

RIDDOR Background Quality Report

Injury statistics as reported under RIDDOR

Contents

Introduction	2
Background	2
Purpose of the document	3
Assessment of statistics against quality dimensions and principles	3
Relevance	3
Accuracy and reliability	4
Timeliness and punctuality	5
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	8
Confidentiality, transparency and security	9



Introduction

This document constitutes a background quality report, for the publication of quarterly and annual data from incidents reported by employers and others in Great Britain under RIDDOR (The Reporting of Injuries, Diseases and Dangerous Occurrences Regulations 2013).

The statistics included in this release are for:

- Quarterly figures on fatal injuries.
- Annual figures on fatal and non-fatal injuries.

Background

Employers and others (also known as duty-holders) in Great Britain have a legal obligation to report certain incidents where a person is killed or suffers a specified injury whilst at work. This reporting requirement is called RIDDOR, and is designed to inform the authorities (HSE, local authorities and for railways, the Office of Rail Regulation), who may then decide to investigate. Whilst this type of data is generally referred to as administrative data, a key use of this information is statistical, enabling trends and areas of higher risk to be identified, with a view to development of accident prevention strategies. Although HSE considers it is aware of virtually all relevant fatal injuries, it is important to note that non-fatal injuries are substantially under-reported.

In Great Britain health and safety regulation of workplaces, including the requirements of RIDDOR is the responsibility of HSE. Incidents occurring in Northern Ireland, and consequent statistics, are the responsibility of HSE Northern Ireland (HSENI).

Guidance for duty-holders is available online at the HSE website, as to what is and is not reportable to HSE. Actual reporting is also made online, with an option to report fatal and other serious incidents by telephone. See www.hse.gov.uk/riddor. Alternatively companies reporting frequently can use an automated data transfer as an option. Persons covered by RIDDOR are workers (i.e. employees and the self-employed), and persons not at work but still 'affected' by a work activity (often referred to as members of the public).

To produce these statistics, data is regularly extracted from the RIDDOR database, and a copy made for statistical purposes. The statistical data is taken through a series of validation checks, where implausible data values are looked into, and if necessary adjustments are made prior to publication. In the case of fatal injuries, additional checks are made by correlating with additional sources of information, usually related to an investigation of the incident.

Currently around 75,000 worker injury notifications are made each year. This data forms part of a statistical series dating back to the 1970's (under separate legislation), and has always been publicly available. For RIDDOR, aggregated statistics are available by 'planning year' (1 April - 31 March). For most statistical dimensions, we provide injury incidence rates as well as counts. We do this by using employment estimates provided by the Office for National Statistics (ONS) as a denominator.

These statistics are released through the HSE website, alongside related injury statistics from complementary sources, particularly the LFS (Labour Force Survey). The format of the release is a combination of data tables (e.g. Excel) and written commentary. Many outputs also conform to European standards, meaning comparisons can often be made with EU member states.

Documentation on the statistical processes and known data quality issues is provided at www.hse.gov.uk/statistics/sources.htm. This also describes changes in the reporting system from September 2011; also the legal changes to RIDDOR from April 2012 and October 2013 and effects these changes may have on the data series.

Purpose of the document

This paper aims to provide users with an evidence-based assessment of the quality of the statistical output from the RIDDOR system. 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, 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.”

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.

HSE has published these statistics for many years. Statistics covered by this publication fall into the following legally-defined categories:

- Fatal injuries to workers and members of the public.
- Defined ‘specified’ injuries to workers. This classification changed from ‘major’ in October 2013.
- Other injuries to workers, resulting in more than seven days off work. The previous over-three-day threshold changed from April 2012.
- The most recent legal change in October 2013 has affected all the non-fatal data series. In addition as the change happened halfway through the year 2013/14, the full-year statistics for 2013/14 contained data based on the old as well as the new reporting requirements. Outputs for each of these years clearly reference these changes, for example discontinuity lines in charts.

Typical top-level analyses are fatal and non-fatal injuries (including specified/major and over-7-day injuries to workers) over annual (planning year) time periods. Further detail is also available by the following categories:

- Workplace characteristics (type of industry, geographical location of the workplace, and type of incident leading to the accident).
- Person characteristics (age, gender).
- Injury outcome (type of injury and part of body affected).

