Health Survey for England

Respiratory health

2010

User Guide

A survey carried out on behalf of the NHS Information Centre

Joint Health Surveys Unit

NatCen Social Research

Department of Epidemiology and Public Health, University College London

CONTENTS

1. Background

The data files contain data from the Health Survey for England 2010 (HSE), the twentieth year of a series of surveys designed to monitor trends in the nation's health. The 2010 Health Survey was commissioned by The NHS Information Centre for health and social care and carried out by the Joint Health Surveys Unit of NatCen 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.

The main focus of the HSE in 2010 was respiratory disease and lung function. The survey also provided updates on core topics including smoking, drinking and fruit and vegetable consumption. Additional modules of questions were also included covering contraception and sexual health, well-being, kidney disease and dental health.

The report on this survey, including a detailed Methods and Documentation volume, is published at www.ic.nhs.uk/pubs/hse10report

2. Survey Design

The HSE 2010 included a general population sample of adults and children, representative of the whole population at both national and regional level, and a boost sample of children aged 2-15. For the general population sample, 8,736 addresses were randomly selected in 672 postcode sectors, issued over twelve months from January to December 2010. Where an address was found to have multiple dwelling units, one dwelling unit was selected at random

and where there were multiple households at a dwelling unit, one household was selected at random.

In each selected household, all individuals were eligible for inclusion in the survey. Where there were three or more children aged 0-15 in a household, two of the children were selected at random. A nurse visit was arranged for all participants who consented.

In addition to the core general population sample, a boost sample of children aged 2-15 was selected using 17,136 addresses, some in the same postcode sectors as the core sample and some in an additional 168 postcode sectors to supplement the sample obtained in the core sectors. As for the core sample, where there were three or more children in a household, two of the children were selected at random to limit the respondent burden for parents. There was no nurse follow up for this child boost sample.

A total of 8,420 adults and 5,692 children were interviewed, with 2,074 children from the core sample and 3,618 from the boost. A household response rate of 66% was achieved for the core sample, and 70% for the boost sample. Among the general population sample, 5,587 adults and 1,327 children had a nurse visit.

Height was measured for those aged two and over and weight for all participants. Nurses measured blood pressure (aged 5 and over), waist and hip circumference (aged 11 and over), and spirometry measurements (lung function measurements) for those aged 7 and over. Non-fasting blood samples (for the analysis of total and HDL cholesterol, glycated haemoglobin, serum creatinine and Vitamin D) and urine samples (for the analysis of sodium, potassium, creatinine, albumin and, for those aged 35 and over, melatonin) were collected from adults aged 16 and over. Saliva samples for cotinine analysis were collected from adults aged 16 and over and children aged 4-15. Nurses obtained written consent before taking samples from adults, and parents gave written consent for their children's samples. Consent was also obtained from adults to send results to their GPs, and from parents to send their children's results to their GPs.

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 2010 data consists of two files; one at individual level and one at household level:

HSE10ai.sav	14112 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).

HSE10ah.sav	21793 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 in order to plan your analysis. It includes:

- Major categories of variables (e.g. Blood Sample, General Health)
- Sub categories of variables (e.g. Measurements (within Blood Sample), Diabetes (within General health)
- Source of each variable (e.g. 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 Multicoded questions

Multicoded questions are stored in the archived HSE 2010 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 2010 Health Survey.

As an example, question CONSBX1 on the 2010 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 2010, the variables CONSBX11-15 store the answer to this question by category as follows:

> CONSBX11 - coded 1 for those who ate in the last 30 minutes and 0 for those that didn't.
> CONSBX12 - coded 1 for those who smoked in the last 30 minutes and 0 for those that didn't.
> CONSBX13 - coded 1 for those who drank alcohol in the last 30 minutes and 0 for

those that didn't.

CONSBX14 - coded 1 for those who did vigorous exercise in the last 30 minutes and 0 for those that didn't.

CONSBX15 - coded 1 for those who did none of the above in the last 30 minutes 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.

Documentation for the CAPI questionnaires (household and individual) shows only the name of the first variable (which stores the number of mentions). So, for the example given above this variable name is ConSubX.

4.3 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 or self completion variables when the respondent is not of the given age range.

-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.

4.4 Valid cases

In the 2010 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.

