Welsh Health Survey Quality Report

This paper details how the Welsh Health Survey adheres to the six dimensions of the European Statistical System definition of quality.

The Welsh Health Survey provides unique information about the health and health-related lifestyles of people living in Wales. It presents a picture of the health of the Welsh population, variations between sub-groups and areas, and changes over time, and makes an important contribution to informing and monitoring public health strategy in Wales.

1. Relevance

The degree to which the statistical product meets user needs for both coverage and content.

Background

The Welsh Health Survey (WHS) covers a range of health-related issues, including health status, lifestyle and health behaviours, and health service use. The survey was established in 2003 and ran all year round and ended in 2015. Results are published annually.

An achieved sample of around 15,000 adults and 3,000 children is aimed for per year, to include a minimum of 600 adults from each local authority area.

Survey design

The survey is based on a representative sample of people living in private households in Wales, selected using a random sample from the Post Office's Postcode Address File – an up to date list of all addresses maintained by the UK Post Office.

The sample is stratified by local authority. The smaller authorities are oversampled to allow the production of survey estimates at this level following the collection of two years of data. The survey collects information on households (through a short interview) and on individuals (through a self-completion questionnaire). At each household, all adults and a maximum of two children are eligible for inclusion in the survey.

Interpretation of the data

As the survey relies on a self-completion questionnaire, the results of the survey reflect people's own understanding of their health rather than a clinical assessment of their medical condition, and their own interpretation of the health services they have used. Interpretation of the results should take account of the questionnaire design, as the mode of collection and the questions themselves affect the information collected. The small proportion of people not covered by the Postcode Address File, including those living in institutions, were not covered by the survey. It should be noted that people in institutions are likely to be, on average, in poorer health than those in private households – this should be kept in mind when considering the survey results.

Weighting

The survey results are weighted[1] to take account of unequal selection probabilities and for differential non-response i.e. to ensure that the age and sex distribution of the responding sample matches that of the population of Wales.

Who are the key users?

The survey is used by a wide range of users, including those working in national and local government, NHS organisations, research and academic settings. It is also useful to a wider general audience, particularly those in Wales.

We encourage users of the statistics to contact us to let us know how they use the data. A list of key users of the survey is maintained, who regularly receive updates on survey developments and outputs.

If you are a user and would like to be added to our circulation list, please let us know by e-mailing stats.healthinfo@wales.gsi.gov.uk.

What is the data used for?

The survey meets a range of needs and is used in many ways, including to:

- provide national estimates of health and health-related lifestyle.
- examine differences between population sub-groups (e.g. age, sex, social class) and local areas (health boards and local authorities).
- provide evidence to inform and monitor targets, indicators and policies for promoting better health, such as Programme for Government, Together for Health and Our Healthy Future.
- provide local authority level information for development of joint local health, social care and wellbeing strategies / single integrated plans.

User evaluation 2009

A user evaluation[2] was conducted in 2009 – this showed that the survey worked well to meet user needs, but that there was scope for improving the dissemination of results. A dissemination review[3] was conducted in 2010; this was subsequently followed by the production of a dissemination plan. Subsequently, a number of improvements were implemented such as the introduction of a teaching dataset on the UK Data Archive, the publication of a WHS user guide to provide a brief overview of the WHS and the main reports and datasets available, and a number of other changes including improvements to the WHS annual report and WHS theme pages.

User consultation on outputs 2011

A user consultation on WHS outputs ran from September to December 2011. The consultation sought feedback from users on outputs from the 2010 Welsh Health Survey to help us improve future publications. This is in line with the Official Statistics Code of Practice which states that effective user engagement is fundamental both to trust in

---

[1] NatCen Welsh Health Survey Technical Report
http://gov.wales/statistics-and-research/welsh-health-survey/?lang=en
[2] Welsh Health Survey User Evaluation
http://gov.wales/statistics-and-research/welsh-health-survey/?lang=en
[3] Welsh Health Survey Dissemination Review
http://gov.wales/statistics-and-research/welsh-health-survey/?lang=en

statistics and securing maximum public value. We published a summary of consultation responses in early 2012.

User consultation 2014

A consultation to seek the views of WHS users on the future approach to the survey took place between May-August 2014. A summary of the responses was published in November 2014. The overwhelming majority of responses were positive. However one issue that arose was around timeliness of when the data is released, to address this the Welsh Government has implemented a new publication plan for the 2014 survey. This consists of replacing the Annual report and LA/LHB bulletin with a series of topic based statistical bulletins throughout the year, these bulletins will include local authority/ local health board information. This change allows the prompter release of statistics which users highlighted as most important.

