Background Quality Report

Self-reported workplace injury and work-related illness statistics from the Labour Force Survey

Contents

Introduction 2
Assessment of statistics against quality dimensions and principles 3
Relevance 3
Accuracy and Reliability 5
Timeliness and punctuality 6
Accessibility and clarity 6
Coherence and comparability 7
Trade-offs between output quality components 8
Assessment of user needs and perceptions 8
Performance, Cost and Respondent Burden 9
Confidentiality, transparency and security 9

This document is available from www.hse.gov.uk/statistics/
Introduction

This report provides users with an evidence based assessment of the quality of the statistical output from the module of questions on workplace injury and work-related illness included annually in the Labour Force Survey (LFS). It does this by reporting against the nine quality dimensions and principles of the European Statistical System (ESS) quality framework.

In doing so, this meets our obligation to comply with the UK Statistics Authority (UKSA) Code of Practice for Official Statistics 8, particularly Principle 4, Practice 2 which states:

“Ensure that official statistics are produced to a level of quality that meets users’ needs, and that users are informed about the quality of statistical outputs, including estimates of the main sources of bias and other errors, and other aspects of the European Statistical System definition of quality.”

The Labour Force Survey (LFS) is a nationally representative survey currently consisting of around 44,000 responding households each quarter. It provides a rich source of information about the UK labour market, and is designed, developed and managed by the Office for National Statistics (ONS) in Great Britain and by the Department of Finance and Personnel in Northern Ireland on behalf of the Department of Enterprise, Trade and Investment (DETINI).

The Health and Safety Executive (HSE) commission a set of questions annually in the LFS, taking advantage of existing arrangements for sampling and interviewing a large nationally representative sample, to gain a view of work-related illness and workplace injury based on individuals' perceptions. This includes a core set of questions, and a small number of additional questions asked periodically. Whilst information is collected from Northern Ireland, this information is not routinely published since HSE's jurisdiction is restricted to Great Britain only.

An annual series of workplace injury questions have been included in the LFS since 1993/94 and work-related illness questions since 2003/04 (and periodically prior to this since 1990). Information about the number of days off work due to workplace injury and work-related ill health has also been collected since 2003/04 (and periodically prior to then). In 2012/13, the ill health data collection was suspended for one year, but from 2013/14 returned to annual data collection.

The ONS carry out extensive testing and quality assurance work before releasing the LFS data sets. Any new questions or changes to the LFS are tested and evaluated before being rolled out. On receipt of the full LFS data set from the ONS, various processes are completed including validating the HSE questions, adjusting for non-response to the screening questions (which identify cases of work-related illness and workplace injury) and deriving a range of variables for analysis purposes.

The analysis is completed using the statistical software package Stata, which presents 95% confidence intervals, as well as estimates of levels and rates of work-related ill health and workplace injury.

Statistics are released annually through the HSE statistics web pages (www.hse.gov.uk/statistics/), alongside statistics on work-related illness and workplace injury from other data sources. The release includes a range of detailed tables (www.hse.gov.uk/statistics/lfs/) and written commentary embedded throughout the web pages. For more detailed analyses, users are able to define (within certain parameters) their own tabulation requests using an interactive tool called HandS-On (handson.hse.gov.uk/hse/public/home.aspx). This is currently available for 2001/02-2012/13 data.

Results are only released where sample numbers are sufficiently large to provide reliable estimates and disclosure standards are adhered to. Published tables and outputs from HandS-On include 95% confidence intervals.

Documentation on the data source and methodological issues can be found at:

- Data sources (www.hse.gov.uk/statistics/sources.htm#lfs)
- Technical note (www.hse.gov.uk/statistics/lfs/technicalnote.htm)
Assessment of statistics against quality dimensions and principles

Relevance

This dimension covers the degree to which the statistical product meets user need in both coverage and content.

LFS respondents are asked to recall events over the previous 12 months, and so results reflect an average of the previous 12 month period. For example, respondents administered the questions in the LFS quarter one in 2008 (January 2008-March 2008) cover the 12 months prior to each interview. The results reflect an average of 12-month periods ending January 2008 to March 2008, and this is described as 2007/08.

