Background Quality Report

Annual Population Survey: UK Armed Forces veterans residing in Great Britain, 2016

The purpose of a background quality report is to inform users of the statistics about the quality of the data used to produce the publication, and any statistics derived from that data. Existing uses of the statistics and user requirements are also discussed.

This assessment relates to the 'Annual Population Survey: UK Armed Forces veterans residing in Great Britain' Official Statistic published by Defence Statistics (Health).

1. Introduction

1.1 Overview

This statistical bulletin is an annual publication providing estimates on the size and socio-demographic characteristics of the UK Armed Forces veteran population living in households in England, Scotland and Wales using responses provided in the 2016 Annual Population Survey (APS) produced by the Office of National Statistics (ONS). Summary figures and comparisons to the non-veteran population residing in Great Britain (GB) are presented on: people characteristics; regional location; health; employment status; education and accommodation (housing). The non-veteran population is defined as those aged 16+ who have not served in the UK Armed Forces or are currently serving.

These statistics are published as Official Statistics, adhering to the UK Statistics Authority (UKSA) protocols on pre-release access[1].

The sections below review this statistical release in line with the Quality assurance Framework of the European Statistical System[2], alongside other considerations which are relevant to measuring and reporting on quality.

1.2 Background

The Annual Population Survey (APS) (formerly the Integrated Household Survey (IHS)) is a quarterly survey of households in the UK conducted by the Office of National Statistics. The surveying of UK households resulted in over 289,000 individual responses in 2016. Veteran questions were asked of respondents aged 16 and over and residing in England, Scotland or Wales. Veterans were identified as those who had previously served; non-veterans were those who had never served or were currently serving. Using these criteria 13,000 respondents were identified as veterans and 214,000 as non-veterans.

More information on the coverage of the APS (formerly known as the Integrated Household Survey (IHS)) and the survey itself can be found at; http://www.ons.gov.uk/ons/guide-method/method-quality/specific/social-and-welfare-methodology/integrated-household-survey/index.html.

[1] UKSA Protocol on Pre-Release Access: http://www.statisticsauthority.gov.uk/about-the-authority/uk-statistical-system/legislation/pre-release-access/index.html

[2] Quality Assurance Framework of the European Statistical System: http://ec.europa.eu/eurostat/documents/64157/4392716/qaf_2012-en.pdf/8bcff303-68da-43d9-aa7d-325a5bf7fb42

The MOD have sponsored the veteran questions to provide the evidence base required by Government, third party and the private sector to aid policy development in support of the Armed Forces Covenant. This report also supports MOD's commitment to release information wherever possible.

Defence Statistics has engaged with subject matter experts in the ONS ensuring the design of the APS, the large sample size and the difference in the age and gender structure of the veteran population to the non-veteran population have been taken into account in the analysis.

Previously the only estimate on the number of UK Armed Forces veterans and information on the socio-demographic characteristics was from the Royal British Legion 2014 UK Household Survey of the Ex-Service Community. We have greater confidence in the estimates within this Official Statistic as they have been calculated from a larger sample size (RBL sample size was approximately 25,000).

1.3 Methodology and production

Data Sources

Defence Statistics received electronic data, via a secure portal, from the ONS on all responses to all questions asked in the APS. The data was held in SPSS.

Each respondent in the data set had a unique number which was used to link their responses across all questions

More information on the APS can be found on the ONS website**:** http://www.ons.gov.uk/ons/guide-method/method-quality/specific/social-and-welfare-methodology/integrated-household-survey/index.html

The 2014 publication of the RBL Household Survey, which analyses responses of the veteran population and Ex-Service Community, was used to identify key areas of interest. More information on the RBL Survey can be found at: http://www.britishlegion.org.uk/get-involved/campaign/public-policy-and-research/the-uk-ex-service-community-a-household-survey/

Data Coverage

These statistics included respondents from individuals living in households in England, Wales and Scotland (Great Britain) aged 16 and over. The veteran questions were not asked in Northern Ireland due to security concerns, therefore respondents living in Northern Ireland were not represented.

Changes in APS variables

From January 2016 the smoking variables within the APS survey data changed to focus more specifically on cigarette smoking status. The variables analysed within this report are:

- CIGEVER – Yes/No: Ever smoked cigarettes regularly. Includes only ordinary tobacco which is smoked. Exclude any reference to snuff, tobacco or tobacco products that are chewed or sucked or herbal tobaccos. Excludes e-cigarettes. Applies to all respondents aged 18 and over.
- CIGNOW – Yes/No: Smoke at all nowadays. Applies to all respondents aged 18 and over and when response in CIGEVER is 'Yes'.

