Health Survey for England 2015
Trend Tables – Population Number Estimates User Guide

Published 14th December 2016
This is a National Statistics publication

National Statistics status means that official statistics meet the highest standards of trustworthiness, quality and public value.

All official statistics should comply with all aspects of the Code of Practice for Official Statistics. They are awarded National Statistics status following an assessment by the Authority’s regulatory arm. The Authority considers whether the statistics meet the highest standards of Code compliance, including the value they add to public decisions and debate.

It is NHS Digital’s responsibility to maintain compliance with the standards expected of National Statistics. If we become concerned about whether these statistics are still meeting the appropriate standards, we will discuss any concerns with the Authority promptly. National Statistics status can be removed at any point when the highest standards are not maintained, and reinstated when standards are restored.


These tables and user guide may be of interest to people working in public health, policy officials, commissioners of health and care services and to the public to see estimates of the numbers of people in the population for lifestyle characteristics influencing health such as body mass index (BMI), obesity, smoking, alcohol and fruit and vegetable consumption and physical activity.
## List of tables

<table>
<thead>
<tr>
<th>Adults</th>
<th>Years available</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Adults: number estimates for body mass index (BMI) categories, by survey year, age and sex</td>
</tr>
<tr>
<td>2</td>
<td>Adults: number estimates for self-reported cigarette smoking status, by survey year, age and sex</td>
</tr>
<tr>
<td>3</td>
<td>Adults: number estimates for maximum alcohol consumption on any day in the last week, by survey year, age and sex</td>
</tr>
<tr>
<td>4</td>
<td>Adults: number estimates for fruit and vegetable consumption, by survey year, age and sex</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Children</th>
<th>Years available</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Children: number estimates for overweight and obesity prevalence, by survey year and sex</td>
</tr>
<tr>
<td>7</td>
<td>Children: number estimates for fruit and vegetable consumption, by survey year, age and sex</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bases</th>
<th>Years available</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>ONS mid-year population estimates for adults aged 16 and over living in private households</td>
</tr>
<tr>
<td>10</td>
<td>ONS mid-year population estimates for children aged 2-15</td>
</tr>
</tbody>
</table>

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 commissioned by NHS Digital (formerly 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 NatCen Social Research and the Research Department of Epidemiology and Public Health at UCL.

The survey consists of an interview and nurse visit. There is 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:

**Each year**
- General health
- Social care
- Smoking
- Drinking
- Height measurements
- Weight measurements
- Blood pressure measurements
- Adult blood samples
- Child saliva samples

**Alternate years**
- Fruit and vegetable consumption
- General Health Questionnaire (GHQ-12)
- Overall health (EQ-5D)
- Adult urine samples
- Adult saliva samples

Special topics covered in selected years include, for example, cardiovascular disease, physical activity, accidents, lung function measurement and certain blood analytes.

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 were included in 2005 to 2010, and in 2015.

In 2015, interviews were completed with 8,034 adults and 5,714 children.

Reports on HSE data

A report is published on a selection of findings for each survey year, including topic reports covering survey results, and a volume giving full details of the survey methods and documentation. The report on HSE 2015 is available at [http://digital.nhs.uk/pubs/hse2015](http://digital.nhs.uk/pubs/hse2015)

Trend tables, updated each year, show survey estimates of prevalence for a number of key measures (e.g. the percentage of the population who are overweight, obese, current smokers). The 2015 trend tables are available at [http://digital.nhs.uk/pubs/hse2015trend](http://digital.nhs.uk/pubs/hse2015trend). The trend tables focus on key changes in core topics and measurements, and a brief commentary is provided.
The population number estimates tables show estimates of the numbers of people in the population for key variables such as body mass index (BMI). These are available at [http://digital.nhs.uk/pubs/hse2015trend](http://digital.nhs.uk/pubs/hse2015trend). This user guide introduces the tables, and explains how they are produced.

**Population number estimates**

The number estimates presented here convert the prevalence 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 26.9% for all adults aged 16 and over in 2015 has been converted into a number estimate of around 11.8 million (between 11,293,000 and 12,361,000). The tables in this series relate to lifestyle characteristics influencing health:

- overweight and obesity
- fruit and vegetable consumption
- smoking
- drinking
- physical activity.


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, children in 2015).

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, then multiplied by the estimated proportion of people in the relevant age-sex group in the HSE. The mid-year population estimate was adjusted to represent the population living in private households, excluding those aged 65 years and over living in institutions (44.0 million adults aged 16 and over in 2015).

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.