4.5 Notes about particular variables

4.5.1 Cholesterol results

From the 12th of April, 2010, the laboratory that carries out the analyses on the blood and urine samples taken during the HSE interview introduced new analytical equipment. This had no affect on most analytes, but has meant a slight change in the reference range for total and HDL cholesterol (see Table 28 in Volume 2 of the 2010 report at www.ic.nhs.uk/pubs/HSE10report). A flag variable has been computed (called CHOLFLAG) which shows whether a sample was tested before or after the equipment change, and analysts may make allowance for this difference if they wish to adjust the cholesterol results. Note that the difference (an average of 0.1mmol/L) was very small and in most analyses this difference is unlikely to be statistically significant.

4.5.2 Children's longstanding illness

Due to a routing error in the questionnaire, children (i.e. those aged 15 and under) were <u>not</u> asked the questions which relate to longstanding illness (variable names LongIll – LastYr).

4.5.3 Spirometry data

Details of how spirometry was measured are given in Chapters 3 and 5 of the 2010 report, and the nurse protocol is provided in Appendix B of Volume 2, Methods and documentation (www.ic.nhs.uk/pubs/HSE10report). There were a number of exclusion criteria for the spirometry module, including people with a very high pulse rate. The routing in the spirometry module excluded both those who had a high pulse rate reading, and those who did not have a valid reading. For a case to be routed to a value of LFOUTC=8 ('Not assigned') they had to fulfil the following criteria:

- Be aged under 7 **OR**;
- Be pregnant (NCPregJ=pregnant)**OR**;
- Be taking medication for tuberculosis (TBMed=yes) **OR**;
- Have an average of their second and third pulse measurements over 120 bpm (AvPulse>120) **OR**;
- Have any of a list of specified medical conditions (e.g. heart attack in the last three months, detached retina).

There were 118 adults and 41 children who were excluded from having their lung function measured because their second and/or third pulse measurement was missing, and it was computed as '999' by the CAPI program. Therefore, their average pulse rate was calculated

as >120 and they were excluded.

Each spirometry test session (consisting of all the manoeuvres attempted by the participant) was assigned a grade which denoted the session quality. This grade was based on the number of technically acceptable blows and the reproducibility of these. As an additional quality control, 25% of all measurements were over-read by an experienced respiratory physiology consultant. In 22% of the 1,298 records which were over-read, the session quality was changed to a lower grade, while in 24% of cases, the session quality grade was improved; overall 9% of grades that had been recorded as acceptable by the programme (A-C) were rejected, and 12% of cases that had not been acceptable (D) were re-graded to be usable. Variables included in the dataset are those after the quality control procedure, i.e. showing the amended grade where appropriate.

4.5.4 Melatonin data

It was necessary to re-code some values which fell outside of the range which the laboratory was able to measure. Those values which were <1.5 ng/ml have been re-coded as 1.4 ng/ml, and those which were >25 ng/ml to 26 ng/ml.

5. Weighting variables

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 2010 core general population sample data. (For more detailed information on how the weights were produced see Health Survey for England 2010: Volume 2: Methods and documentation).

A household weight has been generated for the general population sample which adjusts for non-contact and refusal of households, this is described in more detail in section 5.1 Individual level non-response weights have also been generated for the general population and are described in section 5.2 onwards.

The individual weights adjust for the additional non-response among individuals in participating households and additional weights take into account respondents' participation in different elements of the survey: interview, nurse visit, blood sample, cotinine (from the saliva sample), urine sample and spirometry. A weight to correct for seasonal effects for

Vitamin D has also been generated.

5.1 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 the selection of a single household at addresses with more than one.) Note that the population control totals used for the calibration weighting were the ONS projected mid-year population estimates for 2009, but with a small adjustment to exclude the population aged 65 and over living in institutions, based on data from the 2001 census.

5.2 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 both the Core and the Boost sample, the ***wt_child*** variable can be used. For time series analysis of children aged 2-15 in the Boost only sample the ***wt_chsel*** variable can be used.

5.3 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.

5.4 Blood weight

A blood weight has been generated for all adults who had a nurse visit, were eligible for, agreed and were able to give a blood sample. This weight (***wt_blood***) should be used on all analysis of questions asked relating to blood samples.

5.5 Cotinine weight

A cotinine weight (from the saliva sample) has been generated for all adults and children that were aged 4-15yrs who had a nurse visit and were eligible for a saliva sample. This weight (***wt_cotinine***) should be used on all analysis of questions asked relating to saliva samples.