The LFS gives estimates on the levels of workplace injury. All estimates are based on an individual's most recent workplace injury in the 12 month reference period, and exclude injuries caused by road accidents since they fall outside HSE's jurisdiction. Information is presented as estimated incidence and rates of all non-fatal injuries, over-3-day absence injuries and over-7-day absence injuries where:

- Estimated injury incidence is the estimated number of people reporting a workplace injury in the 12 months prior to interview (the reference period);
- Injury incidence rate is defined as the injury incidence estimate divided by the annual estimate of employment (this is taken as the number of individuals reporting themselves as currently employed);
- All non-fatal injuries include all self-reported workplace injuries;
- Over-3-day absence injuries include those self-reported injuries resulting in more than three consecutive (working and non-working) days away from work (not counting the day on which the accident happened);
- Over-7-day absence injuries include those self-reported injuries resulting in more than seven consecutive (working and non-working) days away from work (not counting the day on which the accident happened).

Injury measures by demographic and employment related variables are generally presented as three-year averages to provide more robust estimates.

More details about the definitions and formulae can be found at www.hse.gov.uk/statistics/lfs/injury.htm

For illness, the LFS gives estimates of the number of people who have conditions which they think have been caused or made worse by work (regardless of whether they have been seen by doctors). Estimates are based on the most serious work-related illness, if an individual has more than one. Information is presented as estimated prevalence and rates of self-reported illness and estimated incidence and rates of self-reported illness where:

- Estimated prevalence is the estimated number of people with a work-related illness at any time during the 12-month reference period. It includes the full range of illnesses from long standing to new cases;
- Prevalence rate is defined as the prevalence estimate divided by the population at risk of having a work-related illness;
- Estimated incidence is the estimated number of new cases of work-related illness occurring in the 12 month reference period i.e. people first becoming aware of their illness in this 12 month period;
- Incidence rate is defined as the incidence estimate (restricted to individuals working in the 12 month period) divided by the population at risk of experiencing a new case of work-related illness during the reference period.

As a result of a routing error in the LFS questionnaire 2007/08 and 2008/09 ill health measures are restricted to people working in the last 12 months rather than people ever employed (as in earlier surveys). Hence, all published estimates are generally restricted to people working in the last 12 months for comparison purposes.

More details about the definitions and formulae can be found at www.hse.gov.uk/statistics/lfs/illness.htm
The LFS also provides estimates and rates of the total number of days off work due to work-related illness and/or workplace injury where:

- Working days lost are expressed as full-day equivalent days to allow for variation in daily hours worked and includes days lost due to all non-fatal injuries and all work-related illness (new and long standing cases);
- Working days lost due to work-related illness is a measure of the total time lost due to all episodes of the illness over the 12 month reference period whereas working days lost due to workplace injury is a measure of the elapsed time between injury and returning to work and does not include any subsequent time taken off work;
- Rates presented are in the form of average annual working days lost (full-day equivalent) per case of work-related illness or workplace injury and average annual working days lost (full-day equivalent) per full-time equivalent worker.

More details about the definitions and formulae can be found at [www.hse.gov.uk/statistics/lfs/dayslost.htm](http://www.hse.gov.uk/statistics/lfs/dayslost.htm)

All measures of both workplace injury and work-related illness from the LFS are referred to as 'self-reported'. For ill health in particular, estimates represent an individual's perception of the contribution that work made to the illness, rather than a medically verified estimate. Research indicates that self reports of work-related illness are broadly reliable [www.hse.gov.uk/research/rrhtm/rr970.htm](http://www.hse.gov.uk/research/rrhtm/rr970.htm).

Each measure, with corresponding 95% confidence intervals (see [www.hse.gov.uk/statistics/lfs/errors.htm](http://www.hse.gov.uk/statistics/lfs/errors.htm)), is presented by a range of demographic and employment-related variables, including age and sex, region, occupation, industry and workplace size. The level at which the data is presented is dependent on the level at which it was collected in the LFS (and also the number of sample cases for any given analysis – see [www.hse.gov.uk/statistics/lfs/errors.htm#a13](http://www.hse.gov.uk/statistics/lfs/errors.htm#a13)). Individual characteristics such as age and sex and region are available for all survey subjects. However, employment-related information such as occupation and industry are only available for those who are currently in work or whose most recent job was in the last 8 years.