Statistical outputs are usually in three tranches:

- Fatal injuries are released for the most recent quarter. As this data is based on early results and intended to give an ‘emerging picture’, the data is not complete or detailed, nor is the data badged as ‘National Statistics’.
- Fatal injuries are released annually. This data is not as timely as above, partly due to additional information on fatality rates being included, although the data is complete and allows wider comparisons to be made. The latest year is always ‘provisional’, with the previous year being ‘finalised’.

- Non-fatal injuries are released annually. The latest year is always 'provisional' and denoted 'p', with the previous year being 'finalised' and denoted 'r' (to show the previously-published figure has been revised). This data series is subject to substantial under-reporting, current levels of reporting for employees is estimated at around a half; and the reporting of injuries to the self-employed a much lower proportion. No adjustment is made for this, and figures published are 'as reported'.

To ensure these statistics meet the needs of the user, they are reviewed using a variety of methods. In 2011 and 2014 HSE ran external user-engagement exercises, showing RIDDOR to be a very widely-used data source by many different types of user. Details can be found at www.hse.gov.uk/statistics/about/engagement/previous-consultations.htm. We also make available on the website a feedback form, 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

This dimension covers, with respect to the statistics, their proximity between an estimate and the unknown true value.

Accuracy

All data are published 'as-is', currently with no adjustment to counts of accidents (i.e. no weightings applied, no estimations). However, we sometimes apply rounding and/or data suppression on small counts to avoid disclosure, or if the denominator (employment data) is inconsistent with the numerator (injuries).

For fatal injuries, the data outputs conform to the structure of RIDDOR, as described in 'relevance' above. However, due to the sensitive and serious nature of fatal accidents, additional sources of information are also used by operational staff, just one of which is a formal RIDDOR notification from employers. In addition all fatal injuries are investigated (a large majority of non-fatals are not), meaning the statistics are highly accurate.

For HSE-enforced fatalities, data is recoded on an internal operational database, and maintained continuously with the most recent information in readiness for publication. For LA-enforced fatalities, no central database exists as there are around 380 LAs. In these cases, HSE uses the original RIDDOR notification as a basis, but checks with the relevant LA to confirm the accuracy of the details. The enforcement of railway safety is the responsibility of the Office of Rail Regulation (ORR), and ORR provides HSE with the relevant data for fatalities recorded on their systems.

As we use a variety of sources for fatal injuries, they are not subject to under-reporting so no adjustment to the outputs is necessary.

For non-fatal injuries, numerically there are far more of these reports than fatal injuries, with a relatively small percentage actually investigated. Hence the information provided by the employer when reporting is taken largely at face value, as we have no cost-effective way to verify the information provided.

Quality assurance is carried out on all data. Case-by-case fatal data is assured as described above, and below is the outline processing for non-fatal injuries. The initial quality focus is to make sure that the top-level number of reports is in line with expectations (severity by employment status). Subsequent quality assurance examines individual fields in detail to ensure that any unexpected changes or data values are understood and represent genuine changes.

The main quality assurance strategy can be summarised as follows. The descriptions may run sequentially, but where necessary some or all elements are re-run as necessary:

- The original RIDDOR database contains some internal validation checks. These 'hard' checks stop notifiers from entering incorrect data, with a warning message to explain what is wrong so that they can correct it, for example an incident date in the future.
- For the statistical dataset, analysts perform data cleaning checks on data items that are not covered adequately by the database internal validation checks. All analysts have considerable knowledge of RIDDOR, for example whether certain types of incident are technically reportable or not (if not, the records are removed).

- We perform basic checks on file size and characteristics to make sure they are in line with expectation. Detailed quality assurance checks are run on the raw data such as analysis of the new dataset against previous datasets, looking for changes in means; levels of change in each data item; the incidence of missing values; changes in categorical variables; extreme values.
- Where possible records that require re-coding due to inaccuracies are done en-block by automation. It also often happens that batches of records require case-by-case examination of certain data fields. In both scenarios a judgement is made by analysts as to the correct codes to use. Where necessary additional expert opinion on a case is obtained from operational and policy colleagues.
- Tabulations are produced and compared with the outputs generated during the same period the previous year. We look at the statistics from an operational and policy point of view to ensure that any unexpected changes or data values are understood and represent genuine changes. Where uncertainties exist in data outputs, these are highlighted and if possible an explanation given.

Data published for the most recent year is given a 'p' status (provisional), and a year later is released as final (with an 'r' notation to depict revised). This allows for late reports and where necessary additional quality checks, although the changes are usually small, typically increasing by around 1-2%.