Therefore comparisons with previously published estimates on smoking cannot be made.

Statistical Methods

In comparing the veterans to the non-veteran population there were three main statistical concerns which were addressed: the difference in the population structures; the large sample size and the sampling design.

- Differences observed may be as a result of the varying population structures. **Standardising** was used to enable us to take the population structures into account and be confident that any differences observed were true differences

- The large sample size can produce significant results even though the difference is only trivial. The **measure of effect** was used to identify differences which were large enough to note; not just those which were significantly different.

- The standard significant test formula assumes the data is from a simple random sample. The formula had to be adjusted using the **design factor** to take account of both of the standardisation and the sample design.

Weighting

A weight (PWTA17) was applied to the entire sample, by the ONS, in order to inflate the sample size to the population size ensuring it represents the population in terms of age, gender and location. This weight took the sampling design into account. More information on the weighting and other adjustments used by the ONS can be found at: http://www.ons.gov.uk/ons/guide-method/method-quality/specific/social-and-welfare-methodology/integrated-household-survey/index.html in the 'IHS user guide 2014' page 125

The veteran population was predominantly male and older than the non-veteran population. This difference had to be taken into account when comparing veterans to non-veterans to ensure any differences identified were true differences and not due to the different age and gender profiles.

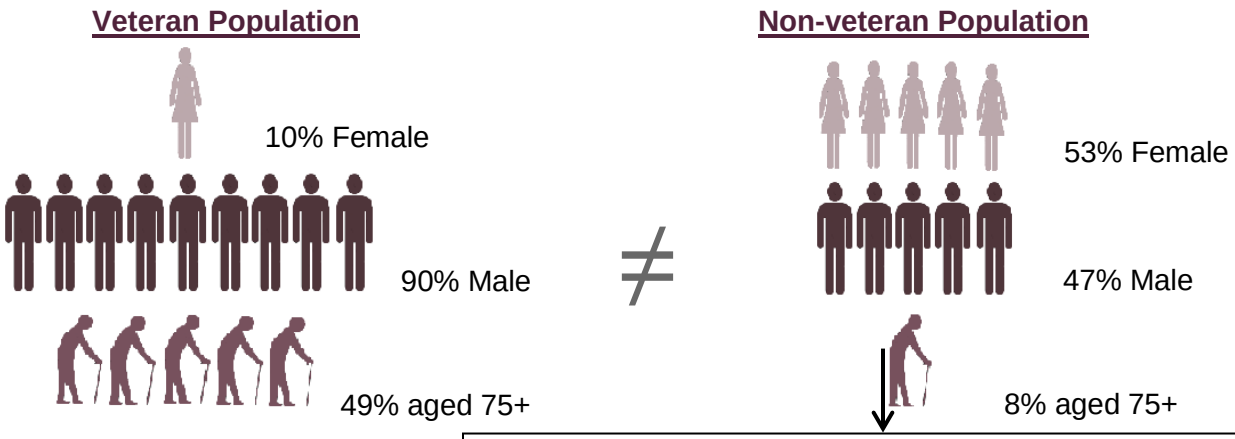
Additional weights (standardising) were assigned to the non-veteran population so the age and gender distribution mirrored the veteran population. This did not affect the numbers in the veteran population or the non-veteran population but inflated some responses and reduced others.

This was calculated for each age and gender group by:

$$\frac{\%\text{ of estimated veteran population}}{\%\text{ of estimated non-veteran population}}$$

Further weights were applied to both the veteran and non-veteran populations to account for the 9,973 individuals who took part in the APS but did not provide a response to the veteran questions. The percentage of people who answered the veteran question differed by age and gender; therefore probability of the non-responders being a veteran differed by age and gender.

Before standardisation; (People characteristics dissimilar across populations)



Using the following logic, the demographic structure of the non-veteran population was altered:

Examples:

- Every 5 female responses in the non-veteran population was reduced into an average of 1 response

- Every 1 response from those aged 75+ in the non-veteran population was increased into an average of 5 responses

After standardisation, (People characteristics similar across populations)

Percentage

Percentages enabled comparisons to be made between two populations i.e. veterans and non-veterans. The estimated number who gave particular response for the population was divided by the estimated total number in the population and multiplied by 100.

