreshment sample core members (FINSTAT="C3CM") in the Phase 2 data than in the Phase 1. This is due to the inaccuracy since discovered with the sampling algorithm.

One case has a changed value of FINSTAT in Version 2 of the Phase 2 data. This is because their date of birth was given as being in 1953 in their HSE interview (meaning they were originally given a FINSTAT of "C3CM") but it has been noted that in their Wave 3 interview they gave their date of birth as being in 1959 (meaning they have now been given a FINSTAT of "C3YP").

WAVE FOUR DATASET INFORMATION

There are a total of 11,050 cases in the Phase 2 deposit. This differs from the number of cases deposited at Phase 1 and, although we are asking analysts to delete all copies of this data that they may hold, we still think it is useful to illustrate the differences in the number of cases between the two datasets.

There are two cases which were in the Phase 1 data but are not in Phase 2 (IDAUNIQ= 110861 and 121418). The first has been excluded as they refused to answer almost all of the questions in the interview, and have now been classed as being unproductive. The second has been excluded as they were in fact an erroneous duplicate interview.

There are seven cases which have a different IDAUNIQ in Phase 2 in order that they match their values of IDAUNIQ in Wave 3 Phase 2. These cases used to be identified as 150081, 150328, 150535, 150583, 150597, 150870, and 151487 and are now identified as 701951, 700754, 700108, 704339, 702643, 700564 and 701668 respectively.

There are 192 cases that were not in the Phase 1 data but are in Phase 2. 68 of these cases had an institution interview, while 119 cases had problems with their feedforward data. A further five cases (IDAUNIQ= 108592, 120652, 118292, 112695 and 112832) have been added as we have been able to resolve previous issues with the processing and delivery of their data.

There are also some cases which have a different FINSTAT in the Phase 2 data than they did in the Phase 1 data. Eight cases who had a FINSTAT of "C1YP" (or similar) in Phase 1 have been re-classified as "C3CP" in Phase 2. These were cases which would have been re-classified as "C3CM" at Wave 3, but were unproductive at that Wave.

There are 92 cases that were erroneously classified as being "C1CM" in Phase 1 who have been correctly re-classified to "C3CM" in Phase 2.

Two further cases have a changed value of FINSTAT in the Phase 2 data. One is the same case whose status was changed from "C3CM" to "C3YP" at Wave 3 – they have now been changed from the former to the latter between Phase 1 and Phase 2 of Wave 4. The other (IDAUNIQ=161873) has changed FINSTAT from a "C4CM" to a "C4NP4" as they were erroneously assigned the former at Phase 1.

6.2 Multi-coded and Dichotomous Variables – Wave three and beyond

A major change in the way that ELSA data is archived occurred from wave 3 and beyond regarding multi-coded variables. A variable is defined as multi-coded if more than one response could be given e.g. favourite activities are reading *and* going to the cinema. For Wave 3, and beyond, such variables will be archived as dichotomous variables, with each answer option being coded as 'mentioned' or 'not mentioned'. Please note that although the way the data is presented is different, no change was made to the way the questions were asked during the interview.

For example, data from the series of questions **EXRELE01 – EXRELE10** ("Whether having too little money stops them from doing things") was outputted as multi-coded variables at previous waves. Therefore, there was a variable for the respondent's first response to the question, second response and so on. Each variable had the following answer options:

1 Buy your first choices of food items

- 2 Have family and friends round for a drink or meal
- 3 Have an outfit to wear for social or family occasions
- 4 Keep your home in a reasonable state of decoration
- 5 Much better off
- 6 Pay for fares or other transport costs to get to and from places you want to go
- 7 Buy presents for friends or family once a year
- 8 Take the sorts of holidays you want
- 9 Treat yourself from time to time
- 96 None of these.

As of Wave 3, the data was outputted as dichotomous variables so there is a variable for each response category for that question which shows whether or not the respondent gave that response.

The variables in the data are:

EXRELEFO: Things interviewee is not able to do because of having too little money: Buy your first choices of food items.

EXRELEME: Things interviewee is not able to do because of having too little money: Have family and friends round for a drink or meal.

EXRELEOU: Things interviewee is not able to do because of having too little money: Have an outfit to wear for social or family occasions.

...

EXREL96: Things interviewee is not able to do because of having too little money: None of these.

