National Child Development Study

2008-2009 Follow-Up (INTERIM DEPOSIT)

Deposited with the UK Data Archive

A Guide to the Dataset

February 2009

Compiled by

Matt Brown, Jane Elliott, Margaret Hancock,

Peter Shepherd and Brian Dodgeon

Centre for Longitudinal Studies

Institute of Education, University of London

NCDS 2008-2009 Follow-Up Team

The NCDS 2008-9 follow-up was designed and conducted by a joint team drawn from CLS and the National Centre for Social Research (NatCen). The key members of the team are listed below:

CLS Team

Research Design and Implementation	Jane Elliott, Peter Shepherd, Matthew Brown, Lisa Calderwood, Maggie Hancock, Brian Dodgeon, Samantha Parsons
Cohort Studies Administration	Denise Brown, Kevin Dodwell
Tracing Unit	Kevin Dodwell, Mary Ukah, George Andrew, Peter Deane
Database Team	Robert Browne, Denise Brown, Jon Johnson, Maggie Hancock

NatCen Team

Research Design and Implementation	Marie Thornby, Elizabeth Hacker, Caroline Killpack, Carli Lessof, Reg Gatenby, Suneeta Bhamra, Richard Boreham
Field administration	Coral Lawson, Helen Selwood, Janice Morris, Mercedes Damm
CAPI/CASI Programming	Nafiis Boodhumeah, Peyman Damestani, Richard Akers, Sue Corbett

Acknowledgements

PREFACE

This document has been prepared to accompany the deposit, with the UK Data Archive at the University of Essex, of an interim dataset from the 2008-2009 follow-up of the National Child Development Study(NCDS), a continuing, multidisciplinary, national, longitudinal study.

The elements of the deposit, to which reference will be made throughout this document, are identified below. Users are advised that they will need to consult all elements of the documentation to gain a full understanding of the data.

NCDS 2008-2009 Interim Deposit:

Title	Format
National Child Development Study - Data collected in 2008/9 (at age 50) - INTERIM DEPOSIT	SPSS
National Child Development Study 2008-2009 Follow-Up (INTERIM DEPOSIT) – A Guide to the Dataset	PDF
NCDS 2008-2009 Follow-up: Draft Questionnaire Documentation	PDF
NCDS 2008-2009 Follow-up: Self-completion Questionnaire	PDF

Contents

1. **Introduction**

This document accompanies the National Child Development Study (NCDS) 2008/9 follow-up interim dataset which has been deposited with the UK Data Archive at the University of Essex.

The interim dataset is based on 2,997 interviews which were completed between August and December 2008 and is comprised of a subset of the full list of variables that will be included in the final version of the file which will be deposited with the Data Archive later this year.

The 2008/9 follow-up began in August 2008 and is scheduled to be complete by May 2009. The follow-up partially overlaps with a telephone follow-up of the 1970 British Cohort Study (1970 cohort). The NCDS survey was designed by the Centre for Longitudinal Studies of the Institute of Education, University of London (CLS), and the fieldwork is being carried out by the National Centre for Social Research (NatCen). The work is funded by the Economic and Social Research Council.

The interim dataset is intended to facilitate preliminary analyses of the data from the 2008/9 follow-up. The preliminary nature of the dataset is such that it should not be used as the basis for published work. Users of the data are requested to base any publications using data from the 2008/9 follow-up on the final dataset to be released later this year.

Users of this preliminary dataset must also be aware that the interim data have yet to be fully cleaned or edited and as such inconsistencies or apparent errors may be discovered. Users of the dataset are asked to inform CLS about any such inconsistencies using the feedback page on the CLS website: www.cls.ioe.ac.uk/feedback

1.1 NCDS Background

NCDS started life as the Perinatal Mortality Survey, which was designed to examine the social and obstetric factors associated with stillbirth and infant mortality. In the first survey, data were collected about the births and families of 17,638 babies born in Great Britain during one week in March 1958. There have subsequently been a further eight surveys which have sought to gather information from respondents living in England, Scotland and Wales3, in order to monitor their health, education, social and economic circumstances.

These surveys were carried out in 1965 (age seven), 1969 (age eleven), 1974 (age sixteen), 1981 (age 23), 1991 (age 33), 1999/2000 (age 42), 2004/2005 (age 46) and 2008/2009 (age 50). As part of the 1991 survey, information was additionally collected from co-resident partners and for a third of the sample data was also collected from any co-resident natural or adopted children of the cohort member.

Surveys of sub-samples of the cohort took place in 1976 (age 18), 1978 (age 20) and 1995 (age 37). The most recent sub-study, in 1995, involved conducting basic skills assessments with 10% of the cohort.

In addition, a 'Biomedical Survey' was conducted in 2002/3 which sought to obtain objective measures of ill-health and biomedical risk factors.

The 2004/2005 survey took the form of a 30 minute telephone interview which focused on updating the key changes in circumstances which had been experienced by cohort members in the time that had elapsed since their last interview.

The 2008/2009 survey is comprised of two parts: a 'core' face-to-face interview and a paper self-completion questionnaire. As in all recent follow-ups the main aim is to update information gathered in previous surveys in order to explore the factors central to the formation and maintenance of adult identity in each of the following domains:

• Lifelong learning
• Relationships, parenting and housing
• Employment and income
• Health and health behaviour
• Citizenship and values

1.2 Survey Design

A number of organisations were involved in the development and delivery of the 2008/2009 follow-up.

Centre for Longitudinal Studies (CLS) – CLS are an ESRC resource centre based at the Institute of Education, University of London and have been responsible for the study since 1991. CLS were responsible for the development of the 2008/2009 survey and commissioned the fieldwork.

Funders – The 2008/2009 survey was funded by the Economic and Social Research Council (ESRC).

Advisors – The content and design of the 2008/2009 survey was developed in collaboration with a number of advisory panels comprised of researchers, policy-makers and funders.

Fieldwork subcontractors – Following competitive tendering the National Centre for Social Research (NatCen) were selected to conduct the 2008/2009 survey on behalf of CLS. NatCen assisted CLS with the development of instrumentation, conducted the fieldwork and carried out initial data preparation (including coding and post field editing) and documentation.   NatCen also conducted the follow-ups in 2000 and 2004/2005.

2. Survey instruments

The 2008/2009 survey is comprised of the following elements:

1. A 55 minute 'core' interview which includes:

   a. Computer Assisted Personal Interview (CAPI) – 40 minutes
   b. Computer Assisted Self Interview (CASI) – 10 minutes
   c. A series of cognitive assessments – 5 minutes

2. A paper self-completion questionnaire

The paper self-completion questionnaires were (in most cases) posted to cohort members once the interviewer had arranged an appointment to conduct the core interview. In the majority of cases cohort members had then completed these questionnaires in advance of their core interview meaning that they could be collected by the interviewer when they visited. **The 2,997 cases included in this interim deposit have all completed the paper self-completion questionnaire.**

In addition, cohort members completing the survey were also asked for their consent to link their responses with administrative data held by DWP and NHS and if applicable to provide their consent for CLS to approach their parents about a possible study concerned with examining inter-generational transfers and relationships between cohort members and their parents.

More detail about each of these elements is provided below.

3. Survey content and key variables

Many of the questions in the 2008/9 follow-up interview had been asked in earlier waves of the NCDS and the BCS, which will allow for the making of comparisons both across the sweeps of NCDS and with the BCS cohort. There are, in addition, a number of new areas of data collection which have been introduced this time around including a new module covering symptoms of the menopause and a module featuring a series of cognitive ability tests.

Table 1 below provides details of some of the key variables which may be of interest.

Table 1 - Some key variables

Information	Variables	Variable label
Identifier	ncdsid	ncdsid serial number
Sex	n8CMsex	CM's sex
Legal marital status	n8MS	CM's legal marital status
Cohabitation status	n8Cohab	Whether CM cohabiting as a couple
Accomodation	N8accom	Type of accommodation
Number of rooms	n8numrms	Number of rooms in the house
Tenure	n8ten	Home ownership / tenure status
Mother alive	nd8maliv	(Derived) Whether CM's mother alive (incl prev swp data)
Father alive	nd8paliv	(Derived) Whether CM's father alive (incl prev swp data)
Non-residential relationships	n8othrea	Whether CM is currently in a non-residential relationship
Receipt of state benefits	N8stbe	Whether receiving any of state benefits specified (on card)
Economic activity	n8Econ02	(Derived) CM's current economic activity
Net pay	n8CNetWk	Computed : weekly amount of take-home pay
Partner's economic activity	n8pecac2	(Derived) CM partner's current economic activity
Partner's net pay	n8PNetWk	Computed : Partners net pay - weekly amount
Highest academic qualification	nd8achq	Highest Academic Qualification (1991,2000,2004,2008 data)
Health	n8HlthGn	CM self-assessment of health
Registered disabled	n8khldsl	Whether CM registered disabled
Health limits activities	n8khllt	Whether health limits everyday activities
Smoking	n8smokig	Smoking frequency
Alcohol consumption	n8drinks	Frequency of drinking alcohol

A broad summary of the content of the 2008/9 follow-up interview and self-completion questionnaire is provided in sections 3.1 to 3.4.

3.1 Core CAPI interview:

The CAPI interview collects updated information about household composition, housing, relationships, births and other pregnancies, periods of lone parenthood, adopted children, absent and older children, parents, family income, economic activity, education and qualifications, work-related training, use of computers, health, smoking, drinking, exercise, height, weight, social participation and social support.

3.2 CASI interview:

The CASI interview collects information on voting behavior and party support, experience of symptoms of the menopause, problematic drinking behavior, well-being, relationship satisfaction, domestic division of labour, job commitment, attitudes towards pensions and retirement, childhood, efficacy and life satisfaction.

The CASI interview included several established scales which are listed below along with the variables of which they are comprised. Overall scores for each scale have also been derived and included within the interim deposit. Additional information about the derivation of these overall scores and copies of the syntax used can be found in Appendix 2.

3.2.1 Kanungo's Job Involvement Scale

Four items from a ten-item scale (Kanungo, 1982) as used by Frone and Rice (1987).