For more detailed analyses, users are able to define (within certain parameters) their own tabulation requests using an interactive tool called HandS-On ([handson.hse.gov.uk/hse/public/home.aspx](http://handson.hse.gov.uk/hse/public/home.aspx)) for the period 2001/02-2012/13. Whilst the tool accesses a restricted subset of the LFS data set, it will only report aggregated data where sample numbers are sufficiently large to provide reliable estimates and where disclosure standards are adhered to. The data set includes industry, occupation, region (at Government Office Level), country, age (in 10 year age bands) and gender, allowing cross-classification with ill health and accident data. Standard outputs include estimates of levels and rates, and associated 95% confidence intervals.

The LFS provides an essential source of regular outcome data for HSE, ensuring an evidence base for decisions about targeting and prioritisation. It also provides supplementary data on workplace injuries to meet European requirements, and furthermore, it forms a major component of the economic cost model of health and safety failings. These statistics are also used by the wider community e.g. industry, academics, members of the public, students.

To ensure these statistics meet the needs of the user, they are reviewed using a variety of methods. In 2011 HSE ran an external user-engagement exercise and in 2014 HSE hosted a statistics user conference for external users. The aim of the conference was to allow users to meet with data source leads to find out more about the information currently available, as well as to provide HSE with the opportunity to engage directly with its external customers. Results of the consultation and details about the conference can be found at [www.hse.gov.uk/statistics/about/engagement/previous-consultations.htm](http://www.hse.gov.uk/statistics/about/engagement/previous-consultations.htm). We also make available contact details of the responsible Statistician in the main statistical documents, and provide a feedback form on the website, to allow users to comment on any aspect of the statistical outputs. In addition, we have a popular electronic email bulletin service, providing regular information to users and reminding them of forthcoming releases and other relevant key events.
Accuracy and Reliability

The proximity between an estimate and the unknown true value

Estimates from the LFS are subject to various sources of error, namely sampling and non-sampling error, where non-sampling error includes coverage error, non-response error, measurement error and processing error. The errors encountered are described in detail in the Labour Force Survey User Guide Volume 1: Background and Methodology (www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/labour-market-statistics/volume-1-2011.pdf).

Sampling errors and confidence intervals

The LFS survey data is used to make inferences about the whole population. When data obtained from a sample is used in this way, there is an element of sampling error, or uncertainty, about the sample estimate. Sampling errors relate to the fact that the chosen sample is only one of a very large number of samples which may have been chosen, each giving rise to different sample estimates. The spread of these results is the sampling variability, which generally reduces with increasing sample size. A confidence interval is a range of values, defined by a lower and upper bound, which indicates the variability of an estimate. Statistical methods are used to calculate the sampling variability from which the confidence interval can be determined. For example, with a 95 per cent confidence interval, it is expected that in 95 per cent of the survey samples, the resulting confidence interval will contain the true value that would be obtained by surveying the whole population (see www.hse.gov.uk/statistics/lfs/errors.htm).

Results are only released where sample numbers are sufficiently large to provide reliable estimates (www.hse.gov.uk/statistics/lfs/errors.htm#a13) and disclosure standards are adhered to. Published tables (www.hse.gov.uk/statistics/lfs/) and outputs from HandS-On (handson.hse.gov.uk/hse/public/home.aspx) include 95% confidence intervals.

A difference between two rates is ‘statistically significant’ if there is less than 5% chance that it is due to sampling error alone (see www.hse.gov.uk/statistics/lfs/errors.htm#a14). Most of the published tables are accompanied by a table containing statistical significance tests (located via a link in the footnote of published tables). For example, testing whether rates for independent groups, such as males and females, or a rate for a particular group compared with the overall rate, such as the construction industry compared to all industries, are statistically significantly different.

One way of increasing the reliability of survey data is to increase the sample size on which it is based. Whilst the annual sample size is fixed, several years’ worth of data can be pooled to produce estimates for the average of the combined years. Some results – particularly for workplace injury but also for work-related illness by industry and occupation where the number of annual sample cases at detailed levels tend to be low - have been produced in this way by pooling the three years worth of data. Although published tables only pool three years worth of data, HandS-On has the facility to pool three, five or seven years of data. However, results should be treated with caution, particularly injuries which follow an overall downward trend.