Unless otherwise specified, ‘don’t know’ and ‘not applicable’ responses were ignored and percentages were based only on the numbers of respondents who chose the remaining item response options.

Margin of Error

Each estimate carries a margin of error; which is presented in the supplementary tables. Margins of error provide a measure of the level of uncertainty in the estimate; or a measure of how reliable the survey is. The higher the margin of error, the less likely the results of the survey are true for the population. Large error margins are usually the result of having a small number of respondents within a particular group. Where the margin of error is large the results should be interpreted with caution.

The estimate plus and minus the margin of error provides the confidence interval around the estimate. The confidence interval provides a measure of the likely variation of the given statistic.

The margin of error calculation took account of the design factor.

Significant test

The z test: difference between two proportions was used to identify if there was a significant difference between the estimated percentages from the veteran and non-veteran responses. The significance test gave us the confidence to state that an observed difference between the percentages was a real difference. The confidence intervals, from the test, have been presented in the supplementary tables. If they do not contain zero we can conclude that the estimated veteran percentage and the percentage for (standardised) non-veterans were significantly different.

In addition to carrying out a z test, a significant difference between two estimates can be identified by comparing the confidence intervals around the estimates. However, the wider the confidence intervals around an estimate the least likely you are to find a significant difference. Therefore you may result in the confidence intervals overlapping where there is a true significant difference. Where the confidence intervals around two estimates do not overlap you can be confident that the two estimates are significantly different even if they have a large confidence interval or margin of error.

Due to the number of significance tests being carried out there was a higher likelihood that differences will be classed as significantly different when they are not (a false positive). The significance test was therefore carried out at the 99% level meaning there should be less than 1% (1 in 100) chance that differences observed in the APS results aren’t representative of the population as a whole

The significance test calculations took account of the design factor.

Design Factor

The weighting and standardisation applied can either improve the precision of a survey estimate or make them worse. In this case the weighting and standardisation decreased the precision. By using the formula for a simple random sample we were increasing the risk of stating there was a significance difference when there wasn't one. The design factor is applied to the margin of error and significance test calculations to account for this.

The design factor is the square root of the design effect.

The design effect is the ratio of the variance of the estimate from the complex sample design to one from a simple random sample – it quantifies the extent to which the expected sampling error in the survey departs from the sampling error that can be expected under simple random sampling.

The design effect is calculated by taking the average squared weights and dividing by the sum of the average weight.

Measure of Effect

When comparing two estimated proportions from large samples, a significant difference is more likely to be found even if the difference is only trivial. The measure of effect enables us to confirm that any difference observed were large enough to note. The effect size is independent of the sample size.

The standardized difference (d) for categorical responses, assuming each response option is a separate binary outcome, has been used[3].

$$d = \frac{(\hat{p}1 - \hat{p}2)}{\sqrt{\frac{[\hat{p}1(1-\hat{p}1) + \hat{p}2(1-\hat{p}2)]}{2}}}$$

$\hat{p}1$ = estimated percentage for veterans
$\hat{p}2$ = estimated percentage for non-veterans

Cohen's rule of thumb has been applied to identify small (d>=0.2), medium (d>=0.5) and large (d>=0.8) effects.

Non-response

Missing values, where respondents have not provided a response/valid response have not been included in the analysis.

An assumption has been made that those who do respond to the survey or a particular question are missing at random (MAR) and therefore equally as representative of the population as those who do respond.

The information presented in this publication has been structured to release information into the public domain in a way that contributes to the MOD accountability to the British public but which doesn't risk breaching individual's rights to medical confidentiality. In line with Defence Statistics' rounding policy, and in keeping with the Office for National Statistics Guidelines, all numbers less

[3] http://support.sas.com/resources/papers/proceedings12/335-2012.pdf

than three have been suppressed and presented as '..' to prevent the inadvertent disclosure of individual identities.

Relevance

The key users of these statistics are Government departments, Devolved Administrations, the third party and the private sector. The information will aid policy development in support of the Armed Forces Covenant.

The Armed Forces Covenant recognises that the Government's commitment to ensure the Armed Forces Community (currently serving, veterans and dependants) are not disadvantaged compared to other citizens.

This statistical bulletin provides a key statistic: estimated number of UK Armed Forces veterans residing in Great Britain. In addition it provides information on the socio-demographic breakdown of veterans and identifies if there are any differences in relation to the non-veteran population. This will aid policy makers identify where efforts should be focused in relation to the Armed Forces Covenant.