Variable name	Variable label
n8comjb1	CASI: Whethr most imp things in CM's life relate to presjob
n8comjb2	CASI: Whether job is only small part of who CM is
n8comjb3	CASI: Whether CM personally involved in his/her job
n8comjb5	CASI: Whether most CM's interests centre around their job
nd8jobinv	(Derived) Overall job involvement score (1-6)

3.2.2 AUDIT (Alcohol Use Disorders Identification Test)

The AUDIT test replaces the CAGE scale as a measure of problematic alcohol consumption. The AUDIT consists of 10 questions covering alcohol consumption, problems and dependency. Responses to each question are scored from 0 to 4 giving a maximum score of 40. Scores of 8 or more are associated with harmful or hazardous levels of drinking; scores of 13 or more for women and 15 or more for men are likely to indicate alcohol dependence.

Variable name	Variable label
n8drinks	Frequency of drinking alcohol
n8Audi02	CASI: Num drinks containing alcohol CM has typ drinking day
n8Audi03	CASI: Frequency of six or more drinks on one occasion
n8Audi04	CASI: Frequency CM unable to stop drinking in last yr
n8Audi05	CASI: Frq drink caused CM failed to do as expctd in last yr
n8Audi06	CASI: Freq CM need drink first thing in morning in last yr
n8Audi07	CASI: Frequency CM felt guilt after drinking in last yr
n8Audi08	CASI: Frq drink made CM unable to rem prev night in last yr
n8Audi09	CASI: Whether CM/other injured due to CM drinking in last yr
n8Audi10	CASI: Wh rels/friends concern about CM drinking in last yr

3.2.3 Malaise inventory

A set of 24 self-completion questions which combine to measure levels of psychological distress, or depression (Rutter et al, 1970). The 2008-9 follow up uses 9 of the original 24 items.

Variable name	Variable label
n8mal02	CASI: Whether CM feels tired most of the time
n8mal03	CASI: Whether CM often feels miserable and depressed
n8mal05	CASI: Whether CM often gets worried about things
n8mal09	CASI: Whether CM often gets into a violent rage
n8mal12	CASI: Whether CM often suddenly scared for no good reason
n8mal14	CASI: Whether CM is easily upset or irritated
n8mal16	CASI: Whether CM is constantly keyed up and jittery
n8mal20	CASI: Whether every little thing gets on CM's nerves
n8mal21	CASI: Whether CM's heart often races like mad
nd8mal	(Derived) Total Malaise score (9 questions)
nd8malg	(Derived) Total Malaise score - grouped

3.3 Cognitive assessments:

The cognitive assessment module is comprised of four tests as described below. Variable names and variable labels are provided for reference.

3.3.1 Word list recall

A test of verbal learning and recall where participants are required to learn a list of 10 common words. The CAPI program randomly selects one of four lists of words which are presented to the respondent by the computer using a recorded voice. In cases where the computer voice is not audible the list is read aloud by the interviewer, who is asked to imitate the pace and clarity of the recorded voice, reading the words at approximately 2 second intervals.

Once the list had been read out, cohort members have up to two minutes to recall as many as they can. Interviewers make a note of each word correctly recalled and enter the total into the CAPI program.

Variable name	Variable label
n8CfLisn	Number of words correctly recalled

3.3.2 Animal naming

A test of verbal fluency which measures how quickly participants can think of words from a particular category, in this case naming as many different animals as possible within one minute.   Interviewers make a note of each named animal and enter the total into the CAPI program. Repetitions, named animals (e.g. Bambi) and redundancies (e.g. white cow, brown cow) are excluded from the total score.

Variable name	Variable label
n8CfAni	Number of animals mentioned

3.3.3 Letter cancellation

A test of attention, mental speed and visual scanning. The participant is given a page of random letters of the alphabet and asked to cross out as many "Ps" and "Ws" as possible within one minute.

(This exercise is scored post-interview. Scores are not included in the interim deposit).

3.3.4 Delayed word list recall

A test of delayed memory which asks the participant to recall as many words as they can from the original list presented to them during the first word-recall task. The word lists are not repeated and participants have again two minutes to recall as many as they can. Interviewers make a note of each word correctly recalled and enter the total into the CAPI program.

Variable name	Variable label
n8CfLisD	Number of words recalled after delay

3.4 Comparability with other studies

When conducting the cognitive assessments interviewers working on the 2008-9 follow-up followed exactly the same procedures as were employed by interviewers working on the English Longitudinal Study of Aging (ELSA) (Insert Ref to ELSA) and as such the data collected by the two studies will be directly comparable.

Word-list recall exercises and the letter cancellation task have also been included in the 1946 cohort study (the National Survey of Health and Development). There are however a number of small differences in protocol which will have an impact on the comparability of results:

- The word list recall exercise asks 1946 cohort members to recall 15 words whereas ELSA and NCDS ask respondents to recall 10 words.

- In the 1946 study word list recall exercise the words are shown to the respondent in a flip book (at intervals of two seconds) whereas in the ELSA/NCDS exercise the words are read to the respondent by the computer (unless respondent unable to hear well in which case the words are read by interviewer). A person's ability to recall words which they have read may differ from their ability to recall words which they have heard spoken.

- In the 1946 study, respondents are asked to recall the words on 3 occasions whereas ELSA/NCDS respondents are only asked to recall the words twice. On the 1946 study, once the word-list recall task has been completed for the first time it is immediately repeated a second time whereas on ELSA the task is only completed once at first. Each of the studies then include a delayed word-list recall exercise but the 1946 respondents will be at an advantage as they will have had an extra opportunity to commit the words to memory.

- 1946 cohort members are given one task between the original word list recall exercise and the delayed word-list recall exercise (the letter-cancellation task) whereas ELSA/NCDS respondents are given two tasks (the letter-cancellation exercise and the animal naming exercise).

The word-list recall exercises (immediate and delayed) were also included in the 1993, 1995, 1996, 1998, 2000, 2002 and 2004 sweeps of the Health and Retirement Study. The protocols followed by interviewers working on the HRS were exactly the same as those working on NCDS and ELSA meaning the data collected by each of the studies will be comparable. The only difference is that between the immediate word-list recall exercise and the delayed word-list recall exercise the HRS has included 5 minutes of questioning rather than additional cognitive assessments as included in ELSA/NCDS.

3.5 Paper self-completion:

The paper self-completion questionnaire collects information on leisure activities, personality type, health, sleep, values and attitudes, sense of community and quality of life. In addition, cohort members were asked an open question where they were asked to write about the life that they imagined they would be living at the age of 60.

The paper self-completion included several established scales. As mentioned above, overall scores for each scale have also been derived and included within the interim deposit. Additional information about the derivation of these overall scores and copies of the syntax used can be found in Appendix 2.

3.5.1 Personality Inventory

50 questions from the International Personality Item Pool (IPIP) (Goldberg, 1999).

Variable name	Variable label
n8scq2a	SC:I am the life of the party
n8scq2b	SC:I feel little concern for others

n8scq2c	SC:I am always prepared
n8scq2d	SC:I get stressed out easily
n8scq2e	SC:I have a rich vocabulary
n8scq2f	SC:I don't talk a lot
n8scq2g	SC:I am interested in people
n8scq2h	SC:I leave my belongings around
n8scq2i	SC:I am relaxed most of the time
n8scq2j	SC:I have difficulty understanding abstract ideas
n8scq2k	SC:I feel comfortable around people
n8scq2l	SC:I insult people
n8scq2m	SC:I pay attention to details
n8scq2n	SC:I worry about things
n8scq2o	SC:I have a vivid imagination
n8scq2p	SC:I keep in the background
n8scq2q	SC:I sympathise with others' feelings
n8scq2r	SC:I make a mess of things
n8scq2s	SC:I seldom feel blue
n8scq2t	SC:I am not interested in abstract ideas
n8scq2u	SC:I start conversations
n8scq2v	SC:I am not interested in other people's problems
n8scq2w	SC:I get chores done right away
n8scq2x	SC:I am easily disturbed
n8scq2y	SC:I have excellent ideas
n8scq2z	SC:I have little to say
n8scq2aa	SC:I have a soft heart
n8scq2bb	SC:I often forget to put things back in their proper place
n8scq2cc	SC:I get upset easily
n8scq2dd	SC:I do not have a good imagination
n8scq2ee	SC:I talk to a lot of different people at parties
n8scq2ff	SC:I am not really interested in others
n8scq2gg	SC:I like order
n8scq2hh	SC:I change my mood a lot
n8scq2ii	SC:I am quick to understand things
n8scq2jj	SC:I don't like to draw attention to myself
n8scq2kk	SC:I take time out for others
n8scq2ll	SC:I shirk my duties
n8scq2mm	SC:I have frequent mood swings
n8scq2nn	SC:I use difficult words
n8scq2oo	SC:I don't mind being the centre of attention
n8scq2pp	SC:I feel others' emotions
n8scq2qq	SC:I follow a schedule
n8scq2rr	SC:I get irritated easily
n8scq2ss	SC:I spend time reflecting on things
n8scq2tt	SC:I am quiet around strangers
n8scq2uu	SC:I make people feel at ease
n8scq2vv	SC:I am exacting in my work
n8scq2ww	SC:I often feel blue
n8scq2xx	SC:I am full of ideas

3.5.2 Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) –

A 14 positively worded item scale with five response categories. It covers most aspects of positive mental health (positive thoughts and feelings) including both hedonic and eudaimonic perspectives (Tenner et al, 2007).

Variable name	Variable label
n8scq3a	SC:I've been feeling optimistic about the future
n8scq3b	SC:I've been feeling useful
n8scq3c	SC:I've been feeling relaxed
n8scq3d	SC:I've been feeling interested in other people
n8scq3e	SC:I've had energy to spare
n8scq3f	SC:I've been dealing with problems well
n8scq3g	SC:I've been thinking clearly
n8scq3h	SC:I've been feeling good about myself
n8scq3i	SC:I've been feeling close to other people
n8scq3j	SC:I've been feeling confident
n8scq3k	SC:I've been able to make up my own mind about things
n8scq3l	SC:I've been feeling loved
n8scq3m	SC:I've been interested in new things
n8scq3n	SC:I've been feeling cheerful
nd8wemwbs	(Derived) Warwick Edinburgh Mental Well-Being Scale

3.5.3 SF-36

A widely used multi-purpose health survey comprised of 36 questions. It yields an 8-scale profile of functional health and well-being scores as well as psychometrically-based physical and mental health summary measures and a preference-based health utility index (Ware, Snow, Kosinski, & Gandek, 1993).