These questions now have much simpler answer options of:

0 Not mentioned (*i.e. respondent did not give this answer*)

1 Mentioned (*i.e. respondent did give this answer*).

The new dichotomous variables have been given different names in order to distinguish them from their multi-coded equivalents at previous waves. The variable names now have a common 'stem' for each question (**EXRELE** in the example above), and then two or three letters to indicate the answer option. Where answer options are the same between questions, the same two or three letters have been used each time for consistency.

6.3 Serial Numbering

Constant Serial Numbers

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. Each respondent will have a unique value for **IDAUNIQ**, which will remain constant.

Wave-specific household serial number

The five digit household serial number (for example, **for wave 5 is called IDAHHW5**) was randomly generated for the archived dataset, and does not relate to the serial number used during interviewing. The dataset 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. This number is used for several variables. For example, it is used to identify the respondent's partner within the household (see variable **CPID**), or to identify the person in the household that responded to particular questions such as the Income and Assets module (see variable **IAPID**).

6.4 Interview Outcome

WAVE ONE

Individual and household level interview outcomes are given in **Indoc** and **Rthhout** respectively. The individual outcome is a two-digit code. The first digit indicates whether the interview was full (1) or partial (2). The second digit indicates whether the interview was in person (1) or by proxy (3). The household outcome is a three-digit code. The first digit indicates whether all eligible residents were interviewed (1) or some interviewed (2).

WAVE TWO

Individual and household level interview outcomes are given in **W2INDOUT** and **W2HHOUT** respectively. The individual outcome is a two-digit code. The first digit indicates whether the interview was full (1) or partial (2). The second digit indicates whether the interview was in person (1) or by proxy (3). The household outcome is a three-digit code. The first digit indicates whether all eligible residents were interviewed (1) or some interviewed (2).

WAVE THREE

Individual and household level interview outcomes are given in **W3INDOUT** and **W3HHOUT** respectively. The individual outcome is a two-digit code while the household outcome has three digits. There are three additional outcome codes for the self-completion questionnaires – **OUTSCW3** for the core self-completion, and **OUTSHW3** and **OUTSWW3** for the health and work vignette self-completions respectively. For the latter two outcomes, if a respondent has a value of '-1' it means they were not part of the sub-sample selected to receive that questionnaire.

WAVE FOUR

Individual and household level interview outcomes are given in **OUTINDW4** and **OUTHHW4** respectively. The individual outcome is a two-digit code, whereas the household outcome is a three-digit code.

WAVE FIVE

Individual and household level interview outcomes are given in **W5INDOUT** and **W5HHOUT** respectively. The individual outcome is a two-digit code, whereas the household outcome is a three-digit code.

6.5 Survey Module Identifiers

Some modules in the interview were answered by all interviewed individuals whereas others were not. Each of the modules that could be answered by one individual on behalf of others has a variable that identifies the person who answered the module:

- Household Demographics = **HHRESP**
- Income and Assets = **IAPID**
- Housing = **HOPID**

Where modules were answered by one person on behalf of others, information was copied directly to the other member(s) of the household or financial unit.

6.6 Sample Type

FINSTAT indicates the status of the individual. This means whether they were interviewed as a core sample member (C1CM), younger partner (C1YP) or new partner (C1NP1). Note that the data set contains other status variables, for example the status at sampling (**Elsa** and **Partner**), but **FINSTAT** is the final version that should be used for analysis.

6.7 Age

DIAGR is the age given in the individual interview. Again, all respondents aged 90 or over have been given a value of 99 for this variable for confidentiality reasons. **DIAGR=99**.

INDOBYR is derived from the date of birth variables from the household grid (**DHDOB**, not archived) and individual interview (**DIDBN**, not archived). If the variable from the individual session was answered (**DIDBNY**) then this value was used, otherwise the value from the household grid (**DHDOB**) was used in the derivation. **INDOBYR** provides the year of date of birth only; the day and month of birth have been removed from the dataset to retain confidentiality.

INDAGER is computed from date of birth (**INDOB**, not archived) and date of interview (**INTDAT**, not archived). All respondents over age 90 have been classified as 99 years old for confidentiality reasons.

For age analysis, **INDOBYR** and **INDAGER** are the suggested variables to use.

6.8 Gender

DISEX was the sex given in the respondent's individual questionnaire session. **DHSEX** was given in the household demographics module. **INDSEX** is derived from these two variables. If the variable from the individual session was answered (**DISEX**) then this value was used, otherwise the value from the household grid (**DHSEX**) was used in the derivation.

