



Health Survey for England

2012

**Trend tables:
Population number
estimates**

A survey carried out on behalf of the Health and
Social Care Information Centre

Joint Health Surveys Unit

NatCen
Social Research that works for society



Department of Epidemiology and Public Health,
University College London

Health Survey for England – 2012 trend tables: Population number estimates User Guide

Tables	Contents	Years available
Adults		
1	Adults: number estimates for body mass index (BMI) categories, by survey year, age and sex	2003-2012
2	Adults: number estimates for self-reported cigarette smoking status, by survey year, age and sex	2003-2012
3	Adults: number estimates for alcohol consumption on heaviest drinking day in the last week, by survey year, age and sex	2003-2012
4	Adults: number estimates for fruit and vegetable consumption, by survey year, age and sex	2003-2011
5	Adults: number estimates for levels of physical activity, by survey year, age and sex	2003, 2004, 2006, 2008, 2012
Children		
6	Children: number estimates for overweight and obesity prevalence, by survey year and sex	2003-2012
7	Children: number estimates for fruit and vegetable consumption, by survey year, age and sex	2003-2011
8	Children: number estimates for levels of physical activity, by survey year, age and sex (new method and 2011 guidelines)	2008, 2012
Bases		
9	ONS mid-year population estimates for adults aged 16 and over living in private households	2003-2012
10	ONS mid-year population estimates for children aged 2-15	2003-2012

Introduction

Background

The Health Survey for England (HSE) is a series of annual surveys designed to measure health and health-related behaviours in adults and children living in private households in England. The survey is currently commissioned by the Health and Social Care Information Centre (HSCIC), and before April 2005 was commissioned by the Department of Health. Since 1994, the survey has been carried out by NatGen Social Research and the Research Department of Epidemiology and Public Health at UCL (University College London).

The survey consists of an interview and nurse visit. It has a series of core elements that are included every year (or alternate years), and special topics that are included in selected years. Core topics include general health, smoking, drinking, fruit and vegetable consumption, height, weight, blood pressure measurements and blood, urine and saliva samples. The 2011 survey introduced a major new core topic of social care. Special topics include, for example, cardiovascular disease (CVD), physical activity, accidents, lung function measurement and certain blood analytes. In 2012 there was a focus on physical activity, and additional modules of questions were also included, covering average weekly alcohol consumption, gambling, well-being and sexual health.

Each year there is a general population sample in which adults and children in selected households are eligible for inclusion. Adults aged 16 and over have been included since the start of the HSE series, children aged 2-15 were first included in 1995, and infants aged 0-1 have been included since 2000. In some years a boost sample is used to increase the proportion of respondents from certain population groups, such as in 2002 when a boost sample of children and young adults was included, and 2005 when a boost of older people aged 65 and over was included. Boost samples of children have been included in 2005 to 2010. In 2012, there was a general population sample of 8,291 adults and 2,043 children. There was no child boost in 2011 and 2012 and so base sizes may be small for some categories compared with recent years

The trend tables focus on key changes in core topics and measurements.

- The trend tables, updated each year, show survey estimates of *prevalence* for a number of key measures (the percentage of the population overweight, obese, current smokers etc). The 2012 trend tables are available at www.hscic.gov.uk/pubs/hse2012trend.
- The *number estimates* presented here convert the prevalences in the key trend tables into estimates of the numbers of people in the population in England that they represent. As an illustration, the obesity prevalence estimate of 24.8% for all adults aged 16 and over in 2012 has been converted into a number estimate of around 10.8 million (between 10,285,000 and 11,337,000). The tables in this series relate to lifestyle characteristics influencing health: overweight and obesity, fruit and vegetable consumption, smoking, drinking and physical activity. The number estimates cover 2003-2012. Estimates for adult physical activity, however, are only available for 2003, 2004, 2006, 2008 and 2012, and child physical activity for 2008 and 2012; results for the latter two years present a new method of estimating physical activity, and are based on the new guidelines introduced in 2011. Fruit and vegetable consumption was not covered in 2012.

The prevalence trend tables present the results from the representative general population sample, and in some years boost sample data are also included to increase the precision of sub-group estimates (e.g. young adults in 2002, persons aged 65 and over in 2005).