Variable name	Variable label
n8scq4a	SC:Vigorous activities, such as running, lifting heavy objects. participating in strenuous sports
n8scq4b	SC:Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf
n8scq4c	SC:Lifting or carrying groceries
n8scq4d	SC:Climbing several flights of stairs
n8scq4e	SC:Climbing one flight of stairs
n8scq4f	SC:Bending, kneeling or stooping
n8scq4g	SC:Walking more than one mile
n8scq4h	SC:Walking half a mile
n8scq4i	SC:Walking 100 yards
n8scq4j	SC:Bathing or dressing yourself
n8scq5a	SC:Past 4 wks phys Health led to cut down amount of time spent

	on work/other activities
n8scq5b	SC:Past 4 wks phys Health led to accomplished less than would like
n8scq5c	SC:Past 4 wks phys Health led to limited in the kind of work or other activities able to do
n8scq5d	SC:Past 4 wks phys Health led to difficulty performing work/other activities
n8scq6a	SC:Past 4 wks emo probs led to cut down amount of time you spent on work or other activities
n8scq6b	SC:Past 4 wks emo probs led to accomplished less than would like
n8scq6c	SC:Past 4 wks emo probs led to not done your work/other activities as carefully as usual
n8scq7	SC:Past 4 wks, what extent has phys health/emo probs ifamily friend etc.
n8scq8	SC:How much bodily pain have you had during the past 4 weeks?
n8scq9	SC:During the past 4 weeks, how much did pain interfere with your normal work (including both work outs
n8scq10a	SC:Did you feel full of life?
n8scq10b	SC:Have you been a very nervous person?
n8scq10c	SC:Have you felt so down in the dumps nothing could cheer you
n8scq10d	up?
n8scq10e	SC:Have you felt calm and cheerful?
n8scq10f	SC:Did you have a lot of energy?
n8scq10g	SC:Have you felt downhearted and low?
n8scq10h	SC:Did you feel worn out?
n8scq10i	SC:Have you been a happy person?
n8scq10j	SC:Did you feel tired?
	SC:Has you health limited your social activities (like visiting friends, relatives, etc.)?
n8scq11a	SC:I seem to get ill a little easier than other people
n8scq11b	SC:I am as healthy as anybody I know
n8scq11c	SC:I expect my health to get worse
n8scq11d	SC:My health is excellent
nd8phhe	(Derived) SF-36 Physical functioning score
nd8rlmp	(Derived) SF-36 Role-limitations due to physical health
nd8rlme	(Derived) SF-36 Role-limitations due to emotional problems
nd8enfa	(Derived) SF-36 Energy/fatigue score
nd8emwb	(Derived) SF-36 Emotional Well-Being score
nd8socf	(Derived) SF-36 Social Functioning score
nd8pain	(Derived) SF-36 Pain score
nd8genh	(Derived) SF-36 General health score

3.5.4 Medical Outcomes Study Sleep Scale

4 items from a 12 item scale measuring quality of sleep (Hays & Stewart, 1992).

Variable name	Variable label
n8scq12	SC:During the last 4 weeks,what was usual time taken to fall asleep
n8scq13	SC:During the last 4 weeks, average number of hours sleep per night
n8scq14	SC:During the last 4 weeks, how often did you waken and have trouble falling back to sleep
n8scq15	SC:During the last 4 weeks, how often did you sleep enough to feel rested upon waking ?

3.5.5 CASP-14

A scale designed to measure quality of life in the 'third age' by using Likert-scaled questions which cover four theoretical domains: control, autonomy, self-realisation and pleasure. The 2008-9 follow up uses 14 items from the full 19 item scale (Wiggins et al, 2004).

Variable name	Variable label
n8scq19a	SC:My age prevents me from doing the things I would like to
n8scq19b	SC:I feel what happens to me is out of my control
n8scq19c	SC:I feel left out of things
n8scq19d	SC:I can do the things I want to do
n8scq19e	SC:Family responsibilities prevent me from doing what I want to do
n8scq19f	SC:I feel that I can please myself what I do
n8scq19g	SC:Shortage of money stops me from doing things I want to do
n8scq19h	SC:I look forward to each day
n8scq19i	SC:I feel that my life has meaning
n8scq19j	SC:I enjoy the things that I do
n8scq19k	SC:On balance I look back on my life with a sense of happiness
n8scq19l	SC:I feel full of energy these days
n8scq19m	SC:I feel that life is full of opportunities
n8scq19n	SC:I feel that the future looks good for me
nd8casp14	(Derived) Overall CASP-14 Quality of Life Score

4. NCDS Interim Dataset

The interim dataset for the NCDS 2008-9 follow-up (has been supplied to the UK Data Archive) in the form of an SPSS dataset as follows:

National Child Development Study - Data collected in 2008/9 (at age 50) - INTERIM DEPOSIT

The file contains 2,997 cases. All cases completed the CAPI interview, CASI interview and the paper self-completion questionnaire. The file contains no partial interviews and no proxy interviews.

The cognitive assessment module was completed by 2,985 cases.

The variables included in the interim dataset comprise of a subset of the full list of variables which are to be included in the final version of the file. **In total the file contains 932 variables**. Broadly speaking the variables included file are those which required little or no post-fieldwork cleaning and editing in order to make them interpretable.

Details of the CAPI/CASI program and copies of paper questionnaires may be found in the following, which also accompany the data deposit:

NCDS 2008-2009 Follow-up: Draft Questionnaire Documentation
NCDS 2008-2009 Follow-up: Self-completion Questionnaire

Excluded variables fall into 4 main types:

1) Variables relating to the four 'history modules' – housing history, relationship history, fertility history and economic activity history.
2) Variables relating to questions where routing was dependent on responses given in a previous sweep of the study.
3) Variables relating to 'unfolding bracket sequences' – variables which are triggered when a cohort member either reports that they do not know or they do not wish to answer a question with a numerical response (e.g. their net income from employment). The unfolding bracket sequences present the respondent with a sequence of questions which attempt to collect an approximation of the missing value by asking whether the response is more or less than x.
4) Open questions where responses were entered verbatim.

Appendix Table 1 provides a detailed summary of the core interview variables which have been included in, or excluded from the interim dataset. All variables from the paper self-completion questionnaire have been included in the data set, with the exception of answers to Q21, an open question asking cohort members to imagine that they are 60 years old and to write a few lines about the life they are leading.

Additional information about the variable names, labeling of variables and about CASI variables, cognitive assessment variables, paper self-completion variables, identifiers and derived variables is given below.

4.1 Variable names

As noted above, the core interview (including the cognitive assessment module) was conducted using CAPI and CASI and as such the variable names in the dataset are based on those automatically allocated by the CAPI program (Blaise).

Within the Blaise, each question has a variable name (rather than number), made up of a maximum of 8 characters, and this is used as the root of the variable name on the dataset. Where the question is repeated (eg: the same question is asked for each birth, relationship, job, qualification, etc reported), Blaise automatically allocates a number suffix (eg: name, name2, name3, name4).

Unfortunately, where the variable name in the Blaise program was originally more than 6 characters long, Blaise truncates the name to allow for the suffix. As a result, there is not always a simple match between the Blaise program documentation and the data.

As many of the questions asked in the core interview were identical to the questions asked in the 2004/5 sweep the variable names allocated by Blaise were also identical. In order to ensure that variable names in the 2008/9 data are different to those in the 2004/5 data, all variables in the interim 2008/9 data set have been given the prefix n8. In some cases the variable names have also been truncated in order to limit the name to 8 characters. Where a variable was included in the 2004/5 sweep the variable name has, as far as possible been kept identical for the 2008/9 sweep although prefixed with "n8" rather than "n7".

In addition it should be noted that a number of derived variables have been included in the deposited dataset. Names of these variables are given the prefix "nd8" and variable labels are given the prefix "(Derived)". A list of all derived variables in this deposit is provided in Appendix 2.

A number of variables are automatically derived within the CAPI program. These variables have the usual "n8" prefix but the variable labels are "prefixed" with "(Recoded)"

Variables from the paper self-completion questionnaire are derived from the question numbers as they appear of the printed questionnaire and prefixed with "n8sc". They have the following form: n8scq1a, n8scq1b, n8scq2a, n8scq2b etc.

4.2 Variable labels

The variable labels included on the datasets relating to the core interview are also initially derived from the CAPI/CASI program. In exporting the SPSS dataset from Blaise, labels based on the wording of questions were automatically allocated. The variable labels in this interim dataset have not yet been individually reviewed. This will occur in advance of the deposit of the final data set and, where necessary, labels will be modified in an effort to ensure that labels are comprehensible and accurate.

Variable labels for the data derived from the paper questionnaire have been derived from the question wording used on the printed questionnaire.

4.3 Value labels

The value labels are similarly derived from the Blaise program or printed questionnaire. Value labels will be individually reviewed and amended, where necessary, in advance of the deposit of the final data set.

4.4 Missing values

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

-9 = Refusal
-8 = Don't Know
-2 = Schedule not applicable
-1 = Item not applicable

For some variables, additional missing values may occur. They will be in the range -3 to -6 and have been labelled to indicate the reason the data is missing.

4.5 Variable order

The order in which variables appear in the dataset will broadly follow the order of sections, and of questions within sections of the survey instruments – CAPI/CASI or paper. However, for the core interview elements (CAPI,CASI, cognitive assessments) the order is determined by the structure of the Blaise program, which does not necessarily hold each question in the order in which they are put to the respondent. This change in order is typically, but not exclusively associated with question sequences which are repeated to produce grid-like data structures (eg: birth, relationship, job, qualification histories, etc).

4.6 Value labels

The use of CAPI/CASI has served to ensure that the values should be within the specified range for each variable in datasets relating to the core interview. In advance of depositing the final data set the range of values for all variables in all datasets will be checked and, where necessary, updated during the assessment of data quality to be undertaken by CLS (see below).

4.7 Consistency

Again, the use of CAPI/CASI should ensure that all filters have been correctly followed within the datasets relating to these elements of the core interview, and that the data is consistent throughout.

The paper self-completion questionnaire contained no routing. The consistency of all data will be checked during the assessment of data quality to be undertaken by CLS in advance of the deposit of the final dataset (see below).

