

**Health Survey for England**

**The health of older  
people**

**'05**

**User Guide**

# 1. Background

The data files contain data from Health Survey for England 2005 (HSE), the fifteenth year of a series of surveys designed to monitor trends in the nation's health. The 20045Health Survey was commissioned by the Information Centre and carried out by the Joint Health Surveys Unit of the *National Centre for Social Research* and the Department of Epidemiology and Public Health at Royal Free and University College Medical School.

The aims of the Health Survey series are:

- to provide annual data about the nation's health;
- to estimate the proportion of people in England with specified health conditions;
- to estimate the prevalence of certain risk factors associated with these conditions;
- to examine differences between population subgroups in their likelihood of having specific conditions or risk factors;
- to assess the frequency with which particular combinations of risk factors are found, and which groups these combinations most commonly occur;
- to monitor progress towards selected health targets;
- since 1995, to measure the height of children at different ages, replacing the National Study of Health and Growth;
- since 1995, monitor the prevalence of overweight and obesity in children.

## 2. Survey Design

The Health Survey for England 2005 was designed to provide data at both national and regional level about the population living in private households in England. The sample for the HSE 2005 comprised of three components: the core (general population) sample, a boost sample of people aged 65 and over (those living in institutions were not included) and a boost sample of children aged 2-15. The core sample was designed to be representative of the population living in private households in England and should be used for analyses at the national level.

A random sample of 720 postcode sectors was then selected with probability proportional to the total number of addresses within them. Once selected, the PSUs were randomly allocated to the 12 months of the year (60 per month) so that each quarter provided a nationally representative sample.

Within each selected postcode sector, a sample of 26 delivery points was selected, giving a total selected sample of 18,720 (720 x 26) addresses. From the 26 addresses within each postcode sector, ten were selected at random and allocated to the core sample, and the remaining 16 addresses were allocated to the boost sample.

All private households in the general population sample are eligible for inclusion in the survey (up to a maximum of three households per address). For the core sample, up to two children aged 0-15 are interviewed in each household, as well as up to 10 adults aged 16 and over.

At boost addresses interviewers screened for households containing at least one person of either of the age groups covered in the boost: persons aged 65 and over, or (for certain months) children aged 2-15 years. Because of funding restrictions, the boost included children only during January, February, October, November and December. At each household where people of the eligible ages were found, all persons aged 65 and over and up to two eligible children were selected by the interviewer for inclusion in the survey.

An interview with each eligible person was followed by a nurse visit both using computer assisted interviewing. The 2005 survey for adults focused on the health of older people. All adults were asked modules of questions on general health, alcohol consumption, smoking, fruit and vegetable consumption and complementary and alternative medicine. Older informants were also asked about use of health, dental and social care services, cardiovascular disease (CVD), chronic diseases and

quality of care, disabilities and falls. Older informants in the boost sample were asked a slightly shorter questionnaire, omitting questions about fruit and vegetable consumption and complementary and alternative medicines.

Children aged 13-15 were interviewed themselves, and parents of children aged 0-12 were asked about their children, with the child interview including questions on physical activity and fruit and vegetable consumption.

Interviewing was conducted throughout the year to take account of seasonal differences.

### 3. Documentation

The documentation has been organised into the following sections

- Interview (contains the CAPI documentation for household and individual questionnaires, nurse visit questionnaires, self-completion booklets and showcards)
- Data (contains the list of variables and list of derived variables)
- Other instructions (contains interviewer, nurse and coding & editing instructions).

### 4. Using the data

The 2005 data consists of one individual level file and one household level file:

HSE05ai.sav	13,297 records	contains data for all individuals in Household who gave a full interview. It contains information from the household questionnaire, main individual schedule, self-completions and the nurse visit (where one occurred).
-------------	----------------	--

HSE05ah.sav	17,095 records	contains data on household composition, sex, age and marital status for all individuals in co-operating households.
-------------	----------------	---

#### 4.1 Variables on the files

Each of the data files contain questionnaire variables (excluding variables used for administrative purposes) and derived variables. The variables included in the individual file are detailed in the “**List of Variables**” document in the data section of the documentation. This document is the best place to look at in order to plan your analysis. It includes:

- Major categories of variables (eg Accidents, Anthropometric measurements)
- Sub categories of variables (eg Attitudes to cycling, Major accidents within the Accidents category)
- Source of each variable (eg Individual questionnaire, Nurse visit, Derived variable etc.)