For the number estimates, the prevalence estimates for each year (expressed as a proportion) were multiplied by a scaling factor equal to the total mid-year population estimate published by the Office for National Statistics (ONS) multiplied by the estimated proportion of people in the relevant age-sex group in the HSE. The mid-year population estimate published by ONS¹ was adjusted to represent the population living in private households, excluding those aged 65 years and over living in institutions (43.7 million adults aged 16 and over in 2012).

Details of the method used to compute the number estimates and accompanying margin of error (i.e. the width of the 95% confidence interval divided by two) are provided in the technical annex.

How to use these tables

The eight tables presented here show estimates of how many people in England, living in private households, have particular characteristics or behaviour. For instance the tables show the number of men and women who are estimated to be obese, who eat the recommended five or more portions of fruit and vegetables a day, or who currently smoke cigarettes. The tables also show breakdowns by age groups, for instance the number of children aged 11 who meet the government's recommendation of eating five or more portions of fruit and vegetables per day, or the number of 16-24 year olds who drink above recommended levels of alcohol.

Each table for adults shows results separately for men and women, usually within age categories, followed by results for all adults (men and women combined). Similarly tables for children show results for boys and girls within age groups, and then results for all children.

- Numbers in all tables are presented in thousands, so 10,301 in the tables represents an estimate of 10,301,000 people in the population in England.
- As results are based on survey data they are affected by sampling error. The margin of error is shown for each estimate in the tables, again in thousands. For instance '± 389' in the tables represents ± 389,000. The lower and upper limits of the 95% confidence interval can be obtained from the margin of error as follows:

Upper limit = estimate + margin of error

Lower limit = estimate – margin of error

- The confidence interval shows the range within which the true result is likely to fall (at the 95% confidence level). Thus for the estimate of 10,301,000 people and a margin of error of ± 389,000, the true value is expected to lie between 9,912,000 and 10,690,000 in 95 cases out of 100.
- The confidence interval is a measure of the survey's reliability, and is also influenced by the size of the sample. For a large sample such as all men or all women, the confidence interval is narrow. Where the sample size is smaller, for instance when looking at a particular age group, the confidence intervals are larger (reflecting the larger uncertainty in the estimation). Estimates for adults aged 16 and over based on a full-size general population sample of 15,000-16,000 adults (2003, 2006 and 2008) have a narrower confidence interval than in years with smaller samples (2004, 2005, 2007 and 2009-2012).
- The ONS mid-year population estimate bases for 2003-2012 are shown in Tables 9 and 10 for adults and children respectively. It should be noted that there may be slight differences between the sum of estimates in the tables and the bases in the final two tables. (For instance the sum of the % low levels of physical activity, % doing some activity and % meeting recommendations may not sum to the exact population size). The reason for this is the rounding of estimates to the nearest thousand.
- Before 2003, no weighting was applied to the adult sample, whereas from 2003 survey estimates have been weighted for non-response.² The sample of children each year was weighted to adjust for the probabilities of selection,³ and from 2003 non-response weighting was also introduced for children. While the prevalence tables show estimates from 1993 onwards, the number estimates series are published from 2003 onwards. Number estimates before 2003 are not shown because the change in weighting means a different calculation must be used for the confidence intervals, as described in the technical annex.
- Non-response weighting brings the profile of the survey sample very close to the profile

of the total population; however, at the time that weighting is applied the ONS mid-year population estimates for the relevant year are not available, and the previous year's population estimates are used. Estimates for the 2012 *prevalence* trends, therefore, have been computed on a non-response weight based on 2011 mid-year population estimates. In most years the *number estimates* have been computed using a grossing-up factor calculated from the ONS mid-year population estimates for the previous year; for the 2011 number estimates, the grossing-up factor was calculated from the 2011 census.

- The adult population of England living in private households has increased from around 39.7 million in 2003 to 43.7 million in 2012. Therefore an increase in numbers in a particular category from one year to the next is not unexpected. Increases in the numbers should be considered alongside any change in *prevalence* (i.e. the percentage of the population) for that category. For full details on key trends in the health of adults and children since 1993 on selected topics see the commentary for the prevalence trend tables at www.hscic.gov.uk/pubs/hse2012trend.

Technical annex

This section provides a technical description of the methods used to calculate the number estimates and accompanying margins of error. The section concludes with a worked example. The method used is valid only from the introduction of non-response weighting into the HSE series. The number estimates are, therefore, presented for years 2003 onwards.