6.9 Financial Variables

ELSA contains very detailed information on all aspects of finances. Summary variables have been derived by IFS, for all waves and are available from the Data Archive. Pension and mortgage data is also available as a separate level file.

6.10 Dropped Variables

All variables in the questionnaire documentation with an @ symbol next to their name have been deleted from the archived dataset.

The following types of variables have been deleted in order to reduce the potential to identify individuals and for other reasons (specified below):

1. Those containing text
2. Those which contained a personal identifier (e.g. name/address)
3. Those considered to be disclosive, such as:
 - Detailed ethnicity
 - Specific country of birth
 - Full interview date
 - Full date of birth
 - Council tax payments (Different councils charge different amounts and therefore the amount may reveal the area the respondent lives in)
 - Water and sewerage charges (These vary in different areas and therefore the amount may reveal the area the respondent lives in)
4. Timing variables, which give the time at specific points in the interview (used for administration purposes)
5. 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.

There are no geographical variables in the archived dataset. Geographical variables can be requested via the NatCen Data Release Panel. Please contact the ELSA Data Manager for more details (see the Contact Details section of this User Guide).

6.11 Missing Values

For most questions there are the following missing values:

- 1 Not applicable
- 8 Don't know
- 9 Refusal

For some questions, a response of 'don't know' or 'refusal' was not permitted. This is indicated in the questionnaire.

For various reasons, some respondents did not complete the entire interview. Where possible, for these 'partial' interviews, the questions that were not asked were coded as '-1' (not applicable).

As mentioned earlier, some individual values were re-coded as a result of issues with feedforward data. We have re-coded all possible erroneous values to the following missing value:

- 2 CAPI/Interviewer error

Where cases have this value, it does not necessarily mean that they would have had a 'positive' value for that question – equally they may not have been eligible for a question and so would have had a value of -1.

The only module where the missing values differ is the Self Completion module (SC). Here, the missing values are:

- 1 Not applicable
- 2 Schedule not applicable
- 9 Not answered

6.12 Coding and Editing

Additional coding and editing tasks were performed after the interviews were conducted. The Codebook and Edit Instructions document relevant to each wave (deposited with this User Guide) provides details of the tasks that were conducted.

Coding

The coding of responses was mostly dealt with by the CAPI questionnaire through the use of fully closed questions. There were, however, a number of questions where a code frame was not used in order to capture all responses to a particular question verbatim. The responses to these 'open' questions were coded into separate variables after the interview was conducted. The coded responses to open questions are in the core dataset, but as mentioned above the original text responses have been removed to reduce the potential to identify individuals.

Other questions in the CAPI questionnaire had code frames that included an 'other' option (e.g. **DIKLIV**). In these 'other-specify' questions, interviewers could use this option if the respondent's answer did not fit any of the codes or if they were not confident of coding into the prescribed codes. In these cases, the interviewer recorded the full 'other' answer at a follow up question (e.g. 'other' responses relating to **DIKLIV** were recorded in **DIKLO** – not archived).

If these other-specify questions were 'single coded', i.e. when only one option could be chosen (e.g. **DIKLIV**) then the text answers were coded and incorporated into the original coded variable (i.e. **DIKLIV**).

For all previously multi-coded variables (and from wave three dichotomous variables) that were coded, there are two sets of variables. The first are the original variables that contain the answers recorded by the interviewer (e.g. **SPTRBUNA** to **SPTRB95** - named **SPTRAB1** to **SPTRAB7** in previous waves). The second set of variables contain the original coding plus the codes assigned to "other answers" (e.g. **SPTRMUNA** to **SPTRM86** - named **SPTRM01** to **SPTRM07** in previous waves). Note that the letter 'm' added to these variable names after the 'stem' (**SPTR**) means that they contain 'merged' original and coded answers. This naming convention was followed consistently so the final merged variables can be identified by name. The merged variables should be used instead of the original variables.

The original variables provide multi-coded answers in the order in which they were mentioned by the respondent. Please note that in some cases the order of the respondent's answers in the merged variables is not the same as in the original variables.

EDITING

As with the coding, most of the editing was carried out by the interviewers in the field. However, there were some additional checks that related to inconsistencies in the data that were carried out after the interview.