Non-sampling errors

Many of the sources of non-sampling error are difficult to measure. However, in the LFS ‘Performance and Quality monitoring Report (PQM) (www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/labour-force-survey/) detailed response rates are published.

The LFS collects data on a sample of the population. To convert this information to give estimates for the population, the sample data is weighted. Each case is given a weight which can be thought of as the number of people that case represents. This weighting factor takes account of differential non-response among different sub-groups in the population. This weighting procedure involves grossing data to sub-regional population estimates and then adjusting for the estimated age and sex composition, by region. These weighting factors are provided to HSE by the Office for National Statistics (ONS) along with the LFS dataset.

Whilst the LFS weights take account of non-response to the LFS itself, they do not take account for non-response to individual questions, in particular the screening questions which identify cases of work-related illness and workplace injury. To reduce non-response bias in the health and safety specific estimates, an adjustment is made to the weights to take account of this non-response (see www.hse.gov.uk/statistics/lfs/weighting.htm).

ONS periodically revise the LFS weights. HSE assess the impact on the work-related illness and workplace injury published estimates, and a judgement about the need to revise is made. Any subsequent revisions made are noted in the revision log (www.hse.gov.uk/statistics/about/revisions/revision-log.htm).
Workplace injury and work-related ill health measures from the LFS are based on self reports. The very nature of ill health means that there will always be some level of under-reporting. Ill health is more difficult to define than a workplace injury as the person first has to acknowledge they have an illness and then attribute that illness to current or previous work. Self reports from the Labour Force Survey capture the most widely based definition of work-related illness from a large well established representative sample. However, just as some individuals will not recognise or report conditions genuinely caused by work this will to some extent be counter balanced by individuals ascribing the cause of their illness to work when there is no such link. Even with these discrepancies, individuals are uniquely well-placed to assess the role that work factors play in their illness. They are in a position to follow in detail how particular aspects of work have impacted them and to observe their body’s response to this.

There is no gold standard measure of work-related illness, all available sources are subject to some kind of error, and we cannot rely on just one to cover all aspects of ill health. Research undertaken in 1995 (www.hse.gov.uk/statistics/2002/swi95.pdf) and 2010 (www.hse.gov.uk/research/rrpdf/rr970.pdf) indicate a reasonable degree of reliability in self reports of work-related ill health in the LFS, and when sensibly interpreted, such surveys provide valid and relevant information not available from other sources.

**Timeliness and punctuality**

*Timeliness refers to the time gap between publication and the reference period. Punctuality refers to the gap between planned and actual publication dates.*

The time period between the end of the reference period and the delivery of the LFS data set by ONS to HSE is six weeks. The scheduled availability date is included in the LFS Performance and Quality Monitoring report (www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/labour-force-survey/index.html).

The release of results on work-related illness and workplace injuries are brigaded with other health and safety statistics and published at the end of October, approximately five months after the receipt of the LFS data set e.g. 2012/13 data was published in October 2013.

Results are published in accordance with the UKSA Official Statistics Code of Practice. HSE pre-announce the actual date of release one month in advance, on the HSE website and the UKSA publication hub, and through a subscriber-based email bulletin. The processes involved to achieve this publication deadline are well-established, hence have not resulted in a delayed release. Should a delay to the release date be unavoidable, relevant information will be provided to users in advance.

**Accessibility and clarity**

*Accessibility is the ease with which users are able to access the data, also reflecting the format 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.*

Results on work-related illness and workplace injuries from the LFS are published on the HSE statistics web site (www.hse.gov.uk/statistics/) each October. A wide range of detailed tables and charts are available, with interpretation, free of charge.

Top-level results and commentary are presented as conventional web pages, with further analysis contained in stand-alone PDF documents. For more experienced users and those requiring specific data, supporting Excel files containing aggregated tables are also provided. PDF and Excel files are viewable by anyone, using freely-available software.

Again to cater for specific needs and more experienced users, a restricted subset of LFS record-level data is also searchable through an interactive tabulation tool called ‘Hands-On’. This allows users the flexibility of defining (within certain parameters) their own tabulations. Outputs from the tool can be saved to the users’ computer for further re-use.