This background quality report includes known data quality issues affecting the analysis in the release and is produced annually as part of the wider Health and Safety Statistics compendium publication. Ultimately users of the data must make their own assessment of the quality of the data for a particular purpose, drawing on these resources. In addition, their own subject knowledge - or other comparative data sources - may be required to distinguish changes in volume between reporting periods that reflect changes in incidents, from those that are an artefact of changes in data quality or reporting practices. Such issues should be kept in mind when viewing time series analysis since year-on-year changes may sometimes be a product of shortfalls in earlier years, and should not automatically be interpreted as trends in incidents.

Reliability/Known data quality issues

Fatal injury data is highly reliable, due to case-by-case scrutiny by operational and analytical staff. However, non-fatal data is subject to significant under-reporting (current levels of reporting for employees is estimated at around a half; and the reporting of injuries to the self-employed a much lower proportion). Furthermore, previous research into reporting levels suggests variation in reporting between different industries. Reporting levels across other dimensions like geographical region may also differ substantially.

Documentation on the statistical processes and known data quality issues is provided at www.hse.gov.uk/statistics/sources.htm. This also describes the change in the system used for reporting from September 2011, the recent legal changes to reporting, and possible effects these may have on the data series.

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 relevant 'date' for these statistics is the date individual incidents took place, not the date reported or investigated. Similarly if a death occurs some time after the incident, again the incident date is used and not the actual date of death. To allow for late reports, data quality checks and preparing for publication, there will be some delay from the period end to actual release.

- Quarterly fatal statistics are released 3 months after the period of collection ends, for example Q1 2016 (incidents occurring between 1 April to 30 June 2016) would be released end September 2016.
- Annual fatal statistics are released about 3 months after the period end, for example the full-year 2015/16 (1 April 2015 to 31 March 2016) were released early July 2016.
- Due to the volume of records, annual non-fatal statistics require more validation work. In addition they are released alongside other health and safety data, to present a more complete picture, especially in respect of reporting levels. Such statistics are released about seven months after the period end. A full-year such as 2015/16 (1 April 2015 to 31 March 2016) would be released early-November 2016.

In accordance with the UKSA Official Statistics Code of Practice, HSE pre-announce the actual date of release of the above series one month in advance on the HSE website and the Government-wide Statistics Release Calendar and through a subscriber-based email bulletin.

The processes involved to achieve these publication deadlines are well-established, hence have not resulted in a delay to release in recent years. Should a delay to the above release timetables 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.

Accessibility

Alongside this data quality statement, relevant summaries of the results included in this publication are accessible via the HSE website, and to which the Government-wide Statistics Release Calendar also directs users.

Given the wide range and depth of these statistics, a variety of web-based publication formats are used (no charges are made to users). 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. 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

Clarity

The statistics presented as web pages and PDF documents use plain language, and to avoid technical terms use the same descriptors as employers use when filling-in the original notification forms. Where some technical terms are deemed necessary, they are defined where appropriate, usually through a footnote or annex. Where it is considered helpful to explain analyses, charts and maps are used to supplement text and data tables.

When trends over time are presented, they can show possible discontinuities in the series for a variety of reasons. Should this occur, it will be highlighted and if possible explanations provided, for example as a consequence of the change in reporting requirements for over-7-day injuries from April 2012, and for the change from major injuries from October 2013. It has to be remembered that non-fatal injuries are substantially under-reported, and where necessary users are reminded in the presentation of the data. This is a key reason why non-fatal statistics are released alongside equivalent data from the LFS, as it enables a more complete picture to be provided to users (see 'comparability' section below).

'Metadata' can be described as "data which provides information about one or more aspects of the original data", for example, what we mean by 'industry classification' and how this dimension may have changed over time. Separate metadata documents and Excel look-up Tables, describing all the variables, are available at <http://www.hse.gov.uk/statistics/sources.pdf>

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

This publication refers to just one data source (RIDDOR). Where possible outputs use counts (number of injuries) as well as 'rates' of injury. That is, the number of injuries divided by the number of persons 'exposed' (employed) to the same work activity. Estimates of employment are provided by the Office for National Statistics (ONS). By calculating a rate of injury, this enables fairer comparisons across given dimensions to be made. For example, simply comparing the number of injuries in (say) England would be much higher than (say) Scotland. Dividing each figure by the relevant number of workers in the respective countries puts both on a more comparable footing.