4.8 CASI Self-completion

The CASI self-completion was administered toward the end of the core interview. The interviewer handed the laptop computer used for the interview to the cohort member and explained how they should complete the questionnaire. Where the cohort member was unable or reluctant to use the

laptop, the interviewer assisted, and if necessary administered the self-completion as an interview. The variables which hold the data for the CASI self-completion are identified below. Labels for these variables are given the prefix "CASI".

n8CarOwn	n8MenS09	n8MenS20	n8RelSat
n8vote01	n8SymP09	n8SymP20	n8wcooks
n8votewo	n8MenS10	n8MenS21	n8wshops
n8votenw	n8SymP10	n8SymP21	n8wclean
n8prtypp	n8MenS11	n8Audi02	n8wwash
n8polint	n8SymP11	n8Audi03	n8wdiy
n8MenS01	n8MenS12	n8Audi04	n8wcash
n8SymP01	n8SymP12	n8Audi05	n8wtends
n8MenS02	n8MenS13	n8Audi06	n8wteach
n8SymP02	n8SymP13	n8Audi07	n8wcares
n8MenS03	n8MenS14	n8Audi08	n8comint
n8SymP03	n8SymP14	n8Audi09	n8comjb1
n8MenS04	n8MenS15	n8Audi10	n8comjb2
n8SymP04	n8SymP15	n8mal02	n8comjb3
n8MenS05	n8MenS16	n8mal03	n8comjb5
n8SymP05	n8SymP16	n8mal05	n8penint
n8MenS06	n8MenS17	n8mal09	n8PenSec
n8SymP06	n8SymP17	n8mal12	n8penwrk
n8MenS07	n8MenS18	n8mal14	n8worret
n8SymP07	n8SymP18	n8mal16	n8affsvr
n8MenS08	n8MenS19	n8mal20	n8rtfaro
n8SymP08	n8SymP19	n8mal21	n8Enoulc

4.9 Identifiers

This identifier (ncdsid) replaces the old case identifier 'serial'. In August 2008, all historic NCDS datasets were re-deposited at the UK Data Archive with the old case identifier 'serial' replaced by a new identifier 'ncdsid'. This identifier is also to be found on this interim dataset and can be used to link the data longitudinally to earlier sweeps.

The reason for changing the case identifier was to improve the security of the data and increase safeguards on the confidentiality of cohort members.

All datasets carrying the old serial numbers should normally be deleted within three months of receipt of the revised data by the researcher.

For more details on this change, download the CLS Confidentiality and Data Security Review document available from the CLS website:

http://www.cls.ioe.ac.uk/studies.asp?section=00010002000300010005

4.10   Repairing relationship histories

Post-fieldwork editing and cleaning of the relationship history data collected in the previous sweep of the study (the 2004-5 follow-up) revealed that there were a number of problems with the routing of the questions in this module which lead to flaws in the information obtained about the period between the 1999/2000 follow-up and the 2004/5 follow-up. Details of the problems are provided below:

1. Cohort members who were cohabiting at the time of the 1999/2000 follow-up and also cohabiting at the time of the 2004/5 follow-up but with a different partner, they were not asked about the end of the relationship they were having with the person they were cohabiting with at the time of the 1999/2000 follow-up.

2. A number of cohort members who were cohabiting at the time of the 2004/5 follow-up, reported that their current partner was the same partner that they had at the time of the 1999/2000 follow-up, but at the time of the 2000 survey reported not having a partner.

3. When recalling details about previous partnerships, cohort members who were not married at the time they moved in with an ex-partner but who subsequently did get married, were not asked whether they got divorced or when.

4. In both the 1999/2000 follow-up and the 2004/5 follow-up there were a number of cohort members who reported being married or living with their current partner in the relationship history module but did not report a spouse or partner in the household grid.

5. There were also a number of cohort members with missing information about their marital and/or legal marital status at the time of the 1999/2000 follow-up or the 2004/5 follow-up.

Cases where any of the above applied were flagged as 'repair cases'. A revised introduction to the relationship history module was triggered for these cases where it was explained that when they were interviewed as part of the 2004/5 follow-up there was a problem with the data that was collected about their relationship history and that as a result it was necessary to recollect. These cohort members were then asked for a full relationship history from the date of their 1999/2000 follow-up interview through to the date of their 2008/2009 follow-up interview.

The majority of the relationship history variables have not been included in the interim data set.

4.11 Reference dates for retrospective data/histories

The 2008/9 follow-up gathered retrospective information on housing situation, relationships, pregnancies, economic activity and qualifications. The majority of cohort members had participated in one of the last two follow-ups which took place in 1999/2000 and 2004/5. For such cohort members the reference date for retrospective questions was the date of their last interview. For the small number of cohort members who had not participated in either of these follow-ups the reference date was generally the 1st January 2000; the exception to this was the pregnancy history module which for these cohort members used their 16th birthday as the reference date so that a full pregnancy history was collected.

The majority of the retrospective variables have not been included in the interim data set.

4.12 Derived variables

A number of derived variables (including the overall scores from the various scales employed in the questionnaire) have been included in the dataset. The variable names all have the prefix 'ND8' and the variable label are endorsed '(Derived)'. A full list is provided in Appendix 2, together with details of their derivation.

4.13 Further information

Queries about any aspect of the interim data should be sent to CLS using the feedback page on the CLS website: www.cls.ioe.ac.uk/feedback.

5. Survey development and piloting

Each element of the 2008/9 follow-up was rigorously tested prior to the commencement of main-stage fieldwork.

5.1 First pilot

A first pilot took place between the 28th November and the 24th December 2007 with specially recruited members of the public aged 46-55 in 12 areas. This first pilot comprised the first test of the questions and accompanying procedures for both the 'core' interview and the paper self-completion questionnaire. The pilot sought to measure the length of the interview and to identify any routing or filter errors within the program. It was also important to identify any problematic questions (in terms of comprehension or sensitivity) and any administrative or procedural difficulties.

65 individuals completed the 'core' interview (CAPI interview, CASI interview and cognitive assessments) and 42 individuals completed the paper self-completion questionnaire.

The pilot was generally very successful and both the core interview and the paper questionnaire were well received. The key issue was that the core interview was considerably longer than its target length, and as such substantial cuts were required in advance of the second pilot or 'dress rehearsal'.

5.2 Dress Rehearsal

The dress rehearsal took place in 8 areas between the 7th April and the 8th May 2008. It was intended as a test of all changes made as a result of the pilot stage and as a full procedural test of all survey instruments, documents and procedures prior to main stage fieldwork. The dress rehearsal took place with actual cohort members so that in addition to further testing of the elements above it was possible to test contacting and tracing procedures.

In total 77 cohort members completed a core interview, of whom 72 also completed the paper self-completion questionnaire.

As was found in the first pilot, each of the four elements (CAPI interview, CASI interview, cognitive assessments and paper self-completion questionnaire) were well received by the participating cohort members. The post-pilot cuts to the core interview were found to be sufficient to have adequately reduced the interview length so it was not necessary to make any further revisions in advance of the main stage.

6. Fieldwork

6.1 Fieldwork period

Fieldwork began in August 2008 and is scheduled to run until May 2009. The 2,997 cases included in this preliminary dataset were all interviewed between August 2008 and December 2008.

6.2 Issue of sample and fieldwork waves

The sample was issued to the fieldwork contractor in two batches.

The first (and largest batch) was issued in advance of the main stage of fieldwork and was comprised of 11,707 cohort members who had either:

1) Participated in NCDS6 (2000), NCDS Biomedical Survey (2002) or NCDS7 (2004) and had not subsequently died, emigrated or permanently withdrawn from the study (n=11,320).
2) Not participated in any of the above but had confirmed their address by responding to a birthday card mailing or in some other fashion since 2000 (n=387).

The second batch of sample was issued to the fieldwork contractor in December 2008 and was comprised of 609 previously untraced cases for whom new contact details had been obtained either from a tracing exercise conducted in collaboration with the Department of Work and Pensions (DWP) or a tracing exercise conducted via the National Health Service Central Register.

Sample was subsequently issued to interviewers in 8 overlapping waves. The first batch of sample was issued to the first six waves of fieldwork and the second batch of sample was issued to the seventh wave of fieldwork. The eighth wave of fieldwork will be used to 'mop-up' the cases which were difficult to contact in the earlier waves.

The cases included in this preliminary data file were all allocated to the first three waves of fieldwork.

Wave	Fieldwork start date	Number of cohort members
Wave 1	11th August	2561
Wave 2	15th September	2471
Wave 3	13th October	1850
Wave 4	10th November	1734
Wave 5	8th December	1778
Wave 6	12th January 2009	1313
Wave 7	18th February 2009	609
Wave 8 - Mop-Up Wave	30th March 2009	?
Total		12,316

6.3 Making contact

In advance of any contact from NatCen, cohort members were sent a pre-notification mailing which advised that the 2008-9 follow-up would soon be starting. The pre-notification mailing was accompanied by a summary report entitled 'Now we are Fifty' which provided cohort members with a summary of some of the key findings to have emerged from the study since it began in 1958. The pre-notification mailing was mailed to cohort members in 4 batches:

1) Cases allocated to Waves 1 and 2 were mailed 3-4 weeks in advance of the start of Wave 1 fieldwork.
2) Cases allocated to Waves 3 and 4 were mailed 3-4 weeks in advance of the start of Wave 3 fieldwork.
3) Cases allocated to Waves 5 and 6 were mailed 3-4 weeks in advance of the start of Wave 5 fieldwork.
4) Cases allocated to Wave 7 were mailed 3-4 weeks in advance of the start of Wave 7 fieldwork.

(There was no pre-notification mailing in advance of Wave 8 as cases allocated to this wave had previously been allocated to an earlier wave).

Where pre-notification letters were returned because the cohort member no longer lived at the mailed address these cases were prioritised for tracing by the CLS tracing team.
Each wave of fieldwork was preceded by the mailing of an advance letter which advised cohort members that an interviewer would be calling shortly. Once contact had been established interviewers attempted to arrange appointments to conduct the core interview (CAPI, CASI and cognitive assessments). Paper self-completion questionnaires were then posted to cohort members with a letter confirming the date and time of the appointment, so that in most cases the questionnaire could be collected by the interviewer when they visited the cohort member's home to conduct the interview.

