

## Police recorded road collision statistics

### What are these statistics?

Police recorded road collisions statistics provide information about road collisions in Wales which resulted in personal injury. Police forces are required by law to collect information about such incidents, including details about casualties, vehicles, locations and factors that led to the collision.

Provisional summary statistics on StatsWales and an interactive dashboard are updated on a quarterly basis, and detailed final analysis and commentary are published in annual reports.

These statistics only include road collisions which resulted in personal injury and for which information was reported by the police. It is known that there is an element of under-recording of such incidents, particularly for less severe incidents where police officers may not have been in attendance.

These statistics do not include collisions:

- That were not reported to the police
- That occurred on private land i.e. car parks or fields
- Where no personal injury was recorded
- Where the collision was later confirmed by a medical professional or coroner to be a suicide or medical episode

These statistics feed into the [Department for Transport's statistics on reported road collisions and casualties in Great Britain](#).

### Policy and operational context

Road safety is devolved to the Welsh Government and Welsh Ministers are the Highways Authority for motorways and trunk roads in Wales. These statistics are essential to support analysis and policy development on road safety in Wales. Context for road safety interventions by the Welsh Government and its partner organisations is given in the [Road Safety Framework for Wales](#), which was published in July 2013. This included the creation of targets to reduce the number of people killed and seriously injured on Welsh roads by 2020. A new strategy is currently being developed and is planned to be published later this year (2024).

## Users and uses

The data are used by the Welsh Government to assess the effectiveness of current policy, develop and assess the impact of new policies and road safety interventions and monitor trends of the number of people killed and seriously injured on Welsh roads. The Welsh Ministers may analyse the data to promote road safety in Wales. Much of the data is important in understanding the operation of the road network and driver behaviour.

In addition, a variety of other organisations use these data to support road safety assessments and identify and address collision hotspots. They include police forces, road safety groups, local authorities, transport planners and highway authorities. The statistics are also frequently covered by national and local media in reporting the latest picture and trends in road safety.

The statistics included in this release do not cover any of the national indicators, however they may be relevant in the context of some indicators and they may be used by public services boards in relation to their local well-being assessments and plans.

## Strengths, limitations and coverage of the data

### Strengths

- The data are the most comprehensive and reliable source of information on road collisions and casualties in Wales.
- The information is processed and published on a quarterly basis to enable users to identify and act on emerging road safety issues.
- The statistical outputs have been developed to meet user needs and cover a wide range of topics relevant to users.
- The [road collisions dashboard](#) enables users to manipulate the underlying data by a range of geographical and demographic variables to create bespoke analyses to supplement those in the statistical bulletins.
- The statistics are derived from existing administrative systems operated by police forces.

### Limitations

- There are long-standing issues around under-coverage, particularly with less serious collisions, affecting the statistics in Wales and across Great Britain. For example, hospital, survey and compensation claims data indicate that many non-fatal collisions are not reported to or recorded by the police. This is further discussed in the [coverage](#) section.
- Interpreting short-term trends can be problematic because of changes in police recording practices and road users' reporting of incidents.
- Quarterly updates are regarded as provisional because they are often under-estimates compared with final year's data.

- Since data on severity of injuries and factors contributing to incidents are typically based on police officers' subjective assessments, they often don't provide a comprehensive picture. However, this is due to change with a roll out of injury based reporting throughout 2024.
- Delays in data provision have occasionally resulted in postponements to statistical releases.
- From 2024 a new vehicle category "personal powered transporter device" will be included. Prior to this "E-scooters" would be classed as "other" and can only be identified using a free text field.
- Additional validation carried out in 2024 identified some inconsistencies between the speed limit of the road where the collisions occurred recorded by the police officer in attendance, compared with data held by local authorities. This is further discussed in the [accuracy](#) section.

## Coverage

The longstanding issues with under-coverage of collisions is one of the most significant quality concern with these statistics. Under-coverage falls into two categories:

- Collisions resulting in personal injury not reported to the police. Research by DfT<sup>1</sup> with other data sources such as insurance claims, hospital admissions data and survey data indicates that around 50% of collisions involving some degree of personal injury, accounting for two thirds of all non-fatal casualties, are not reported to police. Evidence suggests the majority of these injuries are relatively minor, and it is not known how many would have qualified as recordable injuries had police attended the scene.
- Collisions reported to the police but not recorded by the police. We have identified some instances of this through our QA processes, which flag significant changes in the data which we then query with police. In these scenarios police forces have usually, but not always, been able to retrospectively collate and supply the missing information. But by their nature, these occurrences are only identifiable if there are relatively large discrepancies compared with previous years' data. It is not known to what extent there is under recording on a small scale or on an ongoing basis. Improving awareness around the need to record collisions through the STATS19 system has been a feature of our work with police forces. As with non-reported collisions, the significant majority of non-recorded cases are likely to be 'slight' injury collisions.

Overall, the available sources show that collisions reported to and recorded by police forces represent only a subset of all personal injury road collisions, but that coverage of serious injuries and fatalities is good.

In addition to under coverage of collisions, there is a known under recording of serious injuries in the non-injury based reporting systems. The Department for Transport (DfT) together with the

---

<sup>1</sup> Summarised under 'strengths and weaknesses' in [Reported Road Casualties for Great Britain](#).

