

Health Survey for England

**The Health of
Ethnic Groups**

'99

User Guide

A survey carried out on behalf of The Department of Health

Joint Health Surveys Unit

Social & Community Planning Research

Department of Epidemiology and Public Health, University College London

1. Background

The Health Survey for England 1999 (HSE) data files contain data from the ninth year of a series of surveys designed to monitor trends in the nation's health. The 1999 Health Survey was commissioned by the Department of Health 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 University College, London.

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;
- to assess the frequency with which combinations of risk factors 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.

2. Survey Design

The 1999 Health Survey for England was designed to provide data at both national and regional level about the population aged 2 and over living in private households in England with a specific focus on the health of minority ethnic groups. The sample design of the 1999 survey had two parts: a general population sample that followed the same pattern as in previous years and a minority ethnic 'boost' sample, designed solely to yield additional interviews with members of the most populous six minority ethnic groups: Black Caribbean, Indian, Pakistani, Bangladeshi, Chinese and Irish.

The general population sample was about half the size of previous years, and involved selecting about 6,500 addresses from the Postcode Address File (PAF) in 312 postal sectors that were issued over a 12 month period. All adults in the selected households were surveyed, as were all children if there were no more than two in the household. If there were more than two children, two were randomly selected for inclusion.

In the ethnic boost sample, about 26,500 addresses were randomly selected from PAF, within another 340 postal sectors, which were issued over a 10 month period (January to October 1999). About 37,600 adjacent addresses were also covered by focused enumeration. All the sampled addresses were fully screened and only informants from the specified minority ethnic groups were eligible for inclusion in the survey. Among eligible informants at an address, a maximum of four adults and three children were selected for interview, so that a random selection procedure was used at addresses which contained more than this number of adults and/or children. (This methodology was not used for boosting the number of Chinese informants, who were followed up from an earlier survey in 1998 which had looked at the health of the Chinese in England).

For informants from the specified minority ethnic groups (whether in the general population or the ethnic boost sample), the coverage of the 1999 survey was similar to that for 1998. An interview with each eligible person (Stage 1) was followed by a visit by a nurse (Stage 2), who made a number of measurements and requested permission to obtain a sample of blood from those aged 11 and over. Saliva samples were also collected from those aged 4 and over. Blood and saliva samples were sent to a laboratory for analysis.

Informants in the general population sample, unless they were members of the specified minority ethnic groups, were given a short version of the questionnaire covering only the core topics. They did not have the follow-up nurse visit, except for a small sub-sample selected to provide comparisons with minority ethnic groups in respect of measurements not previously employed on the Health Survey.

Information about children aged 2-12 was obtained from a parent, with the child present.

Computer-assisted interviewing was used by both interviewers and nurses.

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 1999 data consists of two individual level files and one household level file:

HSE99gp.sav	9,640 records	contains data for all individuals aged 2+ in the <i>General Population Sample</i> in co-operating 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).
HSE99eth.sav	10,259 records	contains data for all individuals aged 2+ in the <i>Ethnic Boost Sample and respondents in the General Population Sample who were of the specified ethnic groups</i> in co-operating 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).

N.B. Respondents in the general population who were of the specified ethnic groups (Black Caribbean, Indian, Pakistani, Bangladeshi, Chinese and Irish) will be included in both of the above datasets.

HSE99ah.sav	22,140 records	contains data on household composition, and sex, age and marital status for all individuals in co-operating households. Provided as an aid to household level analysis. Other household level variables are stored on the individual file.
-------------	----------------	--

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

General Population Data (HSE99gp.sav)

There is **no** weighted variable for general population **adult** data.

For **children** in the general population aged 2-15, the weighted variable **ErrorWt** should be used.

Minority Ethnic Group Data (HSE99eth.sav)

ErrorWt should be used to weight the ethnic group data to do any analysis. It takes account of the different selection probabilities within each ethnic group. The weighted base within each ethnic group is the same as the unweighted base so that standard errors are calculated correctly.

ScaleWt is a scaled version of ErrorWt which ensures that the different ethnic groups have weighted bases that are in proportion to the prevalence of each ethnic group in the population. ScaleWt should not be used for standard errors as they would be calculated using the weighted base.

4.3 Multicoded questions

Multicoded questions are stored in the archived HSE 1999 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 1999 Health Survey.