7 Appendices

7.1 Notes about particular variables in the Health module

Variable name (in questionnaire)	Note
All HEOPT, HEDIAA and HEDIAB questions	<p>The idea of the diagnosed disease questions is to track diagnoses across time, but not to attempt to derive an exact history. The first time the respondent is interviewed for ELSA they are asked if they have ever been diagnosed with any of the diseases or conditions specified. For several of these conditions they are asked their age of diagnosis. In subsequent interviews they are asked to confirm that they had been diagnosed and, for most conditions, to say whether or not it is continuing. They are then asked if any additional conditions have been diagnosed since the previous interview and, if so (for several conditions), when.</p> <p>First-time ELSA interviewees are not asked whether they still have the condition but in some circumstances (e.g. heart attack, stroke, and psychiatric conditions) there are subsidiary questions that ask if they have experienced the disease/condition during the previous two years.</p>
	<p>A report of a diagnosis at these questions does not necessarily signify incidence since last interview for two reasons:</p> <ul style="list-style-type: none"> • Respondents interviewed by proxy do not have information fed forward so may report an old diagnosis here; • Some respondents report diseases for the first time although they were diagnosed a long time ago – analysts are advised to look at dates of diagnosis where available.
	<p>Users wishing to ascertain incidence or prevalence are advised to link data together across the waves and to note whether the interview was by proxy or not. Apart from the note above regarding incidence please also note;</p> <ul style="list-style-type: none"> • Proxies are asked if the respondent has ever had any of the listed conditions, even if they have previously reported diseases. They are not asked the questions about confirming a previous report - this is to preserve confidentiality. It has been found that proxies do not report all conditions that a respondent has previously reported in person; • Occasionally a respondent re-reports an old disease or condition that had been ‘dropped’ from the feed forward, because the respondent had told the interviewer at an earlier wave that they no longer had the condition.

	For first time ELSA interviewees and for proxies, Code 96 indicates that the respondent (or proxy) believes they have never been diagnosed with any of the conditions in the set. For subsequent interviews, Code 96 indicates that no additional diagnoses have been made since the last interview.
All HEOPF, HEDIAW and HEDIBW questions	These are the variables that tell the interviewer if a respondent has previously reported a condition. They are not used if the interview is by proxy.
HEDIAW	The feedback takes the answers as originally given to the interviewer, so ignores any recoding of “other” CVD conditions, e.g. if a respondent had said that they had a leaky heart valve at the previous interview, this would be recoded for the archived data set for that wave to Code 5 (variable HEDIMHM). However, for the current interview it would be fed forward as an “other” condition and the wording used by the respondent reported back to them for confirmation.
HEDIAS	Respondents who have previously reported the following conditions, and have not since denied every having them, are not asked if they still have the condition: heart attack, stroke, diabetes or high blood sugar. In Wave 2 this question was asked of those with diabetes or high blood sugar.
HEDIDS	Those who have previously reported the following conditions, and not since denied every having them, are not asked if they still have the condition: psychiatric condition, Alzheimer’s disease.
HEACD	Users will need to use the feed forward variable DHEACD to identify all those who have ever reported diabetes. In Wave 4 the variable applies to those who newly reported a diagnosis of diabetes or high blood sugar and to those who had reported diabetes or high blood sugar previously but had Code 2 at HEACD for their previous interview or who had not answered this question. In Wave 2 this question simply applied to those who had reported diagnosis of diabetes or high blood sugar either at Wave 1 or Wave 2.
HEHIBPB1, HEMDA1	These questions only apply to respondents who had had a stroke but had not already been asked about high blood pressure in the set of questions that specifically follow a diagnosis of high blood pressure (variables HEHIBPB and HEMDA).
HEHRTMD	Eligibility for this question has changed – previously it was those respondents with angina or myocardial infarction, but at Wave 4 it was extended to those with diabetes or stroke.
HEDIDS	The archived name of this variable has changed from HEDBLU at Wave 3 to HEDBSLU at Wave 4.
HEIQA – HEIQQ	Eligibility for this question has changed as a result of a change to the rules for proxy interviewing. The questions are intended to refer to those who are not capable of responding for themselves. Previously, proxy interviewers were only