In cases where interviewers established that a cohort member was no longer living at the issued address they were expected to make reasonable efforts to trace that individual. Interviewer tracing methods included contacting the new occupants of the issued address, contacting the 'stable contact' using contact details provided to them and talking to neighbours. Where the interviewer failed to find the cohort member the case was passed back to the CLS tracing team for further investigation.

In the small number of cases where the cohort member was unable to understand or respond to questions put by the interviewer, short proxy interviews were undertaken with a family member or carer. There is no proxy interview data in the preliminary data set.

6.4 Pre-fieldwork tracing

Between follow-ups efforts are made by CLS, through the mailing of an annual birthday card and other activities, to maintain contact with as many members as possible. Prior to fieldwork, the CLS Tracing Team attempted to obtain a current address for as many cohort members as possible.

The methods used by the CLS tracing team in advance of fieldwork include the use of:

- Contact details previously provided by the cohort member (for themselves and for relatives and friends)
- Other information contained in study records
- Telephone number databases
- Postcode databases
- Electoral register databases
- National Health Service Central Register records of NHS registration, emigrations and deaths

In addition to the efforts of the tracing team, a major tracing exercise of cohort members was undertaken in September 2007 in collaboration with the DWP. DWP attempted to match the details of cohort members who had not been contacted for some time against benefit records. Where details of cohort members were matched to DWP records, tracing letters were sent by DWP where an address different to that held by CLS was identified. Those receiving benefits were sent an opt-out letter and those not receiving benefits were sent an opt-out letter.

NCDS cohort members are also 'flagged' on the National Health Service Central Register (NHSCR). In advance of fieldwork, CLS also gained the agreement of ONS that tracing letters could be mailed to untraced cohort members via Health Authorities identified on the NHSCR.

Updated address information collected via these two methods was not obtained until after fieldwork had started; previously untraced cases for whom new addresses became available were therefore allocated to Wave 7. If new address information was obtained via DWP or NHSCR tracing for cohort members already issued then these new details were provided to NatCen so that interviewers could use them to help with their tracing efforts.

6.5 Briefing

All interviewers involved in the 2008-9 follow-up attended a one-day briefing in advance of commencing their assignments. The briefings covered the background to the NCDS, contact and tracing procedures, the CAPI and CASI interview, conducting the cognitive assessments, collecting contact information, collecting consents and the use of survey documents.

7. Data coding and editing

Interviewers transmit the data they have collected via modem and residual editing will be undertaken in advance of the deposit of the final data set.   The data provided in this preliminary dataset are provided in largely unedited form.

However, the use of CAPI/CASI provides for the incorporation of an extensive range of checks on validity, range and consistency and as such the amount of post-fieldwork editing that will be required should be fairly minimal.

The CAPI interview includes a number of questions where precodes are provided for answers but provision is also made to record additional information where the precode 'other' is used. These questions will be reviewed by the CLS team in order to determine the priorities for coding, and to identify the appropriate coding frames. Where possible, coding frames that had been employed for earlier NCDS surveys will be adopted although it may be necessary to add additional codes or in some cases to develop a coding frame from scratch. Coded responses will be included in the final data set but the data provided in this preliminary dataset has not been coded as yet.

8. References

Babor T.F., De la Fuente J.R., Saunders J., Grant, M (1989), AUDIT: the alcohol use disorders identification test—guidelines for use in primary health care. Geneva, World Health Organization.

Frone, M and Rice, W (1987), Work-family conflict: The effect of job and family involvement, Journal of Occupational Behaviour, Vol. 8, No. 1 (Jan., 1987), pp. 45-53.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), Personality Psychology in Europe, Vol. 7 (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Hays, R. D. and Stewart, A. L. (1992) 'Sleep measures'. In A. L. Stewart and J. E. Ware (eds) Measuring functioning and well-being: The Medical Outcomes Study Approach. Durham, NC: Duke University Press.

Kanungo, R (1982), Measurement of job and work involvement, Journal of Applied Psychology. Vol 67(3), pp. 341-349.

Saunders JB, Aasland OG, Babor TF, De la Fuente JR, Grant M (1993), Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption. Addiction. Vol. 88 pp 791–803.

Tennant, R. et al. (2007) The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): Development and UK validation, Health and Quality of Life Outcomes, 5(63).

Ware JE, Snow KK, Kosinski M, Gandek B (1993), SF-36 Health Survey Manual and Interpretation Guide. Boston, MA: New England Medical Center, The Health Institute.

Wiggins, R. et al (2004) 'Quality of life in the third age: key predictors of the CASP-19 measure', Ageing & Society, 24(5).

Appendices

Appendix 1 - Variables included and excluded from NCDS8 initial data deposit

Module	Variables	Whether included or excluded from preliminary data file.
Introduction	Confirmation of name, date of birth etc.	Excluded
	Confirmation of sex	Included
	Type of residence	Included
Household grid	Legal marital status of cohort member	Included
	Details collected for each member of household:	
	Sex	Excluded
	Age	Excluded
	Relationship to cohort member	Excluded
	Reason those who have left household since last interview moved out	Excluded
	Details of all children not living in household:	
	Sex	Excluded
	Age	Excluded
	Relationship to cohort member	Excluded
	Parenthood of children (living in household or elsewhere):	
	Whether cohort members' partner is parent of cohort members' natural children	Excluded
	Contact between children and absent parents	Excluded
	Ethnicity	Excluded
	National Identity	Included
	Whether English is usual language spoken at home	Included

Housing	**Housing history:**	
	Whether living at same address as at last interview	Excluded
	Details of all previous addresses:	
	Dates of moves	Excluded
	Reasons for moving	Excluded
	Postcodes of addresses (or town if not known)	Excluded
	Current home:	
	Type of accomodation (flat/terraced house etc)	Included
	Number of rooms	Included
	Tenure	Included
Relationships	**Relationship history:**	
	Information collected about all co-habiting relationships since previous interview (up to and including current partner):	
	When started living together	Excluded
	Whether male or female	Excluded
	Age of partner when started living together	Excluded
	Legal status of partner when moved in together	Excluded
	Whether got married	Excluded
	When got married	Excluded
	When stopped living together (ex-partners only)	Excluded
	Whether got divorced (ex-partners only)	Excluded
	When got divorced (ex-partners only)	Excluded
	When partner died (ex-partners only)	Excluded
	Other relationships:	
	(Asked of those with no co-habiting relationship since time of last interview)	
	Whether in a relationship	Included
	Length of relationship	Included
	Sex of partner	Included

Pregnancies	Whether been pregnant (or fathered a pregnancy) since previous interview	Included
	For those who have been pregnant: Number of pregnancies since previous interview	Included
	For each pregnancy: Number of babies carried	Excluded
	For each baby carried: Outcome of pregnancy (Live birth, stillbirth, miscarriage, termination, currently pregnant)	Excluded
	For each live birth:	
	Whether child lives in household	Excluded
	Sex of child	Excluded
	Weight at birth	Excluded
	Date of birth	Excluded
	Whether anything wrong at birth (and details of the problem if so)	Excluded
	Whether current partner is other parent	Excluded
Adopted Children	**Adopted Children**	
	Age of child when adopted	Excluded
	Whether child adopted by both cohort member and partner	Excluded
	Whether child is cohort member's partner's natural child	Excluded
Absent and Older Children	**Absent Children:**	
	Who child lives with	Excluded
	When child last lived with cohort member	Excluded
	Frequency of contact with child	Excluded

	Whether cohort member contributes to child's maintenance (for children aged 18 or under)	Excluded
	Older children (16+):	
	Economic activity of child	Excluded
	Whether still in full-time education / age left full-time education	Excluded
	Highest qualification	Excluded
	Whether cohort member's child has had any (natural) children of their own (and number of children).	Excluded
	Age at which cohort member's child first had a child of their own	Excluded
Other relatives	Whether cohort member's mother is still alive	Included
	Age when mother died	Excluded
	Whether cohort member's father is still alive	Included
	Age when father died	Excluded
	Whether parents live together	Included
	Where parents live (postcode or town)	Excluded
	Frequency of contact with parents (face-to-face and via telephone)	Included
	Closeness of relationship with mother/father	Included
	Worries about mother / father	Included
	Assistance given to parents	Included
	Hours spent caring for parents (per week)	Included
	Whether partner's mother is alive	Included
	Whether partner's father is alive	Included
	Assistance given to partner's parents	Included
	Hours spent caring for partner's parents (per week)	Included
Family Income	Benefits received by cohort member and partner (type of benefit, amount received and details of who receives benefit)	Included
	Other sources of income	Included
	Savings and investments (Types of savings and investments held and total amount)	Included
	Assessment of current financial situation	Included

	Comparison of current financial position with position one year ago	Included
	Expected financial position in one year's time	Included
Employment	**Economic Activity History:**	
	Details of all periods of economic activity:	
	Dates	Excluded
	If employed: Job details to derive SOC	Excluded
	Reason for changing job	Excluded
	Current employment:	
	Number of jobs	Included
	Whether been promoted (if still working for same employer as at last interview)	Included
	Job details (SOC)	Excluded
	If an employee:	
	Whether has managerial / supervisory duties	Included
	Number of employees where works	Included
	Whether a private or public sector organisation	Included
	Net and gross pay – amount and period covered	Included
	Number of hours worked (excluding overtime)	Included
	Whether works overtime	Included
	Number of hours paid overtime per week	Included
	Number of hours unpaid overtime per week	Included
	Total hours worked per week	Included
	Frequency of working in the evening (6 p.m. to 10 p.m.)	Included
	Frequency of working at night (10 p.m. to 4 a.m.)	Included
	Frequency of working in the early morning (4 a.m to 7 a.m.)	Included
	Frequency of working weekends	Included
	Whether job is permanent, temporary or fixed term contract	Included

	If self-employed:	Included
	Whether works on own or with employees	
	How many people employed	Included
	Number of hours worked per week	Included
	Frequency of working in the evening etc. (as above)	Included
	Type of self-employment	Included
	Whether draws up profit and loss accounts	Included
	Net profit in last 12 months (or most recent period for which figures available)	Included
	How much earned before tax in last 12 months (or most recent period for which figures available)	Included
	If employed / self-employed:	
	Travel to work time	Included
	Job satisfaction	Included
	Security of current job	Included
	Whether demands of work interfere with the demands of family and home life	Included
	Whether expects to be doing the same job in a year's time	Included
	If unemployed:	
	Reason for being unemployed	Included
	Whether currently looking for paid work	Included
	Whether would like to be in paid work	Included
	Income from any other paid work, odd-jobs, casual work etc.	Included
	Hours worked for this income	Included
	If looking after children:	
	Whether enjoys being at home looking after children	Included
	Whether would prefer to be in employment	Included
	Main reasons for not working at the moment (if would prefer to be employed)	Included
	Pensions:	
	Pensions currently contributing to or have contributed to in the past	Included