5.6 Urine weight

A urine weight has been generated for all adults that had a nurse visit and were eligible to have a sample of urine taken. This weight (**wt_urine**) should be used on all analysis of questions asked relating to urine samples (including Melatonin analysis).

5.7 Spirometry weight

A spirometry weight has been generated for all adults, and for children aged 7-15, that were eligible to have a sample of urine taken. This weight (**wt_spirometry**) should be used on all analysis of spirometry data.

5.8 Vitamin D weight

A weight has been generated for Vitamin D, which adjusts for the season in which the interview/nurse visit took place. This weight (**vitd10wt**) should be used when a correction for seasonal effects is required for analysis of Vitamin D.

6. HSE 2010 Report

Further information about the Health Survey for England 2010 is available in the following publications:

- Craig R and Mindell J (eds). Health Survey for England 2010: Volume 1. Respiratory health. The NHS Information Centre, Leeds, 2011.

- Craig R and Mindell J (eds). Health Survey for England 2010: Health and lifestyles. Volume 2. Methods and documentation. The NHS Information Centre, Leeds, 2011.

- Craig R and Mindell J (eds). Health Survey for England 2010: Respiratory Health. Summary of key findings. The NHS Information Centre, Leeds, 2011.

These can be found at www.ic.nhs.uk/pubs/HSE10report

Trend tables for the Health Survey for England series can be found at www.ic.nhs.uk/pubs/HSE10trends :

- Adult Trend Tables 2010. The NHS Information Centre, Leeds, 2011.

- Child Trend Tables 2010. The NHS Information Centre, Leeds, 2011.

- Health Survey for England 2010: Trend tables (commentary). The NHS Information Centre, Leeds, 2011.

- Population Number Estimate Tables 2010. The NHS Information Centre, Leeds, 2011.

- Population Number Estimates - user guide. The NHS Information Centre, Leeds, 2011.

Further information about the Health Survey for England in general can be found on the respective websites of The NHS Information Centre, NatCen Social Research and University College London:

http://www.ic.nhs.uk/statistics-and-data-collections/health-and-lifestyles-related-surveys/health-survey-for-england

http://www.natcen.ac.uk/study/health-survey-for-england-2010

http://www.ucl.ac.uk/hssrg/hse.html

7. APPENDIX A

HEALTH SURVEY FOR ENGLAND 2010 – CONTENTS

Household data

Household size, composition and relationships	Household income
Accommodation tenure and number of bedrooms	Type of dwelling and area
Mould and damp, cooking appliances, pets	Smoking in household
Economic status/occupation of household reference person	Car ownership

Individual level information

	Age							
	0-1	2-3	4	5-7	8-10	11-12	13-15	16+
Interviewer visit								
General health, longstanding illness, limiting longstanding illness, acute sickness	●	●	●	●	●	●	●	●
Personal care plans								●
Respiratory health	●	●	●	●	●	●	●	●
Doctor-diagnosed hypertension and diabetes								●
Kidney disease								●
Dental health								●
Fruit and vegetable consumption				●	●	●	●	●
Smoking					● [a]	● [a]	● [a]	● [b]
Drinking (seven day period)					● [a]	● [a]	● [a]	● [b]
Economic status/occupation, educational achievement								●
Ethnic origin	●	●	●	●	●	●	●	●
Height measurement		●	●	●	●	●	●	●
Weight measurement	●	●	●	●	●	●	●	●
Reported birth weight	●	●	●	●	●	●	●	
Consent to linkage to NHS Central Register/Hospital Episodes Statistics								●
Self completion								
Warwick Edinburgh mental wellbeing scale								● [a]
GHQ12							● [a]	● [a]
EQ5D								● [a]
Contraception, sexual health								● [a,c]
Perception of weight					● [a]	● [a]	● [a]	
Strengths and difficulties			● [d]	● [d]	● [d]	● [d]	● [d]	
Nurse visit								
Immunisations	●							
Prescribed medicines and vitamin supplements	●	●	●	●	●	●	●	●
Nicotine replacements								●
Waist and hip circumference						●	●	●
Blood pressure					●	●	●	●
Lung function				● [e]	●	●	●	●
Saliva sample (cotinine)			●	●	●	●	●	●
Blood sample								●
Urine sample								●

[a] This module was administered by self-completion.
[b] This module was administered by self-completion for those aged 16-17 and some aged 18-24.
[c] This module was administered to adults aged 16-69.
[d] This questionnaire was administered by self-completion to parents of children aged 4-15.
[e] Lung function was measured in children aged 7 and over.