For confidentiality reasons, suppression rules are employed to avoid inappropriate disclosure. A HSE statistical confidentiality policy is publicly available at www.hse.gov.uk/statistics/about/confidentiality.htm. Reliability thresholds (www.hse.gov.uk/statistics/lfs/errors.htm#fa13) are also implemented to meet ONS' recommended requirements for publication of LFS data.

The UK data archives at Essex University provide free access to various LFS datasets, and can be contacted via the UK Data Archive website. The HSE module of questions is made available after the annual ‘first release’ of results, but only the standard LFS weights are included on the datasets.

This document is available from www.hse.gov.uk/statistics/
Coherence and comparability

Coherence is the degree to which data which have been derived from different sources or methods but refer to the same topic are similar. Comparability is the degree to which data can be compared over time and domain.

Coherence

A range of data sources are used to measure work-related illness, including the Department for Work and Pensions’ (DWP) industrial disease compensation scheme and the statutory employers reports (under the Reporting of Injuries, Diseases and Dangerous Occurrence Regulations) which are restricted to well established occupational diseases, and voluntary medical surveillance schemes (The Health and Occupation Reporting Network – THOR) which require consultation with a medical practitioner or specialist doctor. Self reports from the LFS are the most inclusive. The self-management of more minor cases of ill health is not included in any of these schemes. However, all available data sources have strengths and weaknesses and we cannot rely on one data source to cover all aspects of work-related ill health. The HSE uses a range of data sources, and a table of preferred statistical sources for different categories of ill health can be found at www.hse.gov.uk/statistics/preferred-data-sources.htm.

Comparisons between self reports of work-related ill health and GP assessed cases (THOR GP) show a good degree of correlation (www.hse.gov.uk/research/rrpdf/rr954.pdf). Results also confirm the preferred statistical sources for different categories of ill health.

The workplace injury questions incorporated in the LFS were designed to cover injuries reportable under the administrative reporting system RIDDOR (Reporting of Injuries, Diseases and Dangerous Occurrences Regulations), as well as all non-reportable workplace injuries. The LFS provides a more complete view than RIDDOR, and by comparing results directly with RIDDOR it provides an indication of the levels of reporting.

Standard questions on workplace injuries and work-related illness form the basis of a Eurostat ad hoc module included in the LFS in EU countries. Two such surveys have been undertaken so far (1999 and 2007). A further ad hoc module on this subject has been included in the 2013 EU LFS.

Comparability

An annual series of workplace injury questions have been included in the LFS since 1993/94 and work-related illness questions since 2003/04 (and periodically prior to this since 1990). Since 2003/04 (and periodically prior to then), information about the number of days off work due to workplace injury and work-related ill health has also been collected. In 2012/13, the ill health data collection was suspended for one year, but from 2013/14 returned to an annual data collection.

The definitions and formulae associated with measures of work-related illness, workplace injury and working days lost can be found at www.hse.gov.uk/statistics/lfs/survey.htm.

Results are available on a consistent basis for:

- workplace injury from 1998/99;
- work-related illness from 2001/02;
- working days lost due to workplace injury from 2000/01 and due to work-related illness from 2001/02.
Whilst the small change in survey period from December-February for years prior to 2006/07 to January-March from 2006/07 onwards may have potentially introduced a discontinuity to these data series, investigations suggested that this change in survey design has not affected the top level injury, ill health and working days lost data (see [www.hse.gov.uk/statistics/pdf/lfsissue1.pdf](http://www.hse.gov.uk/statistics/pdf/lfsissue1.pdf)).

The above surveys had the same design and coverage (with the exception of the small change in survey time period in 2006/07) and used the same methods to adjust for non-response to the screening questions in each survey.

It is only possible to make broad comparisons with the earlier surveys because of differences in design, coverage, the level of information collected and methods used to adjust for non-response to the screening question in each survey. To achieve even these broad based comparisons of illness data, a number of adjustments have to be made to the data to ensure comparability, and these need to be treated with caution.