This bulletin is the primary means in which this information is made available in the public domain

Accuracy and Reliability

The APS sample size was large enough to achieve a small margin of error. A number of questions were only asked of a subset of respondents and they typically carry a larger margin of error. For example the question '*Does your condition or illness reduce your ability to carry out day-to-day activities?*' only applied to those who reported a long-term health problem and many employment questions only included those of working age. The margin of errors have been provided in the supplementary excel tables.

Where comparisons were made between the veteran and non-veteran population, 99% Z-tests were carried out. This level was used to minimise the possibility of finding false positives that can be expected when performing a large number of significance tests on such a heavily weighted data set. In addition to Z-tests, measure of effect tests were carried out ensuring only small (between +/-0.2 - 0.49), medium (between +/-0.5 - 0.79) or large (more than +/-0.8) proportional differences between veterans and the non-veteran population responses were highlighted.

As the APS does not achieve a 100% response rate there is a risk that those who returned questionnaires have differing views from those who did not. We assumed that all non-response rates are Missing At Random (MAR). This means we have assumed that those people who did not respond do not differ from those who did not respond. If those who did not respond have different demographics, lifestyles and experiences then the analysis in this report will be biased and will not be representative of the entire population; instead, our observation will only represent the responding population.

The statistical bulletin does not present any results where the responding group size is less than 3 as results for groups of this size are considered too unreliable.

A number of responses to the veteran questions were inconsistent and their veteran status adjusted. All responses stating they were currently serving but aged over 70 were reclassified to veteran. All responses stating they were currently serving as a regular but unemployed or economically inactive were reclassified to veteran. ONS are investigating the discrepancies and

have included extra advice to those asking the questions to reduce the likelihood of these discrepancies in future data.

Data Revisions

There are no planned revisions of this bulletin. Amendments to figures may be identified in future analysis. To ensure continuity and consistency, figures will only be adjusted during the year where it is likely to substantially affect interpretation and use of the figures otherwise required corrections will be released in future bulletins along with reasons for the corrections.

Timeliness and Punctuality

This release of the bulletin was published seven months after receiving the data from the ONS (ten months after the reference period). This was due to the development time required to develop the report.

This Statistical Bulletin was published on time to meet the pre-announced release date, in line with the Official Statistics' code of Practice. A one year release schedule outlining the following financial year's publication date is published on the Gov.UK website at least one month in advance.

Accessibility and Clarity

This Statistical Bulletin was published on the Gov.uk website and was made available to public from 0930 hours on the day of release: https://www.gov.uk/government/collections/annual-population-survey-uk-armed-forces-veterans-residing-in-great-britain.

24 hour pre-release access to the report was available to a limited distribution list within MOD and the Scottish Government. The full list can be found in the pre-release access list available on the Gov.UK website: https://www.gov.uk/government/statistics/defence-statistics-pre-release-access-list.

All tables and figures included full footnotes to ensure any conditions or caveats were made clear. A glossary of key terms and simple explanations of the statistical methods used were provided in bulletin. Tables and figures from each statistic were separately available in MS Excel format for users to download. This allows for use in individual research and reports.

Key findings have been presented on the first page so that users can quickly focus on the important results, with more detailed commentary presented within the Statistical Bulletin.

Coherence and Comparability

All definitions within this bulletin are consistent with other MOD and wider Government department releases. For example the employment status definitions were consistent with those used by the ONS.

Comparisons have been made to the results published from the Royal British Legion Household Survey; http://www.britishlegion.org.uk/get-involved/campaign/public-policy-and-research/the-uk-ex-service-community-a-household-survey/

Confidentiality, Transparency and Security

These statistics do not contain any identifiable personal data. The information presented in this publication has been structured in such a way to release information to the public domain that contributes to the MOD accountability to the British public but which doesn't compromise data protection.

All Defence Statistics staff involved in the production have signed a declaration that they understand their responsibilities under the Data Protection Act and the Official Statistics Code of Practice. All staff involved in the production process have signed the Data Protection Act, and all MOD, Civil Service and data protection regulations are adhered.

Defence Statistics adhere to the principles and protocols laid out in the Code of Practice for Official Statistics and comply with the pre-release access arrangements. The Defence Statistics Pre-Release Access lists are available on the Gov.uk website.