---

1 Mid-year population estimates are published by the Office for National Statistics (ONS), and can be found at [http://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2015](http://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2015)
How to use these tables

The eight tables 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 8-10 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 2009-2015).
- The ONS mid-year population estimate bases for 2003-2015 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.¹ The sample of children each year was weighted to adjust for the probabilities of selection,² 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 2015 prevalence trends, therefore, have been computed on a non-response weight based on 2014 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 2012 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 44.0 million in 2015. 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 http://digital.nhs.uk/pubs/hse2015trend
References


2 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.
Technical annex

Introduction

This section provides a technical description of the methods used to calculate the number estimates and accompanying margins of error. The section is followed by 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.

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. 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:

\[
W = \sum_{i,j} w_{ij}
\]  
[1]

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

\[
W_i = \sum_j w_{ij}
\]  
[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:

\[
M_i = P \left( \frac{W_i}{W} \right)
\]  
[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 $\phi_{ik}$, 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.

It should be noted, 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. For prevalence estimates users are advised to consult the prevalence trend tables.

### 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 MoE for a number estimate and $\epsilon_{ik}$ is the MoE of the associated prevalence then:

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

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-response 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 MoE. However, the variation is small and the assumption allows considerable computational efficiency gains.

---

3 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 sample with a valid BMI measurement in 2015 is shown below, compared with the ONS population estimates for 2015 (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

<table>
<thead>
<tr>
<th>Age group</th>
<th>ONS 2015%</th>
<th>HSE 2015%</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-24</td>
<td>14.06</td>
<td>13.92</td>
</tr>
<tr>
<td>25-34</td>
<td>17.00</td>
<td>17.15</td>
</tr>
<tr>
<td>35-44</td>
<td>16.14</td>
<td>16.44</td>
</tr>
<tr>
<td>45-54</td>
<td>17.49</td>
<td>17.83</td>
</tr>
<tr>
<td>55-64</td>
<td>14.04</td>
<td>14.24</td>
</tr>
<tr>
<td>65-74</td>
<td>11.90</td>
<td>12.09</td>
</tr>
<tr>
<td>75+</td>
<td>9.37</td>
<td>8.33</td>
</tr>
</tbody>
</table>
Worked example

This worked example illustrates 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. The example used is obesity (including morbidly obese) for men using HSE 2007. Table 1 shows the calculation of the number estimates; Table 2 the calculation of the margin of error.

Table 1  2007 number estimates (obesity for men)

<table>
<thead>
<tr>
<th>Age group</th>
<th>ONS population estimate P</th>
<th>Prevalence of obesity $\phi_{ik}$</th>
<th>Weighted number of respondents $w$</th>
<th>Weighted number of respondents of age-sex group $w_i$</th>
<th>HSE number estimate age-sex group $M_i$</th>
<th>Estimated number obese $N_{ik}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-24</td>
<td>41,034,273</td>
<td>0.091</td>
<td>5,991</td>
<td>461</td>
<td>3,156,621</td>
<td>287,181</td>
</tr>
<tr>
<td>25-34</td>
<td>41,034,273</td>
<td>0.156</td>
<td>5,991</td>
<td>499</td>
<td>3,421,061</td>
<td>534,096</td>
</tr>
<tr>
<td>35-44</td>
<td>41,034,273</td>
<td>0.248</td>
<td>5,991</td>
<td>603</td>
<td>4,127,411</td>
<td>1,025,474</td>
</tr>
<tr>
<td>45-54</td>
<td>41,034,273</td>
<td>0.345</td>
<td>5,991</td>
<td>514</td>
<td>3,523,963</td>
<td>1,216,872</td>
</tr>
<tr>
<td>55-64</td>
<td>41,034,273</td>
<td>0.309</td>
<td>5,991</td>
<td>446</td>
<td>3,057,682</td>
<td>946,352</td>
</tr>
<tr>
<td>65-74</td>
<td>41,034,273</td>
<td>0.284</td>
<td>5,991</td>
<td>296</td>
<td>2,024,746</td>
<td>574,557</td>
</tr>
<tr>
<td>75+</td>
<td>41,034,273</td>
<td>0.219</td>
<td>5,991</td>
<td>189</td>
<td>1,292,632</td>
<td>283,495</td>
</tr>
</tbody>
</table>

- 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:

\[ HSE_{\text{estimate}_{\text{men,16-24}}} = 41,034,273 \left( \frac{461}{5,991} \right) = 3,156,621 \]

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:

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

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 footnote 3). The only ONS mid-year population estimate that can be reproduced is the overall population figure.

### Table 2: Estimated margin of error

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

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:

\[ SE_{\text{number obese}}_{\text{men,16–24}} = 0.017 \times 3,156,621 = 53,756 \]

• 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:

\[ MoE_{\text{number obese}}_{\text{men,16–24}} = 53,756 \times 1.96 = 105,362 \quad (rounded \ to \ \pm 105,000) \]