	Private health insurance:	
	Whether has private health insurance	Included
	Whether paid for directly or via employer	Included
Partner	Age partner left full-time education	Excluded
	Partner's current economic activity status	Included
	Partner's job details (SOC)	Excluded
	Partner's net pay	Included
Qualifications	Age left full-time education (if not previously collected)	Excluded
	Qualifications achieved in reference period	
	For each type of qualification:	Precise details excluded but summary variable included
	Number of qualifications	
	Subject	
	Grade	
	Date qualification achieved	
	Whether studied full or part-time	
	Who paid fees	
Work Related Training	Whether been offered work-related training in reference period	Included
	Number of courses (lasting 20 hours or more) attended in reference period	Included
	Other courses attended for interest or leisure	Included
Computer use	Whether has a computer at home	Included
	Frequency of use of home computer	Included
	Whether uses computer at work	Included
	Frequency of use of computer at work	Included
	Self-rating of computer skills	Included
	Whether would like to improve computer skills	Included
	Whether computer skills limit type of work that can do	Included
	Whether accesses internet for reasons other than work	Included
Health	Self-rating of general health	Included
	Rating of health in comparison with one year ago	Included

	Whether currently suffering from any health conditions (17 broad conditions with follow-up questions to identify specific conditions plus details of any '**other' conditions)**	Included
	Whether seen a doctor about each condition in the last 12 months	Included
	Whether registered disabled	Included
	Whether health limits daily activities (in comparison to most people of same age)	Included
	Number of days spent at hospital as a day-patient or in-patient since previous interview	Included
Cognitive Function	Word-list recall test	Included
	Animal naming task	Included
	Letter cancellation task	Excluded
	Delayed word-list recall test	Included
Smoking and Drinking	**Smoking:**	
	Whether currently smokes regularly or occasionally, or has smoked in the past or has never smoked	Included
	Number of cigarettes smoked per day	Included
	Age started smoking	Included
	Age stopped smoking (if applicable)	Included
	Drinking:	
	Regularity of drinking an alcoholic drink	Included
	Amount of beer drank in last week	Included
	Amount of spirits drank in last week	Included
	Amount of wine drank in last week	Included
	Amount of sherry drank in last week	Included
	Other alcoholic drinks drank in last week	Included
Exercise	Whether regularly takes part in physical activities	Included
	Frequency of taking part in physical activities	Included
	Whether gets out of breath or sweaty when taking part in physical activities	Included
	Amount of physical work involved in cohort member's job (if applicable)	Included

Height and weight	Height	Excluded
	Weight	Included
	Self-assessment of whether about the right weight, underweight or overweight	Included
	Whether currently trying to change weight	Included
Social Participation	**Membership of organisations:**	
	Current and previous membership of organisations	Included
	Regularity of taking part in activities of these organisations	Included
	Political activity:	
	Whether has attended a public meeting, taken part in a public demonstration or protest, or signed a petition in the last 12 months	Included
	Religion:	
	Whether as a child, cohort member was raised according to any particular religion and if so, which religion	Included
	Whether currently practices any religion and if so, which religion.	Included
	How often attends any kind of religious service or meeting	Included
Social Participation	How much cohort member could rely on people around them if they were sick in bed or needed to talk about problems and private feelings	Included
	Whether has visited friends in last 2 weeks	Included
	Whether has had visits from friends in last 2 weeks	Included
	How often has had contact with friends via phone or letter	Included
	Whether feels that most people can be trusted or that you can't be too careful when dealing with people	Included

CASI Self-completion	Car ownership	Included
	Voting behaviour and party support:	
	Whether voted in last general election	Included
	Who voted for	Included
	Party who would vote for if there was an election tomorrow	Included
	Strength of support for party who would vote for	Included
	Level of interest in politics	Included
	Menopause (women only):	
	Use of contraception	Excluded
	Whether has had hysterectomy or oophorectomy	Excluded
	Age had operation (if applicable)	Excluded
	Whether has had period or menstrual bleeding in last 12 months	Excluded
	Reason periods have stopped (if applicable)	Excluded
	Whether has had period in last 3 months	Excluded
	Date of last period	Excluded
	Whether regularity of periods has changed in last few years	Excluded
	When change in regularity of periods was noticed	Excluded
	Whether has experienced symptoms of menopause in the last 12 months	Included
	Whether currently on HRT	Excluded
	Whether has ever had HRT	Excluded
	When HRT started (if applicable)	Excluded
	Whether periods had stopped before starting HRT	Excluded
	Date of last period before HRT	Excluded
	AUDIT Scale (data collected via self-completion):	
	A scale of questions designed to identify problematic drinking behaviour	Included
	Malaise:	
	A series of questions designed to detect levels of emotional disturbance, well-being and stress	Included

	Marital / relationship satisfaction: Happiness of relationship	Included
	Domestic division of labour: How cooking, shopping, cleaning, laundry and ironing, household repairs / DIY, financial management and childcare is shared between cohort member and partner (if applicable)	Included
	Commitment to work: Kanugo's scale of job involvement (4 items from a 10 item scale)	Included
	Attitudes towards pensions / retirement Questions from DWP's Attitudes towards pensions study 6 item agree / disagree scale	Included
	Retrospective questions about childhood:	
	Number of people in household at age 11	Excluded
	Number of rooms in household at age 11	Excluded
	Efficacy / life satisfaction:	
	Whether cohort member feels that they usually get what they want out of life	Included
	Whether cohort member feels that they usually have a free choice or control over their life	Included
	Whether cohort member feels that they can usually run their lives as they wish	Included
	Satisfaction with way life has turned out so far	Included
	Expected level of satisfaction in ten years time	Included

Appendix 2 - Derived variables

2.1 List of derived and recoded variables

Variable name	Variable label
nd8maliv	(Derived) Whether CM's mother alive (incl prev swp data)
nd8paliv	(Derived) Whether CM's father alive (incl prev swp data)
n8Econ02	(Recoded) CM's current economic activity
n8pecac2	(Recoded) CM partner's current economic activity
n8panyb2	(Recoded) Whether partner / spouse had any paid job in the last three months
nd8achq	(Derived) Highest Academic Qualification (1991,2000,2004,2008 data)
nd8anvq	(Derived) Highest NVQ level - Acad qual (1991,2000,2004,2008 data)
nd8aghm	(Derived) Highest Acad qual - GHM measure (1991,2000,2004,2008 data)
nd8vnvq	(Derived) Highest NVQ Level - Voc Qual(1991,2000,2004,2008 data)
nd8nvq	(Derived) Highest NVQ Lev - Acad/Voc Qual (1991,2000,2004,2008 data)
nd8audit	(Derived) Total AUDIT score
nd8audg	(Derived) AUDIT Group
nd8mal	(Derived) Total Malaise score (9 questions)
nd8malg	(Derived) Total Malaise score - grouped
nd8jobinv	(Derived) Overall job involvement score (1-6)
nd8ext	(Derived) IPIP Personality Inventory - Extraversion score 5-50
nd8agr	(Derived) IPIP Personality Inventory - Agreeableness score 5-50
nd8con	(Derived) IPIP Personality Inventory - Conscientiousness score 5-50
nd8emo	(Derived) IPIP Personality Inventory - Emotional Stability score 5-50
nd8int	(Derived) IPIP Personality Inventory - Intellect Score 5-50
nd8wemwbs	(Derived) Warwick Edinburgh Mental Well-Being Scale
nd8phhe	(Derived) SF-36 Physical functioning score
nd8rlmp	(Derived) SF-36 Role-limitations due to physical health
nd8rlme	(Derived) SF-36 Role-limitations due to emotional problems
nd8enfa	(Derived) SF-36 Energy/fatigue score
nd8emwb	(Derived) SF-36 Emotional Well-Being score
nd8socf	(Derived) SF-36 Social Functioning score
nd8pain	(Derived) SF-36 Pain score
nd8genh	(Derived) SF-36 General health score
nd8casp14	(Derived) Overall CASP-14 Quality of Life Score

2.2 Syntax used to derive scores to questionnaire scales

2.2.1   Kanungo's scale of job involvement

**** Scores for Kanungo work involvement items 1,3 and 5 are reversed in order for higher score to represent higher job involvement

recode n8comjb1 n8comjb3 n8comjb5 (6=1)(5=2)(4=3)(3=4)(2=5)(1=6) (else=copy) into n8comjb1r n8comjb3r n8comjb5r .
execute.

value labels n8comjb1r n8comjb3r n8comjb5r (1) strongly disagree (2) disagree (3) somewhat disagree (4) somewhat agree (5) agree (6) strongly agree.

var labels n8comjb1r CASI: Whethr most imp things in CMs life relate to presjob (reverse scored)'.
var labels n8comjb3r 'CASI: Whether CM personally involved in his/her job (reverse scored)'.
var labels n8comjb5r 'CASI: Whether most CMs interests centre around their job (reverse scored)'.
missing values n8comjb1r n8comjb3r n8comjb5r (-9 thru - 1).
formats n8comjb1r n8comjb3r n8comjb5r (F2.0).
execute.

***** Job involvement scale computed as mean of 4 items with 6 responses (items 1,3 and 5 reverse
***** coded to give higher scores=more job involved). Those not answering all 4 items treated as missing.

compute nd8jobinv=(n8comjb1r+ n8comjb2+n8comjb3r+n8comjb5r)/4.
execute.

variable labels nd8jobinv '(Derived) Overall job involvement score (1-6)'.
missing values n8comjb1r ().
execute.
if n8comjb1r = -1 nd8jobinv = -1.
execute.
missing values n8comjb1r (-9 thru -1).
if (sysmis(nd8jobinv)) nd8jobinv = -2.
value labels nd8jobinv -1 'Item not applicable: not employed' -2 'Incomplete information'.
missing values nd8jobinv (-1,-2).
execute.

delete variables n8comjb1r n8comjb3r n8comjb5r.