Once you have decided which variables to include in your analysis, you can look up details of the question wording using the interview section documentation (all variables on the data file are given by name in the copy of the interview schedules provided), or use the “**Derived Variables Specification**” document in the data section of the documentation for derived variables.

## 4.2 Weighting variables

### 4.2.1 Core sample

Prior to 2003, the weighting strategy for the core sample in the HSE was to apply selection weights only, and no attempt was made to reduce non-response bias through weighting. However, following a review of the weighting for the HSE, non-response weighting has been incorporated in the weighting strategy since 2003. The same strategy as in 2003 has been followed for weighting the HSE 2005 core general population sample data.. (For more detailed information on how the weights were produced see Health Survey for England 2005. Volume 5: Methodology and documentation).

Two sets of non-response weights have been generated for the general population sample: household weights which adjust for non-contact and refusal of households, and interview weights which also adjust for the additional non-response among individuals in participating households.

#### *Household weight*

The household weight (**wt\_hhld**) is a household level weight that corrects the distribution of household members to match population estimates for sex/age groups and GOR. These weights were generated using calibration weighting, with the household selection weights as starting values. (The household selection weights correct for where the limit of three households are selected at addresses with more than three.) Note that the population control totals used for the calibration weighting were the ONS projected mid-year population estimates for 2005, but with a small adjustment to exclude (our best estimate of) the population aged 65 and over living in communal establishments.

#### *Individual weight*

For analyses at the individual level, the weighting variable to use is **wt\_int**. These weights are generated separately for adults and children:

- for adults (aged 16 or more), the interview weights are a combination of the household weight and a component which adjusts the sample to reduce bias from individual non-response within households;
- for children (aged 0 to 15), the weights are generated from the household weights and the child selection weights – the selection weights correct for only including a maximum of two children in a household. The combined household and child selection weight were adjusted to ensure that the weighted age/sex distribution matched that of all children in co-operating households.

For analysis of children aged 0-15 in the Core sample, taking into account child selection only and not adjusting for non-response, the **child\_wt** variable can be used.

#### *Nurse weight*

To take into account non-response to the nurse section of the survey, a nurse weight has been generated (**wt\_nurse**) and should be used on all analysis of questions asked during the nurse visit.

#### *Blood weight*

Only individuals aged 65 and over were asked for a blood sample and therefore no core sample analysis of blood data should be undertaken. Consequently no blood weights for the core sample were required.

### 4.2.2 Child sample weights

The child sample is defined as all children aged 0-15 from the core sample and all children aged 2-15 from the boost sample addresses. Child weights were calculated to take into account, household and child selection, plus seasonal adjustment to compensate for any bias when the boost sample was only being carried out during the winter months. The final calibrated weight is called **wt\_intCH**.

There were no nurse visits for the child boost addresses and no blood samples taken for any of the children (core or boost); therefore, no nurse or blood weights were required.

#### **4.2.3 Older people sample – 65 and over**

The sample of older people comprised all adults aged 65 and over from either the core or boost sample. Household weights were not required as there only two cases of households having been selected in the field for a household containing a member of the elderly sample.

##### *Individual weight for older people*

The weighting strategy for the sample of older people was to calibrate the achieved sample to match ONS 2004 mid-year population estimates. For analysis of respondents aged 65 and over, the variable **Wt\_intEL** should be used.

##### *Nurse weight for older people*

To take into account non-response to the nurse visit for older adults, **wt\_nurEL** have been generated.

##### *Blood weight for older people*

Those respondents who had a blood sample taken (applies to respondents aged 65 and over only) have been allocated a blood weight. The variable is called **wt\_bldEL**.