I Number estimates

For each indicator the data presented in the prevalence trend tables were multiplied by a scaling factor to estimate the number of people with particular characteristics or behaviour (number obese, number of current smokers etc). The scaling factor was based on two figures. The first was the ONS mid-year population estimate for the relevant year, adjusted to represent the population living in private households (i.e. excluding those aged 65 years and over living in institutions). The second figure used was an estimate of the proportion of people in the relevant age-sex group (e.g. percentage of the adult population who were male aged 16-24). The proportions in each age-sex group were calculated from HSE data (with the non-response weight applied). It should be noted that a different estimate of this second figure is calculated for each table as only the respondents to the relevant question are included in the calculation. This means that the age-sex distribution estimated from HSE data does not correspond exactly to the age-sex distribution available from the ONS figures.⁴ The calculation chosen allows the prevalence estimates to be re-created from the number estimate tables.

The weighting variable can be denoted as w_{ij} where i denotes the age-sex group of the HSE respondent (e.g. men aged 16-24) and j denotes the HSE respondents to the question within each age-sex group.

Then the weighted number of respondents to the question can be expressed as:

$$w = \sum_{i,j} w_{ij} \quad [1]$$

and the weighted number of respondents of a particular age-sex group as:

$$w_i = \sum_j w_{ij} \quad [2]$$

Dividing [2] by [1], therefore, gives the estimate of the proportion of the population belonging to a particular age-sex group.

If the overall (i.e. men and women combined) ONS mid-year population estimate is labelled as P , then the HSE estimate of the number of persons belonging to age-sex group i (labelled M_i) can be expressed as:

$$M_i = P \left(\frac{W_i}{W} \right) \quad [3]$$

This is the scaling factor described above: the ONS mid-year population total (P) multiplied by the estimate of the proportion of people in the relevant age-sex group. Note that due to item non-response (e.g. refusals and don't knows) the proportion will be slightly different for each table.

Finally, if the number estimate for age-sex group i and health lifestyle indicator k is labelled N_{ik} , and the corresponding prevalence estimate is labelled ϕ_{ik} , then the number estimate can be derived as follows:

$$N_{ik} = P \frac{W_i}{W} \phi_{ik} \quad [4]$$

In other words, the estimated number of, say, obese men aged 16-24, can be expressed as the overall ONS population estimate P , multiplied by the proportion of respondents with a valid BMI who were male aged 16-24, multiplied by the estimated obesity prevalence for men aged 16-24.

Using this calculation method allows the estimates in the prevalence trend tables to be reproduced from the number estimates, ensuring consistency between the two sets of tables.

Take care to note, therefore, that the only ONS mid-year population estimate that can be reproduced from these figures is the *overall* population figure, P (i.e. total population aged 16 and over in the relevant year). The proportion of people in the relevant age-sex group is taken from the HSE (and not ONS) data.

The implication of this is that the prevalence estimates cannot (and should not) be derived by dividing the number estimate for each age-sex group by the corresponding ONS age-sex population estimate.⁴ For prevalence estimates users are advised to consult the prevalence trend tables.

II Margin of error

The margin of error (MoE) of the number estimates is calculated by multiplying the MoE of the prevalence by M_i , the scaling factor used to create the number estimates. If E_{ik} is the margin of error for a number estimate and ϵ_{ik} is the margin of error of the associated prevalence then:

$$E_{ik} = \frac{P W_i}{W} \epsilon_{ik} \quad [5]$$

The standard error of the prevalence is calculated using STATA's survey module commands which take into account the complex sample design. The MoE is then calculated by multiplying the standard error by 1.96.

In order to use equation 5, it must be assumed that no further uncertainty is added to the number estimate when the prevalence is multiplied by the scaling factor, in other words that the scaling factor $P \frac{W_i}{W}$ is constant. It is assumed that the ONS estimates P have small enough variance to discount.

The same assumption can be made for the factor $\frac{W_i}{W}$. This assumption is justified by considering that $\frac{W_i}{W}$ is the weighted proportion of respondents for a particular age-sex group. The purpose of the weighting is to correct the figures for non-response. The weighting scheme uses several variables including age and sex. The weighting is chosen to make $\frac{W_i}{W}$ approximately equal to a fixed value, taken from the ONS population estimates. That is, although the age-sex distribution of the sample is prone to sampling error, the weighting 'fixes' the values of $\frac{W_i}{W}$ to be approximately equal to the ONS population estimates.

