

# Welsh Health Survey

**2003-04**

## User Guide

A survey carried out on behalf of the Welsh Assembly Government  
by the National Centre for Social Research

# 1. Background

The data files contain data from the Welsh Health Survey (WHS) 2003-04, the first year of a new survey commissioned by the Welsh Assembly Government (WAG) and carried out by a consortium comprising the National Centre for Social Research (NatCen), Beaufort Research and the Department of Epidemiology and Public Health at UCL. The WHS replaces two studies previously carried out: the Welsh Health Survey which was conducted in 1995 and 1998, and the Health in Wales Survey which was conducted five times in 1985, 1988, 1990, 1993 and 1996.

The main aims of the Welsh Health Survey are to:

- Provide estimates of health status, health determinants and health service use;
- Contribute to setting and monitoring targets and indicators in the health strategies and National Service Frameworks;
- Examine differences between population sub-groups (such as sex, age, social class) and local areas (health boards and local authorities);
- Provide a direct measurement of need for health care for NHS resource allocation in Wales;
- Provide local health board and local authority level information for the development of joint local health, social care and well-being strategies.

# 2. Survey Design

The first two years of the Welsh Health Survey were designed to achieve a total sample of 30,000 adults over two fieldwork years, 2003-04 and 2004-05 (WHS data for 2004-05 is archived separately). For the two year period from October 2003 to September 2005, 30,240 addresses were randomly sampled from the small user version of the Postcode Address File (PAF), stratified by unitary authority. In Year 1 15,158 addresses were issued, resulting in 10,217 productive households.

Where selected addresses were found to contain more than one household, up to three were included in the WHS. If more than three households were found, interviewers were instructed to select three at random for the WHS. The survey consisted of a short household interview (administered by an interviewer) followed by a self-completion questionnaire, for all members of the household, which was collected by the interviewer. The adult self-completion questionnaire was 18 pages long. The questionnaire for children (aged under 16 years), which was completed by a parent or guardian, was 3 pages long. All survey documents were translated into Welsh, and bilingual interviewers were used where required.

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

# 3. Documentation

The documentation has been organised into two sections, besides this user guide:

- Interview (contains the household questionnaire, show cards and self-completion questionnaires, English versions only)
- Data (contains the list of variables and derived variable specification).

## 4. Using the data

The 2003-04 data consists of two individual level files – one for adults, and one for children:

WH 03_04 adults_archive.sav	16,298 records	contains data for all adults in co-operating households who returned a self-completion booklet. It contains information from the household questionnaire and self-completion.
WH 03_04 child_archive.sav	4,304 records	contains data for all children in co-operating households for whom a self-completion booklet was returned. It contains information from the household questionnaire and self-completion.

### 4.1 Variables on the files

Each of the data files contains 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 Health Service Use, Accidents, Alcohol, Fruit and Vegetables)
- Full list of all variables included in data file (336 in adult file, 47 in child file)
- Source of each variable (eg Household questionnaire, Self-completion, Derived variable)

Once you have decided which variables to include in your analysis, you can look up details of the question wording in the household or self-completion questionnaires, using the interview section documentation, or use the “**Derived Variables Specification**” document in the data section of the documentation for derived variables.

### 4.2 Standard classification lists

The variables nssec8, nssec5 and nssec3 are based on the National Statistics Socio-Economic Classification (NS-SEC). Full details of NS-SEC are in the Office for National Statistics guide ‘The National Statistics Socio-Economic Classification User Manual’<sup>1</sup>.

If appropriate, cases where an ‘other’ answer was given to questions on chronic or long-term illnesses, health problems or disabilities were ‘backcoded’ into the previous pre-coded individual illness questions. This process converted the text at ‘other’ answers on illnesses into ICD groups and chapters which were then matched into the previous illness questions. This process was carried out for both adults and children.

### 4.3 Weighting variables

Weights were calculated for the WHS data to correct for unequal selection probabilities and also for survey non-response. Respondents in the survey did not have equal chances of selection for two reasons: the probability of selecting an address varied by Unitary Authority and at multiple occupancy addresses containing 4 or more households, 3 were selected at random for inclusion in the survey. Weights were also calculated to correct for differences in non-response between groups.

Non-response weights adjusted for non-response at the household and individual level to account for non-contact and refusals of entire households, and for non-response among individuals within responding households. The final weights arrived at are applied at the individual level for both adults and children (**int\_wt**).

More detailed information on the weighting strategy and how the weights were calculated can be found in the technical report (see references).

---

<sup>1</sup> [www.statistics.gov.uk/nsbase/methods\\_quality/ns\\_sec/default.asp](http://www.statistics.gov.uk/nsbase/methods_quality/ns_sec/default.asp)

#### 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.
- 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 Special licence data sets

The sample for both Years 1 and 2 was drawn using a multi-stage stratified method based around Unitary Authority. For this reason both strata and PSU variables pose a potential risk of disclosiveness and have therefore been removed from these data. However they are available through special licence from the Data Archive [DA please insert link/reference here].



## 5. WHS 2003-04 Report

Further information about the Welsh Health Survey Year 2003-04 including the technical and annual report is available in:

<http://new.wales.gov.uk/topics/statistics/publications/publication-archive/health-survey2003-04/?lang=en>