#### **4.3 Multicoded questions**

Multicoded questions are stored in the archived HSE 2005 data sets in two ways. Multicoded questions, where for example the interviewer (or nurse) is instructed to "CODE ALL THAT APPLY" or where an open ended question has elicited more than one answer, were stored as array variables in the QUANTUM DBMS system which was used to read and edit the data. However, in SPSS (which was used for analysis and archiving the data) multicoded variables must be stored as 'flat' variables, coded either **by mention** or **by category**. Questions coded by mention are stored as categorical variables where the complete value set is repeated in each of the variables. Questions coded by category are stored as indicator variables where each value in the set is stored as its own variable. Both approaches have been used in the 2004 Health Survey.

As an example, question CONSBX1 on the 2004 adult nurse schedule is a "CODE ALL THAT APPLY" question which asks "Have you eaten, smoked, drunk alcohol or done any vigorous exercise in the past 30 minutes?". The code frame consists of five values:

- 1 - eaten
- 2 - smoked
- 3 - drunk alcohol
- 4 - done vigorous exercise
- 5 - none of these

If recorded by mention, four variables would record the (up to) four possible responses to the question assigning codes 1-5 in the first variable and codes 1-4 in each of the next three variables. In 2004, the variables CONSBX11-15 store the answer to this question by category as follows:

- CONSBX11 - coded 1 for those who ate in the last half hour and 0 for those that didn't.
- CONSBX12 - coded 1 for those who smoked in the last half hour and 0 for those that didn't.
- CONSBX13 - coded 1 for those who drank alcohol in the last half hour and 0 for those that didn't.
- CONSBX14 - coded 1 for those who did vigorous exercise in the last half hour and 0 for those that didn't.
- CONSBX15 - coded 1 for those who did none of the above in the last half hour and 0 for everyone else.

Because a respondent could have replied with more than one answer, that respondent could have a value 1 for a number of these variables (however, the nature of the question dictates that having a code 1 at CONSBX15 precludes having a code 1 at any of the variables CONSBX11 – CONSBX14). The missing values are the same across all six variables.

In most instances **by category variables** are denoted by a C after the original variable name, **by mention variables** are denoted by an M. Documentation for the CAPI questionnaires (household and individual) shows only the name of the first variable (which stores the number of mentions).

#### 4.4 Missing values conventions

- 1 Not applicable: Used to signify that a particular variable did not apply to a given respondent usually because of internal routing. For example, men in women only questions.
- 2 Schedule not applicable: Used mainly for variables on the self-completions when the respondent was not of the given age range, also used for children without legal guardians in the home who could not participate in the nurse schedule.
- 8 Don't know, Can't say.
- 9 No answer/ Refused

These conventions have also been applied to most of the derived variables. The derived variable specifications should be consulted for details.

#### 4.5 Valid cases

In the 2005 Health Survey report, as in previous reports, cases were excluded from the analysis of anthropometric and blood pressure measurements if their measurement was invalid. For example, those who had smoked, drunk, eaten, or exercised within 30 minutes of having their blood pressure taken were excluded from analysis as this can affect blood pressure.

## 5. HSE 2005 Report

Further information about the Health Survey for England 2005 is available in:

Craig R and Mindell J (eds). *Health Survey for England 2005. Volume1: General Health*. The Information Centre, Leeds, 2007.

Craig R and Mindell J (eds). *Health Survey for England 2005. Volume 2: Chronic Diseases*. The Information Centre, Leeds, 2007.

Craig R and Mindell J (eds). *Health Survey for England 2005. Volume3: Social Capital*. The Information Centre, Leeds, 2007.

Craig R and Mindell J (eds). *Health Survey for England 2005. Volume 4: Mental Health*. The Information Centre, Leeds, 2007.

Craig R and Mindell J (eds). *Health Survey for England 2005. Volume3: Methodology and Documentation..* The Information Centre, Leeds, 2007.