Strengths and limitations

The relatively large sample size for the WHS allows analysis by population sub-groups and local areas (although as with most sample surveys, robust analyses for small areas and groups are not possible). The survey covers a comprehensive range of health-related topics, and relationships between these can also be examined – however, it is not able to focus on individual topics in great detail. The survey has continued in its current form since 2003/04, allowing valuable time series to build up for trend analysis. Where possible, key questions are kept unchanged from year to year to build up these time series, although occasionally it may be necessary to make changes to ensure that the information continues to be relevant to current needs.

2. Accuracy

The closeness between an estimated result and the (unknown) true value.

Sampling error

As with any survey, results are subject to various sources of error. Sampling error arises because the estimates are based on a sample rather than a full count of the population. The results obtained for any single sample are likely to vary slightly from the true population value, and the difference between the estimates derived from the sample and the population values is referred to as the sampling error. In general, the smaller the sample size the larger the potential error.

The WHS was stratified at local authority level, with different probabilities of selection for different authorities. One of the effects of using this complex design is that standard errors for the survey estimates are generally higher than the standard errors that would be derived from a simple random sample of the same size.

Design factors

The ratio of the standard error of the complex sample to that of a simple random sample of the same size is known as the design factor. Put another way, the design factor (or ‘deft’) is the factor by which the standard error of an estimate from a simple random sample has to be multiplied to give the true standard error of the estimate, given the complex design.

True standard errors and design factors for a selection of key WHS variables were calculated using a Taylor Series expansion method, and are shown in NatCen's Technical Report[1].

Confidence intervals

A confidence interval can be calculated around a survey estimate, which gives a range within which the true value for the population is likely to fall. The standard error measures the precision with which the estimates from the sample approximate to the true population values and is used to construct the confidence interval for a survey estimate. NatCen's Technical Report[1] contains confidence intervals for a selection of key variables for Wales. The Initial Headline Results statistical release contains tables of approximate confidence intervals for a range of key variables which can be used as an approximate guide to precision, although as noted in the release it does not take account of complex survey design. For local authority estimates, two years of data combined are used to increase the sample size and provide more robust estimates.

Non-sampling error

Non-sampling errors can be defined as errors arising during the course of all survey activities other than sampling. Unlike sampling errors, they can be present in both sample surveys and censuses. Substantial efforts have been made to avoid non-sampling error in the Welsh Health Survey. For example, the survey results are weighted to take account of unequal selection probabilities, and for differential non-response, i.e. to ensure that the age and sex distribution of the responding sample matches that of the population of Wales. Another cause of bias may be that interviewers systemically influence responses in some way, however extensive interviewer training is provided to avoid this, and interviewer bias is less likely to have a large impact on data collected through a self-completion questionnaire.

Response rates

Response rates are an important dimension of survey quality and are monitored closely. The WHS achieves high response rates - in 2015, 76 per cent of eligible households took part, and self-completion questionnaires were obtained for 77 per cent of adults and 73 per cent of children in participating households.

Question development

Many questions have been subject to cognitive testing – this helps to ensure that they are consistently understood as intended, and that respondents can recall the information needed to answer them. Surveys are widely used to obtain information on health-related behaviours, but this can be a complex area to measure and there may be some differences between what people report and what they do (for instance, they may tend to underestimate their alcohol consumption or overestimate their levels of physical activity – both can be difficult for individuals to assess). However, survey data still provides a reliable means of comparing patterns for these behaviours between different groups and over time.

Data Validation

Data keying and scanning

The household questionnaires are double keyed in-house at NatCen. The self-completion questionnaires are sent to a scanning agency, and following scanning, the self-completion questionnaires, data and electronic images are sent to NatCen and the data linked to the household data through serial numbers (at both household and individual levels).

A report is run comparing the household data to the data booked in at the scanning agency and subsequently scanned. For cases where the data cannot be immediately matched a 'problem file' is produced. Reconciliation procedures are then undertaken to match up household data and self-completion discrepancies (for instance, error in the serial number, individual name or number).

Data coding and editing

The self-completion questionnaires are edited using NatCen's in-house system.