2.2.2   Malaise score (9-item)

****Count number of items answered yes.

count nd8mal = n8mal02 n8mal03 n8mal05 n8mal09 n8mal12 n8mal14 n8mal16 n8mal20 n8mal21 (1).

****exclude cases with sufficient missing values to potentially push them into high malaise group (Scores of 4+).

count n8miss = n8mal02 n8mal03 n8mal05 n8mal09 n8mal12 n8mal14 n8mal16 n8mal20 n8mal21 (-9,-8,-2,-1).

if ((nd8mal + n8miss) ge 4 and n8miss>0) nd8mal = -2.

****recode into 2-category variable - Low malaise and high malaise.

recode nd8mal (0 thru 3=1) (4 thru highest = 2) (-2=-2) into nd8malg.
missing values nd8mal nd8malg (-2).
variable labels nd8mal '(Derived) Total Malaise score (9 questions)'.
variable labels nd8malg '(Derived) Total Malaise score - grouped'.
value labels nd8mal -2 'incomplete info' .
value labels nd8malg 1'Low malaise 0-3' 2' High malaise 4+' -2'incomplete info' .
exe.

delete variables n8miss.

2.2.3   **AUDIT - Alcohol Use Disorders Identification Test**

NOTE: The first official AUDIT question covering regularity of drinking has a five point scale (Never, Monthly or less, 2-4 times a month, 2-3 times a week, 4 or more times a week). This question was replaced with variable 'n8drinks' in the Smoking and Drinking module; n8drinks had an eight point scale (On most days, 2 to 3 days a week, once a week, 2 to 3 times a month, once a month, less often or only on special occasions, never nowadays, never had an alc. drink). The variable 'n8drinks' was recoded to match (as closely as possible the first AUDIT question).

Additionally, an error in the programming of the CASI module meant that the wording of the answer codes for n8Audi02 to n8Audi08 differed slightly from the official version of the AUDIT questionnaire.

AUDIT questions 2 to 8 use the following codeframe: Never, Less than monthly, Monthly, Weekly, Daily or almost daily. However the programming error meant that questions n8audi02 to n8audi08 used the following codeframe: Never, Less than monthly, Monthly, Two to three times a week, Four or more times a week. When calculating overall AUDIT scores 'two to three times per week' has been treated as 'weekly' and 'four or more times a week' has been treated as 'daily or almost daily'.

The two code-frames are roughly comparable but the differences could affect the way the questions were answered meaning AUDIT scores may not be strictly comparable to AUDIT scores collected using the official wording.

recode n8drinks (1=4) (2=3) (3=2) (4=2) (5=1) (6=1) (7=0) (8=0) (else = copy) into nd8audit1.
recode n8Audi02 n8Audi03 n8Audi04 n8Audi05 n8Audi06 n8Audi07 n8Audi08 (1=0) (2=1) (3=2) (4=3) (5=4) (else = copy) into nd8audit2 nd8audit3 nd8audit4 nd8audit5 nd8audit6 nd8audit7 nd8audit8.
recode n8Audi09 n8Audi10 (1=0) (2=2) (3=4) (else = copy) into nd8audit9 nd8audit10.

missing values nd8audit1 nd8audit2 nd8audit3 nd8audit4 nd8audit5 nd8audit6 nd8audit7 nd8audit8 nd8audit9 nd8audit10 (-9 thru - 1).

compute nd8audit = nd8audit1 + nd8audit2 + nd8audit3 + nd8audit4 + nd8audit5 + nd8audit6 + nd8audit7 + nd8audit8 + nd8audit9 + nd8audit10.

if n8drinks = 7 or n8drinks = 8 nd8audit = -1.
if (sysmis(nd8audit)) nd8audit = -2.

execute.

****Scores of 8 or more are likely to be signs of harmful or hazardous drinking.
****Scores of 13 or more for women and 15 or more for men are likely to indicate alcohol dependence.

if nd8audit < 8 nd8audg = 1.

if ((n8cmsex = 1 AND (nd8audit ge 8 and nd8audit le 14))) OR ((n8cmsex = 2 and (nd8audit ge 8 and nd8audit le 12))) nd8audg = 2.
if (n8cmsex = 1 and nd8audit ge 15) or (n8cmsex = 2 and nd8audit ge 13) nd8audg = 3.
if nd8audit = -1 nd8audg = -1.
if nd8audit = -2 nd8audg = -2.
exe.

missing values nd8audit nd8audg (-1,-2).

variable labels nd8audit '(Derived) Total AUDIT score'/ nd8audg '(Derived) AUDIT Group'.
value labels nd8audit -1 'Item not applicable - does not drink alcohol' -2 'Incomplete information'.
value labels nd8audg -1 'Item not applicable - does not drink alcohol' -2 'Incomplete information' 1
'Unproblematic drinking' 2 'Harmful or hazardous drinking' 3 'Alcohol dependent'.

delete variables nd8audit1 nd8audit2 nd8audit3 nd8audit4 nd8audit5 nd8audit6 nd8audit7 nd8audit8
nd8audit9 nd8audit10.

2.2.4 IPIP Personality Inventory – The Big Five

****Extraversion

recode n8scq2f n8scq2p n8scq2z n8scq2jj (1=5) (2=4) (3=3) (4=2) (5=1) (else = copy) into nd8ext1
nd8ext2 nd8ext3 nd8ext4.
recode n8scq2a n8scq2k n8scq2u n8scq2ee n8scq2oo (else = copy) into nd8ext5 nd8ext6 nd8ext7
nd8ext9 nd8ext10.

count extmiss = nd8ext1 nd8ext2 nd8ext3 nd8ext4 nd8ext5 nd8ext6 nd8ext7 nd8ext9 nd8ext10 (-1,-9).

compute nd8ext = nd8ext1+ nd8ext2+ nd8ext3+ nd8ext4+ nd8ext5+ nd8ext6+ nd8ext7+ nd8ext9+
nd8ext10.
if extmiss > 0 nd8ext = -2.
variable labels nd8ext '(Derived) IPIP Personality Inventory - Extraversion score 5-50'.
value labels nd8ext -2 'incomplete information'.
missing values nd8ext (-2).
exe.

****Agreeableness

recode n8scq2b n8scq2l n8scq2v n8scq2ff (1=5) (2=4) (3=3) (4=2) (5=1) (else = copy) into nd8agr1
nd8agr2 nd8agr3 nd8agr4.
recode n8scq2g n8scq2q n8scq2aa n8scq2kk n8scq2pp n8scq2uu (else = copy) into nd8agr5 nd8agr6
nd8agr7 nd8agr8 nd8agr9 nd8agr10.
count agrmiss = nd8agr1 nd8agr2 nd8agr3 nd8agr4 nd8agr5 nd8agr6 nd8agr7 nd8agr8 nd8agr9
nd8agr10 (-9,-1).
compute nd8agr = nd8agr1+ nd8agr2+ nd8agr3+ nd8agr4+ nd8agr5+ nd8agr6+ nd8agr7+ nd8agr9+
nd8agr10.
if agrmiss > 0 nd8agr = -2.
variable labels nd8agr '(Derived) IPIP Personality Inventory - Agreeableness score 5-50'.
value labels nd8agr -2 'incomplete information'.
missing values nd8agr (-2).
exe.

****Conscientiousness

recode n8scq2h n8scq2r n8scq2bb n8scq2ll (1=5) (2=4) (3=3) (4=2) (5=1) (else = copy) into nd8con1
nd8con2 nd8con3 nd8con4.
recode n8scq2c n8scq2m n8scq2w n8scq2gg n8scq2qq n8scq2vv (else = copy) into nd8con5 nd8con6
nd8con7 nd8con8 nd8con9 nd8con10.
count conmiss = nd8con1 nd8con2 nd8con3 nd8con4 nd8con5 nd8con6 nd8con7 nd8con8 nd8con9
nd8con10 (-9,-1).
compute nd8con = nd8con1+ nd8con2+ nd8con3+ nd8con4+ nd8con5+ nd8con6+ nd8con7+ nd8con9+
nd8con10.
if conmiss > 0 nd8con = -2.
variable labels nd8con '(Derived) IPIP Personality Inventory - Conscientiousness score 5-50'.
value labels nd8con -2 'incomplete information'.

missing values nd8con (-2).
exe.

****Emotional Stability

recode n8scq2d n8scq2n n8scq2x n8scq2cc n8scq2hh n8scq2mm n8scq2rr n8scq2ww (1=5) (2=4) (3=3)
(4=2) (5=1) (else = copy) into nd8emo1 nd8emo2 nd8emo3 nd8emo4 nd8emo5 nd8emo6 nd8emo7
nd8emo8.
recode n8scq2i n8scq2s   (else = copy) into nd8emo9 nd8emo10.
count emomiss = nd8emo1 nd8emo2 nd8emo3 nd8emo4 nd8emo5 nd8emo6 nd8emo7 nd8emo8
nd8emo9 nd8emo10 (-9,-1).
compute nd8emo = nd8emo1+ nd8emo2+ nd8emo3+ nd8emo4+ nd8emo5+ nd8emo6+ nd8emo7+
nd8emo9+ nd8emo10.
if emomiss > 0 nd8emo = -2.
variable labels nd8emo '(Derived) IPIP Personality Inventory - Emotional Stability score 5-50'.
value labels nd8emo -2 'incomplete information'.
missing values nd8emo (-2).
exe.

****Intellect

recode n8scq2j n8scq2t n8scq2dd (1=5) (2=4) (3=3) (4=2) (5=1)(else = copy) into nd8int1 nd8int2
nd8int3.
recode n8scq2e n8scq2o n8scq2y n8scq2ii n8scq2nn n8scq2ss n8scq2xx (else = copy) into nd8int4
nd8int5 nd8int6 nd8int7 nd8int8 nd8int9 nd8int10.
count intmiss = nd8int1 nd8int2 nd8int3 nd8int4 nd8int5 nd8int6 nd8int7 nd8int8 nd8int9 nd8int10 (-9,-1).
compute nd8int = nd8int1+ nd8int2+ nd8int3+ nd8int4+ nd8int5+ nd8int6+ nd8int7+ nd8int9+ nd8int10.
if intmiss > 0 nd8int = -2.
variable labels nd8int '(Derived) IPIP Personality Inventory - Intellect Score 5-50'.
value labels nd8int -2 'incomplete information'.
missing values nd8int (-2).
exe.

delete variables nd8ext1 nd8ext2 nd8ext3 nd8ext4 nd8ext5 nd8ext6 nd8ext7 nd8ext9 nd8ext10 nd8agr1
nd8agr2 nd8agr3 nd8agr4 nd8agr5 nd8agr6 nd8agr7 nd8agr8 nd8agr9 nd8agr10
nd8con1 nd8con2 nd8con3 nd8con4 nd8con5 nd8con6 nd8con7 nd8con8 nd8con9 nd8con10 nd8emo1
nd8emo2 nd8emo3 nd8emo4 nd8emo5 nd8emo6 nd8emo7 nd8emo8 nd8emo9 nd8emo10
nd8int1 nd8int2 nd8int3 nd8int4 nd8int5 nd8int6 nd8int7 nd8int8 nd8int9 nd8int10 extmiss agrmiss
conmiss emomiss intmiss.