Or on the Information Centre website:

<http://www.ic.nhs.uk/pubs/hseupdate05>

For the general population, tables showing selected trends from 1993 to 2005 can be found on The Information Centre web page:

<http://www.ic.nhs.uk/statistics-and-data-collections/health-and-lifestyles/health-survey-for-england/health-survey-for-england--updating-of-trend-tables-to-include-2005-data>

## APPENDIX A

### 2005 HEALTH SURVEY FOR ENGLAND – CONTENTS

#### Household data

Household size, composition and relationships	Smoking in household
Accommodation tenure and number of bedrooms	Type of dwelling and area
Economic status/occupation of Household Reference Person	Car ownership
Household income	

#### Individual level information

	0-1	2-3	4	5-7	8-10	11-12	13-15	16-64	65+ Core	65+ Boost
<b>Interviewer visit</b>										
General health, longstanding illness, limiting longstanding illness, acute sickness, fractures	●	●	●	●	●	●	●	●	●	●
Use of health & dental services									●	●
Use of social care									●	●
Carers responsibilities								●	●	●
CVD, including use of services									●	●
Chronic disease & quality of care									●	●
Disabilities									●	●
Falls									●	●
Physical activity	●	●	●	●	●	●	●			
Smoking					● <sup>a</sup>	● <sup>a</sup>	● <sup>a</sup>	● <sup>b</sup>	●	● <sup>c</sup>
Drinking (seven day period)					● <sup>a</sup>	● <sup>a</sup>	● <sup>a</sup>	● <sup>b</sup>	●	●
Fruit and vegetable consumption				●	●	●	●	●	●	
Eating habits		●	●	●	●	●	●			
Complementary and alternative medicine								●	●	
Economic status/occupation, educational achievement								●	●	●
Ethnic origin	●	●	●	●	●	●	●	●	●	●
Social capital								● <sup>a</sup>	●	●
Height measurement		●	●	●	●	●	●	●	●	●
Weight measurement	●	●	●	●	●	●	●	●	●	●
Reported birth weight	●	●	●	●	●	●	●			
Cycling safety					● <sup>a</sup>	● <sup>a</sup>				
Psychosocial health (GHQ 12)							● <sup>a</sup>	● <sup>a</sup>	● <sup>a</sup>	● <sup>a</sup>
Euroqol general health (EQ5D)								● <sup>a</sup>	● <sup>a</sup>	● <sup>a</sup>
Geriatric depression score									● <sup>a</sup>	● <sup>a</sup>
Social support								● <sup>a</sup>	● <sup>a</sup>	● <sup>a</sup>
Strengths and difficulties			● <sup>e</sup>	● <sup>e</sup>	● <sup>e</sup>	● <sup>e</sup>	● <sup>e</sup>			
Perception of weight					● <sup>a</sup>	● <sup>a</sup>	● <sup>a</sup>			
Use of contraceptive pill								● <sup>a</sup>		
Hormone replacement therapy								● <sup>a,d</sup>		
Incontinence									● <sup>a</sup>	● <sup>a</sup>

<sup>a</sup> These modules were administered by self completion.

<sup>b</sup> This module was administered by self-completion for those aged 16-17 and some aged 18-24.

<sup>c</sup> Shortened smoking module for boost sample 65+

<sup>d</sup> This 18+ only (there are no HRT questions in the young adult self-completion).

<sup>e</sup> This module was asked by proxy and administered by self-completion for parents of 4-15 year olds.

	0-1	2-3	4	5-7	8-10	11-12	13-15	16-64	65+ Core	65+ Boost
<i>Nurse visit</i>										
Prescribed medicines and vitamin supplements	●	●	●	●	●	●	●	●	●	●
Nicotine replacements								●	●	●
Immunisations	●									
Blood pressure				●	●	●	●	●	●	●
Infant length	●									
Waist and hip circumference						●	●	●	●	●
Demi-span									●	●
Physical function - grip strength, walking speed , balance, chair rise									●	●
Blood sample – total & HDL cholesterol, ferritin, haemoglobin, glycated haemoglobin, fibrinogen, mean corpuscular volume, serum albumin, serum transferrin, vitamin D, vitamin B12									●	●
Saliva sample – cotinine			●	●	●	●	●			
Urine sample								●	●	