The data are checked to correct cases where routing has not been followed, where respondents have coded more than one answer where only one was required, or where incompatible answers have been entered.

As a separate checking measure all handwritten digits on the questionnaires are verified visually as part of the quality control process.

More information is provided in NatCen's Technical Report[1].

Data testing

The data are organised into three data sets and these are delivered to the WHS team at the Welsh Government for initial testing. Two productive test data sets at the individual level are produced – one for adult data and one for child data. A combined data set is also created containing information from all productive households at the individual level (household data for productive and unproductive individual cases).

The datasets are initially checked by NatCen, and the Welsh Government team ensures that all the required variables are present in the data sets, and frequency checks are run to check that the data set variables are generating the correct results. Syntax for all derived variables is also documented, and is checked by both NatCen and the Welsh Government when first produced.

Missing answers

Missing answers occur for several reasons, including refusal or inability to answer a particular question, and cases where the question is not applicable to the informant. Missing answers are omitted from all tables and analysis.

Revisions, Errors and Postponements

In the unlikely event of incorrect data being published, revisions would be made and users informed in conjunction with the Welsh Government's Revisions, Errors and Postponements arrangements[4].

3. Timeliness and punctuality

Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the time lag between the actual and planned dates of publication.

For the WHS, the lapse of time between the end of the reference period and the publication of first results is around 5 months. Fieldwork for a given year ends in December (with any outstanding data collected in January), and initial headline results for key variables for that year are published the following June. The series of statistical bulletins which cover each topic area in more detail are published on a staggered basis from June-September, all outputs are available on the WHS theme pages[5].

All WHS outputs adhere to the Code of Practice by pre-announcing the date of publication through the Upcoming calendar[6] web pages. Furthermore, should the need arise to postpone an output this would follow the Welsh Government's Revisions, Errors and Postponements[4] arrangements. The publication date has never been missed.

4. Accessibility and Clarity

Accessibility is the ease with which users are able to access the data, also reflecting the format(s) in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the metadata, illustrations and accompanying advice.

Publication

Statistics from the WHS are published in an accessible, orderly, pre-announced manner on the Welsh Government website at 9:30am on the day of publication. An RSS feed alerts registered users to this publication. Simultaneously the releases are also published on the National Statistics Publication Hub. We also publicise our outputs on Twitter[7]. All releases are available to download for free.

Key results for Wales are available to download from the StatsWales[8] website and this can be manipulated online or downloaded into spreadsheets for use offline.

---

[4] The Welsh Government's Revisions, Errors and Postponements arrangements
http://gov.wales/statistics-and-research/about/statement-of-compliance/revisions-errors-postponements/?lang=en
[5] Welsh Health Survey results
http://gov.wales/statistics-and-research/welsh-health-survey/?lang=en
[6] Upcoming calendar web page
http://gov.wales/statistics-and-research/?upcoming=true&lang=en
[7] Twitter
http://twitter.com/StatisticsWales
[8] StatsWales
https://statswales.gov.wales/Catalogue

Alongside the WHS topic based statistical bulletins, additional Excel tables are published which contain more detailed data and breakdowns.

Further information regarding the statistics can be obtained by contacting the relevant staff detailed on the release or via stats.healthinfo@wales.gsi.gov.uk .

UK Data Archive

An anonymised version of the main dataset, together with supporting documentation, is deposited with the UK Data Archive each year (some information is removed to ensure confidentiality is preserved). These datasets may be accessed by registered users for specific research projects. A simplified introductory dataset containing a small number of key variables which may be helpful for new data users is also available. The UK Data Archive is accessed via the UK Data Service website[9]. From time to time, researchers may wish to analyse more detailed data than is available from the Data Archive. Requests for such data are considered on a case by case basis, and confidentiality procedures are adhered to.

Ad-hoc Outputs

A number of ad-hoc outputs have been published using WHS data, including a bulletin analysing the mental health and wellbeing of adults in Wales, and a bulletin on carers. Joint reports with Public Health Wales have also been produced using WHS data and other sources, including a report on smoking in Wales, a profile of alcohol and health in Wales, and profiles of lifestyle and health in local health boards. WHS data is also used in the Chief Medical Officer for Wales' Annual Report and a wide range of other documents.