It should be noted that the variance of $\frac{W_i}{W}$ is not exactly zero. Due to the impact of other

variables included in the weighting and differential item non-response⁴ the value of $\frac{W_i}{W}$ is not exactly the same for all possible samples. This means the assumption causes a slight underestimate in the value of the margins of error. However, the variation is small and the assumption allows considerable computational efficiency gains.

Worked example

The method used to convert the prevalence estimates in some of the standard trend tables into estimates of the numbers of people in the population in England that they represent is illustrated here using a worked example. The example used is obesity (including morbidly obese) for men using HSE 2007. Table A1 shows the calculation of the number estimates; Table A2 the calculation of the margin of error.

Table A1						
2007 number estimates (obesity for men)						
Age group	ONS population estimate (P)	Prevalence of obesity (ϕ_{ik})	Weighted number of respondents (w)	Weighted number of respondents of age-sex group (w_i)	HSE number estimate age-sex group (M_i)	Estimated number obese (N_{ik})
	Column 1	2	3	4	5	6
16-24	41,034,273	0.091	5,991	461	3,156,621	287,181
25-34	41,034,273	0.156	5,991	499	3,421,061	534,096
35-44	41,034,273	0.248	5,991	603	4,127,411	1,025,474
45-54	41,034,273	0.345	5,991	514	3,523,963	1,216,872
55-64	41,034,273	0.309	5,991	446	3,057,682	946,352
65-74	41,034,273	0.284	5,991	296	2,024,746	574,557
75+	41,034,273	0.219	5,991	189	1,292,632	283,495

- Column 1 shows the overall ONS population estimate for adults: showing just over 41 million adults aged 16 and over living in private households in England on 30th June 2007.
- Column 2 shows the estimated prevalence of obesity for each of the seven age groups (expressed as a proportion). 21.9% of men aged 75+ were classed as obese, compared with 9.1% of men aged 16-24.
- Column 3 shows the weighted number of HSE respondents with a valid BMI measurement (using the non-response main interview weight, which is scaled to the achieved sample size, resulting in the weight being standardised around an average of one).
- Column 4 shows the weighted number of HSE respondents in each age-sex group. Of all HSE 2007 respondents aged 16 and over (with valid BMI), about 7.7% were men aged 16-24 (column 4/column 3).
- Column 5 shows that applying the proportion of HSE respondents in each age-sex group to the ONS population estimate P gives an HSE estimate of the number of persons in the population in each age-sex group: equation [3] above.

Applying the proportion of HSE respondents who were male aged 16-24 ($461/5,991 = 0.077$) to the ONS population estimate gives an estimated number of 3,156,621 men aged 16-24:

$$HSE\ estimate_{men,16-24} = 41,034,273 \left(\frac{461}{5,991} \right) = 3,156,621$$

- Column 6 shows the number estimate of men in England who were classed as obese in 2007: which can be calculated using equation [4].

As an illustration, for men aged 16-24, the estimate of the number who were obese can be calculated as follows:

$$obese_{men,16-24} = 41,034,273 \left(\frac{461}{5,991} \right) 0.0910 = 287,181 \text{ (rounded to 287,000)}$$

The number in the population estimated to be obese can be expressed as the prevalence (0.091 for men aged 16-24) multiplied by a scaling factor equal to the ONS mid-year population estimate (41 million) multiplied by the HSE estimate of the proportion of people in the male 16-24 category ($461/5,991 = 0.077$).

Note, therefore, that the HSE estimate of the proportion of the population in England in each age-sex group does not match exactly the equivalent ONS mid-year estimate (see endnote 4). The only ONS mid-year population estimate that can be reproduced is the *overall* population figure.

Table A2					
Estimated margin of error					
Age group	Prevalence of obesity (Φ_{ik})	Estimated standard error (SE)	HSE number estimate age-sex group (M _i)	Standard error of estimated number obese	Margin of error (unrounded)
	Column 1	2	3	4	5
16-24	0.091	0.017	3,156,621	53,756	105,362
25-34	0.156	0.019	3,421,061	66,225	129,801
35-44	0.248	0.019	4,127,411	80,382	157,549
45-54	0.345	0.022	3,523,963	78,154	153,181
55-64	0.309	0.022	3,057,682	68,246	133,763
65-74	0.284	0.023	2,024,746	45,951	90,064
75+	0.219	0.026	1,292,632	32,999	64,677

- Columns 1 and 2 show the estimated prevalence of obesity and accompanying standard error (SE) respectively (both expressed as a proportion).
- Column 3 shows the HSE estimate of the number of residents in England in each age-sex group (described above).
- Column 4 shows the SE of the estimated number that were obese. It is calculated by multiplying the SE of the prevalence by the estimated number of people in the relevant age-sex group (column 2 × column 3).