Comparability

There are two relevant aspects when comparing RIDDOR statistics. Firstly, there are comparisons within RIDDOR itself, but over various dimensions. As well as comparing over time, these also include: geographical location of accident; age and gender of injured person; main industry of employer (eg construction); employment status of injured person; type of injury and part of body affected (e.g. broken arm); type of accident (eg contact with machinery).

For most of the above dimensions, it is possible to make comparisons, but only within each dimension. To enable sensible comparisons, a rate has to be calculated as in 'coherence' above, and it is the rates that are compared, not the number of injuries. By convention, rates are typically expressed in units "per 100,000 employed".

Comparing RIDDOR statistics over time is normally on a yearly basis, where the year is defined in the 'timeliness' section above. Many analyses require the 'latest picture', meaning the most recent year or two's worth of data, although some time series are available back to 1974 (when the Health and Safety at Work Act was introduced) – see <http://www.hse.gov.uk/statistics/history/index.htm>. Where a long time series shows an unexpected break in the trend, it is highlighted and if possible an explanation provided, for example as a consequence of the change in reporting requirements for over-7-day and major injuries from April 2012 and October 2013.

The actual reporting mechanism has changed twice, initially in 2001, and again in 2011. The 2001 change introduced a dedicated call centre, as well as web reporting for the first time (previously it was mainly paper format). The 2011 change enhanced online reporting further. As often happens with system changes, it can have subtle effects on the data series, and most of the effects of the 2001 change are summarised in www.hse.gov.uk/statistics/pdf/discontinuity2.pdf. The 2011 changes are referenced in the 'reliability' section above.

Importantly, when making comparisons it has to be remembered that non-fatal injuries are substantially under-reported. This adds to the complexity of analysis, as reporting levels within one dimension can vary. For example within 'industry', one category (eg construction) could have higher or lower reporting than another (eg agriculture), so different adjustments are needed to see the true picture.

For the second aspect to comparisons, it may be possible to compare statistics from RIDDOR, with other sources, of which there are mainly two:

- European countries and beyond, especially fatal injuries. This is possible, as many of the data fields collected in GB under RIDDOR also align well with data collected by all European member states – there is EU legislation that promotes this consistency (though note RIDDOR statistics do not capture work-related road traffic accidents (RTAs) unlike most other European countries).
- As injury-related questions on the annual UK Labour Force Survey (LFS) are designed to match RIDDOR, again direct comparisons between those sources are helpful. This is especially true in estimating levels of RIDDOR reporting, which currently run at around half for employees (and substantially less for self-employed).

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 timeliness, against accuracy and/or detail.

As mentioned in 'timeliness' above, first releases of fatal and non-fatal data are always provisional ('p'). The following year the figures are then denoted 'r' to indicate a revision. For fatal injuries this means we release as quickly as possible after the period-end, as there is considerable interest in such statistics. One trade-off is a finer level of detail, which has to wait until later in the year. Accuracy also has to be balanced against speed of release. As virtually all fatal injuries are investigated, it can often take considerable time for the full facts to be known, and reflected in the statistics.

For non-fatal injuries and given the much larger volume of records, they cannot be checked for data quality one-by-one. Although HSE and local authorities are aware of all reported incidents, the majority are not investigated, so the data has to be taken at face value. Even those records that are subject to statistical checks, as described in 'accuracy' above, require considerable resource and time. Therefore a balance has to be struck, between the levels of data accuracy considered acceptable, against speed of publication.

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

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. In addition to the formal consultation exercises, we regularly analyse:

- Statistics web page 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.

We also frequently seek views on different statistical issues from subscribers to our statistics e-bulletin.

Performance, Cost and Respondent Burden

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

RIDDOR data in this context is referred to as 'administrative data'. That is, reporting under RIDDOR is a legal requirement, so the data is already available for statistical use. Consequently, there is minimal cost and burden to the respondent (that is, the duty-holder), in providing the data for statistical purposes.

As described in the above quality measures, there is ongoing work carried out by statistical staff in HSE in processing, analysing and presenting the data. The staff costs associated with this work is approximately £150k 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/revisions/

To avoid the disclosure of personal information through statistical outputs, disclosure control is implemented where deemed necessary, especially where small counts are involved.

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: Heidi Edwards

Contact: heidi.edwards@hse.gov.uk

Last updated: November 2016

Next update: October 2017

© *Crown copyright* If you wish to reuse this information visit www.hse.gov.uk/copyright.htm for details.
First published 10/14.