The LFS uses standard classifications e.g. Standard Industrial Classification (SIC) and Standard Occupational (SOC), and incorporates harmonised questions e.g. to collect demographic and employment related information, which are used in other UK surveys.

**Trade-offs between output quality components**

*This dimension describes the extent to which different aspects of quality are balanced against each other.*

The main trade off is one of timelines, against accuracy and or/detail. The aim has always been to publish the full range of LFS results to meet customer needs as quickly as possible, but ensuring that the quality of the product is maintained. In the past, results have been published in two parts, headline results in October (brigaded with other health and safety statistics) and more detailed results in the following March. With more efficient systems in place to produce tables, and the move to publishing results in excel tables, rather than html files, the full range of LFS tables have been published alongside the Annual Statistics Report from October 2011.

Since 2010, users have been able to define (within certain parameters) their own tabulation requests using the interactive HandS-On tool. This allows the production of more detailed tables by the user, without impacting on the release date.

**Assessment of user needs and perceptions**

*This dimension covers the processes for finding out about users and uses, and their views on the statistical products.*

We have published a policy on ‘user engagement’ on the HSE website: [www.hse.gov.uk/statistics/about/engagement/](http://www.hse.gov.uk/statistics/about/engagement/)

This policy explains how we engage with and encourage feedback from users, allowing them to provide views in ways that suit them. We have also published the results from various large-scale user-engagement exercises, the latest in 2011, and information about a HSE hosted user conference in 2014. In addition to the formal consultation exercises, we regularly analyse or engage with:

- Statistics web page usage;
- Hands-On tabulation tool usage;
- Ad-hoc statistical requests;
- New ideas, and pro-actively engage with our statistical peers, to gain insights into ‘what works’ with analysis and presentation of statistics;
- ONS and other Government department developers/users.

We also frequently seek views on different statistical issues from subscribers to our bi-monthly statistics e-bulletin.
Performance, Cost and Respondent Burden

*This dimension describes the effectiveness, efficiency and economy of the statistical output.*

The current average annual cost of sponsoring questions in the LFS is £315k and the estimated annual burden on individuals responding to the sponsored questions is 136 person days. This is based on the number of respondents and the estimated time to respond. As described in the above quality measures, there is ongoing work carried out by statistical staff in HSE in the processing, analysing and presenting the data. The staffing costs associated with this work are approximately £60k each year.

Confidentiality, transparency and security

*The procedures and policy used to ensure sound confidentiality, security and transparent practices.*

A Confidentiality Policy (which also covers data security) is available on the HSE website: www.hse.gov.uk/statistics/about/confidentiality.htm

In summary, this says:

All of our data is handled, stored and accessed in a manner which complies with Government and Departmental standards regarding security and confidentiality, and fully meets the requirements of the Data Protection Act. Access to this data is controlled by a system of passwords and strict business need access control.

To ensure transparency of data release, any revisions to our publications are handled in accordance with the Department’s revisions policy, which is published on the statistics section of the HSE web site. This gives details on the circumstances of when a revision might take place, as well as a log of past revisions: www.hse.gov.uk/statistics/about/revisions/

These statistics also comply with the UK Statistics Authority Code of Practice on release protocols. In particular, pre-release access to the data is strictly controlled. The Policy can be viewed at www.hse.gov.uk/statistics/about/index.htm

LFS results are only released on the HSE Statistics web site or through HandS-On where sample numbers are sufficiently large to provide reliable estimates and disclosure standards are adhered to (see www.hse.gov.uk/statistics/lfs/errors.htm#a13).
National Statistics

National Statistics are produced to high professional standards set out in the National Statistics Code of Practice. They undergo regular quality assurance reviews to ensure that they meet customer needs. They are produced free from any political interference.

An account of how the figures are used for statistical purposes can be found at www.hse.gov.uk/statistics/sources.htm.

For information regarding the quality guidelines used for statistics within HSE see www.hse.gov.uk/statistics/about/quality-guidelines.htm.

A revisions policy and log can be seen at www.hse.gov.uk/statistics/about/revisions/.

Additional data tables can be found at www.hse.gov.uk/statistics/tables/.

Statistician: Jacky Jones
Contact: jacky.hd.jones@hse.gsi.gov.uk

Last updated: October 2014
Next update: October 2015