Methods and definitions

Each of the WHS statistical releases and bulletins[5] contain a methods and definitions section at the end, providing a detailed description of each variable within the bulletin and information on sampling and weighting of the data. It also gives brief descriptions of the socio-demographic classifications and how the age-standardisation is applied. The additional Excel tables published alongside the statistical bulletins contain a 'Notes' sheet with key definitions. Copies of the questionnaires are available online.

[9] UK Data Service website: Welsh Health Survey
http://ukdataservice.ac.uk/get-data/key-data.aspx#/tab-uk-surveys

Disclosure control

Outputs are designed to ensure that individuals are not identifiable from the published results. We adhere to our statement on confidentiality and data access[10], issued in conformance with the requirements set out in Principle 5: Confidentiality of the Code of Practice for Official Statistics.

Language requirements

We aim to use Plain English in our outputs and all publications adhere to the Welsh Government accessibility policy[11]. Furthermore, all our headlines are published in Welsh and English, as is the WHS Initial headline results statistical release.

5. Comparability

The degree to which data can be compared over both time and domain.

Welsh Health Survey results can be compared for different sub-groups (e.g. age, sex, geography, socio-demographic factors). The survey was established in 2003/04, allowing comparisons over time for a number of key measures**.** From 2007 the survey ran on a calendar year basis, and collected more detailed data for children. Some caution should be used when comparing trends over time for sub-groups such as local authorities or children, since changes are not always significant due to smaller sample sizes.

Over time, some variables used in the survey have been subject to minor definitional changes. Where comparisons are not available as a result of changes to variable definitions, this is noted clearly in outputs.

6. Coherence

The degree to which data that are derived from different sources or methods, but which refer to the same phenomenon, are similar.

Differences in methodology and questions mean that in general, results from the WHS are not comparable with health surveys in the other UK countries. However links are provided in WHS outputs to the other UK health surveys for reference. The Scottish Health Survey did publish a Topic Report[12] on UK comparisons, to analyse some results from the four health surveys conducted in the UK where appropriate.

Within Wales, the Census provides some comparability with rates of limiting long term health problem or disability in the WHS. A higher proportion of adults in households reported a limiting long term health problem in the WHS 2011 (34%) than the 2011 Census (26%); although for those being limited a lot the numbers were close (16% in the WHS and 14% in the Census). The higher numbers in the WHS may reflect the different context in which the questions were asked – in the WHS, people were already thinking about their health and limitations, while in the Census the surrounding questions were on other topics not related to health.

---

[10] Statement on confidentiality and data access
http://gov.wales/statistics-and-research/about/statement-of-compliance/confidentiality-data-access/?lang=en

[11] Welsh Government's Accessibility Policy
http://gov.wales/accessibility/?skip=1&lang=en

[12] Scottish Health Survey Topic Report on UK Comparisons
http://www.scotland.gov.uk/Publications/2010/08/31093025/0

There are few alternative sources on illness in the whole population that can be used to gauge the robustness of WHS results at LA level. Survey results for adults reporting limiting long term health problem or disability by local authority from WHS (using data for 2011 & 2012 combined) were compared with similar information on adults in households from the 2011 Census. A correlation coefficient of around 0.8 for any limitation and 0.9 for limited a lot indicated a strong relationship.

Where possible, results are compared with data from other sources to check their credibility, and are generally in line with expectations (although they are not directly comparable because of differences in methodology, definitions etc). For instance, the Quality and Outcomes Framework[13] (QOF) holds individual GP disease registers of patients, and covers some illnesses that can be compared with Welsh Health Survey rates (e.g. heart failure and high blood pressure), and although they are not measuring quite the same thing, rates are generally consistent[14].

Future of WHS

It has been decided to replace existing surveys, including WHS, with a new survey of adults starting during 2016-17 which will include health-related questions. WHS will run until the end of 2015 and then cease in its current form.

Feedback

We welcome comments from users of our publications on content and presentation. If you have any comments or require further information, please contact:

Health Statistics and Analysis Unit
Welsh Government, Cathays Park, Cardiff CF10 3NQ
E-mail: stats.healthinfo@wales.gsi.gov.uk

Produced by Knowledge and Analytical Services, Welsh Government
Last reviewed: September 2016

[13] Quality and Outcomes Framework
http://gov.wales/statistics-and-research/general-medical-services-contract/?lang=en
[14] A Comparison of Health Conditions in the Quality and Outcomes Framework and the Welsh Health Survey
http://gov.wales/docs/statistics/2012/120822healthconditionsen.pdf