As an example, question CONSUBX on the 1999 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 1999, the variables CONSUBX1-5 store the answer to this question by category as follows:

- CONSUBX1 - coded 1 for those who ate in the last half hour and 0 for those that didn't.
- CONSUBX2 - coded 1 for those who smoked in the last half hour and 0 for those that didn't.
- CONSUBX3 - coded 1 for those who drank alcohol in the last half hour and 0 for those that didn't.
- CONSUBX4 - coded 1 for those who did vigorous exercise in the last half hour and 0 for those that didn't.
- CONSUBX5 - 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 CONSUBX5 precludes having a code 1 at any of the variables CONSUBX1 - CONSUBX4). 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.
- 6 Schedule not obtained: Used to signify that a particular variable was not answered because the respondent did not complete or agree to a particular schedule (i.e. nurse schedule or self-completions).

- 7 Refused/ not obtained: Used only for variables on the nurse schedules, this code indicates that a respondent refused a particular measurement or test or the measurement was attempted but not obtained or not attempted.
- 8 Don't know, Can't say.
- 9 No answer/ Refused

These conventions have also been applied to most of the derived variables. Those variables created in earlier years of the Health Survey and used again in 1999 do not on the whole conform to this scheme. The derived variable specifications should be consulted for details.

4.5 Valid cases

In the 1999 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 or eaten within 30 minutes of having their blood pressure taken were excluded from analysis as this can affect blood pressure. The List of Variables document gives details of which variables show only valid codes, and which also include invalid answers.

5. HSE 1999 Report

Further information about the Health Survey for England 1999 is available in:

“Health Survey for England: The Health of Minority Ethnic Groups 1999”. Volume 1: Findings & Volume 2: Methodology and Documentation. Edited by B. Erens, P. Primatesta & G. Prior. HMSO. London. 2000. ISBN 011 322448 6.

or on the Department of Health website at <http://www.doh.gov.uk/public/summary1.htm>

N.B. Additional blood results have been processed since the publication of the Health Survey for England 1999 report (Erens,B., Primatesta, P., and Prior, G. (eds.) (2000) *Health Survey for England 1999: The Health of Minority Ethnic Groups*) which means that the number of blood results on the dataset will be greater than the number reported on in the publication.

APPENDIX A

1999 HEALTH SURVEY FOR ENGLAND - CONTENTS

Household level data

Household size and composition
 Ethnic/Irish screening
 Smoking in household
 Tenure and number of bedrooms

Car and telephone ownership
 State benefits and household income
 Economic status and occupation of HoH
 Type of dwelling and area

<i>Individual level data</i>	Age									
	2-3	4	5-6	7	8-10	11-12	13-15	16-34	35-65	65+
Interviewer visit (All)										
General health, longstanding illness, acute sickness	●	●	●	●	●	●	●	●	●	●
Use of health services: GP, hospital, dental services	●	●	●	●	●	●	●	●	●	●
Accidents	●	●	●	●	●	●	●	●	●	●
Adults' Physical Activity								●	●	●
Smoking								● ^a	●	●
Drinking								● ^a	●	●
Economic status/occupation								●	●	●
Educational attainment								●	●	●
Ethnic origin, country of birth, religion, cultural identity	●	●	●	●	●	●	●	●	●	●
Reported birth weight	●	●	●	●	●	●	●			
Height/weight measurements	●	●	●	●	●	●	●	●	●	●
Additional interviewer visit topics for eligible ethnic groups										
Chest pain (Rose Angina)								●	●	●
CVD, including use of services								●	●	●
Children's Physical Activity	●	●	●	●	●	●	●			
Asthma/Wheezing, including use of services	●	●	●	●	●	●	●			
Self completion (All)										
GHQ12 (psycho-social health)								●	●	●
Social support								●	●	●
Drinking/smoking					●	●	●	● ^a		
HRT								●	●	●
Contraceptive use								●	●	●
Nurse visit (All eligible ethnic groups and all aged 35+ in nurse areas)										
Prescribed Medicines	●	●	●	●	●	●	●	●	●	●
MUAC	●	●	●	●	●	●	●			
Demispan										●
Waist/hip								●	●	●
Blood pressure			●	●	●	●	●	●	●	●
Lung function				●	●	●	●			
Blood sample						●	●	●	●	●
Saliva sample		●	●	●	●	●	●	●	●	●
Fasting blood									●	●
ECG									●	●
Eating habits self completion								●	●	●
SDQ booklet self completion		●	●	●	●	●	●			
a Administered by self-completion for age 16, 17 and optionally by self-completion for 18-24										