For men aged 16-24, the estimated SE of the number who were obese can be calculated as follows:

$$SE \text{ number obese}_{men,16-24} = 0.017 \times 3,156,621 = 53,756$$

- Column 5 gives the margin of error (MoE): the estimated SE of the number estimate multiplied by 1.96.

For men aged 16-24 the MoE is as follows:

$$MoE \text{ number obese}_{men,16-24} = 53,756 \times 1.96 = 105,362 \text{ (rounded to } \pm 105,000 \text{)}.$$

References and notes

- 1 ONS mid-year population estimates are available at www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-england-and-wales/mid-2011--2011-census-based-/stb---mid-2011-census-based-population-estimates-for-england-and-wales.html

- 2 Sproston K and Primatesta P (eds). *Health Survey for England 2003: Volume 3 Methodology and documentation*. The Stationery Office, London, 2004.
http://webarchive.nationalarchives.gov.uk/+www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsStatistics/DH_4098712
- 3 Only up to two children per household are interviewed, and in households with three or more children, two are selected at random. Therefore weighting is required to ensure that children in larger households are not under-represented in the sample. The non-response weighted estimates from 2003 onwards include the necessary adjustment for child selection.
http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4098912.pdf
- 4 The weighted age-sex distribution of HSE respondents does not exactly match the corresponding distribution available from the ONS mid-year population estimates. There are two reasons for this. First, at the time that weighting is applied the mid-year population estimates for the relevant year are not available, and the previous year's population estimates are used. Second, the healthy lifestyle indicators are subject to item non-response. For example, the BMI categories are based on respondents with a valid height and weight measurement. There is some non-response among interviewed respondents to height and weight measurements, and the weighting applied to interview data does *not* correct for this element of non-response. The age profile for the achieved samples in 2003, 2006 and 2007 are shown below, compared with the ONS population estimates for 2007 (adjusted for institutional populations), shown in the shaded column. It can be seen that those in the youngest and oldest categories tend to be slightly under-represented, and this should be borne in mind in interpreting the data.

ONS and HSE age profiles

Age group	ONS	HSE		
	2007 %	2003 %	2006 %	2007 %
16-24	15.00	15.46	14.86	14.45
25-34	16.25	16.79	15.99	16.12
35-44	18.99	19.40	20.30	20.05
45-54	16.17	16.18	16.45	16.88
55-64	14.68	14.67	14.91	15.00
65-74	10.11	9.92	10.22	10.15
75+	8.79	7.58	7.28	7.35

NatCen Social Research

www.natcen.ac.uk

NatCen Social Research is the largest independent social research institute in Britain, carrying out research that works for society. NatCen specialises in research in public policy fields such as health and well-being, society and social change, children and young people, income and work, crime and justice. We offer the full range of quantitative and qualitative research services. Our team includes survey methodologists, data analysts and policy sector specialists. As well as research staff, NatCen has a national panel of over 1,000 interviewers and 150 nurses who work on health-related surveys.

Research Department of Epidemiology and Public Health, UCL (University College London)

www.ucl.ac.uk/epidemiology

The Research Department of Epidemiology and Public Health, chaired by Professor Richard Watt, is a leading centre for research into the social determinants of health, and has a strong interdisciplinary structure. The Department houses 180 staff in 11 main research groups, including the Joint Health Surveys Unit, part of the Health and Social Surveys Research Group (HSSRG). The HSSRG studies population health (including health behaviours and treatments) and inequalities in health. Much of the group's research is carried out using large population surveys that collect data on health, economic and social issues, using a variety of survey methods and statistical techniques, while qualitative methods are also used by the group. The group is multidisciplinary, with epidemiology, sociology, statistics, public health nutrition, demography and geography all represented.

The **Joint Health Surveys Unit** has been created by NatCen Social Research and the Health and Social Surveys Research Group within the Research Department of Epidemiology and Public Health at UCL. The JHSU enables collaborative working, combining the strengths and talents of each organisation, to carry out major health surveys such as the Health Survey for England.



Health & Social Care
Information Centre