2.2.5   Warwick Edinburgh Mental Well Being Scale (WEMWBS)

compute nd8wemwbs = n8scq3a + n8scq3b + n8scq3c + n8scq3d + n8scq3e + n8scq3f + n8scq3g +
n8scq3h + n8scq3i + n8scq3j + n8scq3k + n8scq3l + n8scq3m + n8scq3n.
if (sysmis(nd8wemwbs)) nd8wemwbs = -2.
missing values nd8wemwbs (-2).
value labels nd8wemwbs -2 'incomplete information'.
variable labels nd8wemwbs '(Derived) Warwick Edinburgh Mental Well-Being Scale'.

2.2.6   SF-36

recode n8HlthGn n8khlstt n8scq4a n8scq4b n8scq4c n8scq4d n8scq4e n8scq4f n8scq4g n8scq4h
n8scq4i n8scq4j n8scq5a n8scq5b n8scq5c n8scq5d n8scq6a n8scq6b n8scq6c n8scq7 n8scq8
n8scq9 n8scq10a n8scq10b n8scq10c n8scq10d n8scq10e n8scq10f n8scq10g n8scq10h n8scq10i
n8scq10j n8scq11a n8scq11b n8scq11c n8scq11d (else = copy) into sf361 sf362 sf363 sf364 sf365

sf366 sf367 sf368 sf369 sf3610 sf3611 sf3612 sf3613 sf3614 sf3615 sf3616 sf3617 sf3618 sf3619 sf3620
sf3621 sf3622 sf3623 sf3624 sf3625 sf3626 sf3627 sf3628 sf3629 sf3630 sf3631 sf3632 sf3633 sf3634
sf3635 sf3636.
exe.

****Recode so that all variables are on a scale from 0 to 100 with 100 indicating highest levels of health.

recode sf361 sf362 sf3620 sf3622 sf3634 sf3636 (1=100) (2=75) (3=50) (4=25) (5=0) (else = copy).
recode sf363 sf364 sf365 sf366 sf367 sf368 sf369 sf3610 sf3611 sf3612 (1=0) (2=50) (3=100) (else = copy).
recode sf3613 sf3614 sf3615 sf3616 sf3617 sf3618 sf3619 (1=0) (2=100) (else = copy).
recode sf3621 sf3623 sf3626 sf3627 sf3630 (1=100) (2=80) (3=60) (4=40) (5=20) (6=0) (else = copy).
recode sf3624 sf3625 sf3628 sf3629 sf3631 (1=0) (2=20) (3=40) (4=60) (5=80) (6=100) (else = copy).
recode sf3632 sf3633 sf3635 (1=0) (2=25) (3=50) (4=75) (5=100) (else=copy).
exe.

****The following sections calculate scores for 8 domains.
****In each section the relevant recoded variables are summed and then divided by the number of
questions answered.

****Physical functioning

count physmiss = sf363 sf364 sf365 sf366 sf367 sf368 sf369 sf3610 sf3611 sf3612 (-8,-9).
recode sf363 sf364 sf365 sf366 sf367 sf368 sf369 sf3610 sf3611 sf3612 (-8=0) (-9=0).
if physmiss < 10 nd8phhe = (sf363 + sf364 + sf365 + sf366 + sf367 + sf368 + sf369 + sf3610 + sf3611 +
sf3612) / (10-physmiss).
if physmiss = 10 nd8phhe = -2.
missing values nd8phhe (-2).
variable labels nd8phhe '(Derived) SF-36 Physical functioning score'.
value labels nd8phhe -2 'Information incomplete'.
exe.

****Role limitations due to physical health

count rlmpmiss = sf3613 sf3614 sf3615 sf3616 (-8,-9).
recode sf3613 sf3614 sf3615 sf3616 (-8=0) (-9=0).
if rlmpmiss < 4 nd8rlmp = (sf3613 + sf3614 + sf3615 + sf3616) / (4-rlmpmiss).
if rlmpmiss = 4 nd8rlmp = -2.
missing values nd8rlmp (-2).
variable labels nd8rlmp '(Derived) SF-36 Role-limitations due to physical health'.
value labels nd8rlmp -2 'Information incomplete'.
exe.

****Role limitations due to emotional problems

count rlmemiss = sf3617 sf3618 sf3619 (-8,-9).
recode sf3617 sf3618 sf3619 (-8=0) (-9=0).
if rlmemiss < 3 nd8rlme = (sf3617 + sf3618 + sf3619) / (3-rlmemiss).
if rlmemiss = 3 nd8rlme = -2.
missing values nd8rlme (-2).
variable labels nd8rlme '(Derived) SF-36 Role-limitations due to emotional problems'.
value labels nd8rlme -2 'Information incomplete'.
exe.

****Energy/fatigue

count enfamiss = sf3623 sf3627 sf3629 sf3631 (-8,-9).
recode sf3623 sf3627 sf3629 sf3631 (-8=0) (-9=0).
if enfamiss < 4 nd8enfa = (sf3623 + sf3627 + sf3629 + sf3631) / (4-enfamiss).
if enfamiss = 4 nd8enfa = -2.
missing values nd8enfa (-2).
variable labels nd8enfa '(Derived) SF-36 Energy/fatigue score'.
value labels nd8enfa -2 'Information incomplete'.
exe.

****Emotional well-being

count emwbmiss = sf3624 sf3625 sf3626 sf3628 sf3630 (-8,-9).
recode sf3624 sf3625 sf3626 sf3628 sf3630 (-8=0) (-9=0).
if emwbmiss < 5 nd8emwb = (sf3624 +sf3625+ sf3626+ sf3628+ sf3630) / (5-emwbmiss).
if emwbmiss = 5 nd8emwb = -2.
missing values nd8emwb (-2).
variable labels nd8emwb '(Derived) SF-36 Emotional Well-Being score'.
value labels nd8emwb -2 'Information incomplete'.
exe.

****Social Functioning

count socfmiss = sf3620 sf3632 (-8,-9).
recode sf3620 sf3632 (-8=0) (-9=0).
if socfmiss < 2 nd8socf = (sf3620 + sf3632) / (2-socfmiss).
if socfmiss = 2 nd8socf = -2.
missing values nd8socf (-2).
variable labels nd8socf '(Derived) SF-36 Social Functioning score'.
value labels nd8socf -2 'Information incomplete'.
exe.

****Pain

count painmiss = sf3621 sf3622 (-8,-9).
recode sf3621 sf3622 (-8=0) (-9=0).
if painmiss < 2 nd8pain = (sf3621 + sf3622) / (2-painmiss).
if painmiss = 2 nd8pain = -2.
missing values nd8pain (-2).
variable labels nd8pain '(Derived) SF-36 Pain score'.
value labels nd8pain -2 'Information incomplete'.
exe.

****General Health

count genhmiss = sf361 sf3633 sf3634 sf3635 sf3636 (-8,-9).
recode sf361 sf3633 sf3634 sf3635 sf3636 (-8=0) (-9=0).
if genhmiss < 5 nd8genh = (sf361 +sf3633+ sf3634+ sf3635+ sf3636) / (5-genhmiss).
if genhmiss = 5 nd8genh = -2.
missing values nd8genh (-2).
variable labels nd8genh '(Derived) SF-36 General health score'.
value labels nd8genh -2 'Information incomplete'.
exe.

****Remove surplus variables

delete variables sf361 sf362 sf363 sf364 sf365 sf366 sf367 sf368 sf369 sf3610 sf3611 sf3612 sf3613 sf3614 sf3615 sf3616 sf3617 sf3618 sf3619 sf3620 sf3621 sf3622 sf3623 sf3624 sf3625 sf3626 sf3627 sf3628 sf3629 sf3630 sf3631 sf3632 sf3633 sf3634 sf3635 sf3636 physmiss rlmpmiss rlmemiss enfamiss emwbmiss socfmiss painmiss genhmiss.
exe.

2.2.7   CASP-14

****Recode to 0-3 scale (reverse coding where appropriate so that higher scores represent higher quality of life).

recode n8scq19a n8scq19b n8scq19c n8scq19e n8scq19g (1=0) (2=1) (3=2) (4=3) (else = copy) into casp1 casp2 casp3 casp5 casp7.
recode n8scq19d n8scq19f n8scq19h n8scq19i n8scq19j n8scq19k n8scq19l n8scq19m n8scq19n (1=3) (2=2) (3=1) (4=0) (else = copy)
into casp4 casp6 casp8 casp9 casp10 casp11 casp12 casp13 casp14.
exe.

****Count number of missing values

count caspmiss = casp1 casp2 casp3 casp5 casp7 casp4 casp6 casp8 casp9 casp10 casp11 casp12 casp13 casp14 (-9,-8,-1).

compute nd8casp14 = casp1+ casp2 + casp3 + casp5 + casp7 + casp4 + casp6 + casp8 + casp9 + casp10 + casp11 + casp12 + casp13 + casp14.
exe.

if caspmiss > 0 and caspmiss <14 nd8casp14 = -2.
if caspmiss = 14 nd8casp14 = -3.

missing values nd8casp14 (-9 thru -1).
value labels nd8casp14 -2 'Incomplete information' -3 'No questions answered'.
variable labels nd8casp14 '(Derived) Overall CASP-14 Quality of Life Score'.
exe.

delete variables casp1 casp2 casp3 casp5 casp7 casp4 casp6 casp8 casp9 casp10 casp11 casp12 casp13 casp14 caspmiss.