	done in these circumstances but in Wave 4 the circumstances under which proxy interviews were allowed were extended. So, whereas previously this question applied to all proxies, in Wave 4 a condition was added to limit the questions to cases where the ELSA sample member was ill or away.
HEPAIN – HEDIZ	The pain questions were revamped for Wave 4 to enable a better fit to the ACOVE indicators (please refer to the “Wave 2 Quality of Care Indicators” document, available from the Data Archive, for more information). Although some questions have the same wording and eligibility analysts are reminded that the question order may have changed.
HEPD1, HEPD_9, HEPD2, HEPD_10, HEPD3, HEPD_11, HEPD4, HEPD_13, HEPD5, HEPD_16, HEPD6, HEPD_19, HEPD7, HEPD_22, HEPD8 and HEPD_25	These variables have been deposited although they appear to have been set erroneously to -1. Users are advised to refer to the relevant merged variables instead – we will deposit amended versions of these variables with the next version of the Phase 2 dataset.
HEAND1	This question is not part of the standard Rose Angina question sequence, but was included because of problems with pinpointing whether those who report that they never/cannot walk do in fact experience chest pain that might indicate angina. The reference is to walking at an ordinary pace on level ground if the answer to HEANC is Code 4, otherwise to walking uphill.
HERPE1	This question is not part of the standard MRC respiratory question sequence, but was included because of problems with pinpointing whether those who report that they never/cannot walk do in fact experience being short of breath. The reference is to walking with people of own age on level ground if HERPE is Code 3, otherwise to "uphill or hurry".
HEKNEA	At Wave 3, this question referred to knee pain among those reporting osteoarthritis and knee pain in range 1-5 (HEKNE) and not hip pain in range 6-10 (HEHIP) whereas in Wave 4 it refers more simply and extensively to those who had knee pain whether severity in any of range 1-10 and reported osteoarthritis at question HEART (variable HEARTOA on dataset).
HEHIPA	This question applies to respondents who rated their hip pain between 6 and 10 (with 10 being severe/excruciating) at HEHIP and also reported osteoarthritis at HEART.
HEPMED	This question applies if the respondent reports osteoarthritis and either had knee pain for more than six months with a severity between 6 and 10 (same scale as above), or the equivalent for hip pain.
HEROSMD	In Wave 3 this question applied to those reporting chest pain at HEANK or chest pain at sites 1-4 at HEANI. At Wave 4, those applicable were only those reporting chest pains at HEANK.

HEADLB	Two new activities were added in Wave 4: appreciating risk of danger and communication. This means that the order of the difficulties on the questionnaire has changed but variable names on the dataset have been assigned so that the same name in successive waves refers to the same difficulty, e.g. HEADLMO refers to difficulty with managing money at Waves 4 and 5 even though the position in the list of difficulties has changed.
HEHPB	HEHPB applies to those who reported help at HEHPA, had difficulty with an activity not covered by the previous questions and also had not yet reported any specific source of help. These questions are to be used in conjunction with all the other help questions to find whether a particular source of help was used at all by a respondent. They do not necessarily cover all the sources of help for these residual activities.
HEHPB1 – HEHPM11, HEHPSM5 – HEHPSM9	For these questions, the categories of who has given the help have changed since Wave 3. So, for example, at Wave 3 code 3 on the interview list represented “Whether receives help moving around house from son ”, whereas at Wave 4 it represented “Whether receives help moving around house from daughter ”. However, the variable names on the dataset have been designed such that the same name should apply to the same grouping at each wave, e.g. HEHPBSO refers to help from son at both waves regardless of the code on the interview list.
HEHPSM	The coding frame for these questions was revised for Wave 4. Analysts are advised to check variable labels carefully if making any comparisons to Wave 3. Also note that Code 3 (variable HEHSMEX) may not include all responses of “swimming” or “dancing” mentioned in HEHPSO. The former has occasionally been coded to hydrotherapy (variable HEHSMHY) and both may also be included in ‘other’ (variable HEHSM85 on the dataset)
HEHPLA, HEHPPD	These are new questions which apply to those respondents who reported a particular source of help for at least one activity with which they have difficulty.
HEMEB, HEHYY, HEOVY, HEPEE, HEHRT	The questions on menstruation are intended to update previous information collected during the life history interview.

7.2 Notes about particular variables in the Risk module (wave 5)

ritpcons	This is a categorical variable that indicates whether the time preference choices are consistent on tasks 1-6, on task 7-12, both, neither or both + consistent across the two sets
ridis1c	This is a categorical variable that assigns respondents to a patience category on the basis of tasks 1 to 6. the categories have upper and lower bounds for the weekly discount rate. The top and bottom categories are unbounded ("less than"... "more than") and there is a final category for those whose choices are inconsistent (and hence cant be classified).
ridis2c	As above but for tasks 7 to 12.