Office for National Statistics (ONS) have developed a methodology to quantify the affect of the introduction of injury based reporting systems. on the number of slight and serious injuries reported to the police, and to estimate the level of slight and serious injuries as if all police forces were using injury-based reporting systems.

We do not currently apply any adjustments to our collision or casualty figures, meaning there will be an underreporting of serious injuries and therefore serious collisions.

DfT publish both adjusted and unadjusted figures in their data tables. The unadjusted figures are comparable with the figures we publish.

Dyfed Powys began using the CRaSH tool in mid-May 2023. This means it's likely that there will be an increase in the number of seriously injured casualties reported by Dyfed Powys from this point onwards, and therefore an increase in the number of serious collisions. We currently don't have enough data from Dyfed Powys to understand the scale of the impact on serious collisions.

As more data is collected, we will work with DfT to adjust the figures if appropriate for the other three police forces to ensure comparability within Wales.

This issue does not affect the total number of collisions, or the total number of fatalities.

Throughout 2024 a new specification is being rolled out to all forces on either IBRS or non-IBRS, which will record the severity of a collision based on the severity of the worst injury, instead of the police officer's judgement.

## Data processing cycle

### Data collection

The [Annual Data Requirement \(ADR\)](#) confers a statutory duty on police forces in Great Britain to collect information about road traffic collisions which result in personal injury. There are two different ways of collecting this information.

1. Police officers record this data on a ['STATS19' form](#)
2. Police officers input the data into a standardised reporting tool (called CRASH – Collision Recording and Sharing)

Data are collated at a police force level onto administrative systems and transferred to the Welsh Government on a routine basis. Most of the information is recorded at the scene of the incident by attending officers, however, the Stats19 forms may be completed or updated subsequently when more information is available.

CRASH is an injury based reporting system (IBRS) whereas the STATS19 form is a non-injury based reporting system, meaning that a number of the variables are dictated by the officers judgement, instead of by the injuries of the casualties.

There are four police forces in Wales; South Wales Police; Gwent Police; Dyfed-Powys Police; North Wales Police.

As of May 2023, Dyfed-Powys are using the CRASH system, while the other three police forces remain on the non-injury based reporting system (stats19 2011 format).

## Validation and verification

The Welsh Government runs a suite of automatic validations on the raw data provided by police forces. These checks are intended to identify records containing unusual, invalid or missing data and these are flagged for validation with police forces. Some examples of the types of checks are:

- Age of drivers within expected range
- Invalid collision geographical locations
- Checks for duplicate records
- Internal consistency of records
- Confirming which local authority the collision occurred in when along police force boundaries

When these validation checks are resolved Welsh Government compiles a near-final dataset and runs comparisons against previous years' aggregates. This is done on a quarterly and annual basis for Wales as a whole and for the police force areas. This helps to identify unusual trends in the statistics which can indicate missing data, which can be queried again with the police.

We also compare against other sources of information, for example intelligence from meetings with police forces, information reported to the [Road Safety Wales group coordinated by Royal Society for the Prevention of Collisions \(ROSPA\)](#), and fatal collision reports supplied to Welsh Government.

The data are provided to the UK Department for Transport (DfT) to feed into their statistical releases. DfT also run a suite of automatic validations which can flag up additional quality issues such as duplicate records, but these are typically on a very small scale.

The final data are shared with police forces for final quality assurance and sign off before publication.

## Publication

Once the data has been finalised the statistical bulletin, dashboard and StatsWales tables are compiled. All elements of the release are independently checked and a final sense check is carried out by the responsible statistician prior to publication on the website.

The final data also feed into the Department for Transport's statistical release on [reported road casualties in Great Britain](#).

## Definitions

A full list of notes and definitions relating to these statistics is published by DfT, though the following are particularly important:

**STATS19 data:** The core set of statistical data which the police have agreed to provide about personal-injury road traffic collisions.

**Collision:** Involves personal injury occurring on the public highway (including footways) in which at least one road vehicle or a vehicle in collision with a pedestrian is involved and which becomes known to the police within 30 days of its occurrence. Damage-only collisions, with no human casualties or collisions on private roads or car parks are not included.

**Collision severity:** The severity of a collision is based on the severity of the most severely injured casualty, as outlined below.

**Casualty:** A person killed or injured in a collision. Casualties are sub-divided into killed, seriously injured and slightly injured.

- **Killed (fatal):** Human casualties who sustained injuries which caused death less than 30 days after the collision. Confirmed suicides are excluded.
- **Seriously injured:** Injured casualties are classified as seriously or slightly injured based on the type of injury sustained, as recorded by the police on the basis of information available within a short time of the collision. This generally will not reflect the results of a medical examination, but may be influenced according to whether the casualty is hospitalised or not. Injuries classed as serious include: fractures, concussion, internal injuries, crushings, burns (excluding friction burns), severe cuts, severe general shock requiring medical treatment and injuries causing death 30 or more days after the collision
- **Slightly injured:** An injured casualty that is not classified as seriously injured, having an injury of a minor character such as a sprain (including neck whiplash injury), bruise or cut which are not judged to be severe, or slight shock requiring roadside attention. This definition includes injuries not requiring medical treatment.

**Injury based reporting:** A system of recording collisions through the Collision Reporting and Sharing (CRASH) system. In these systems, casualty injury severity is coded from a list of injuries, rather than being based on the judgement of the reporting officer. In some forces, officers complete the information required on a mobile device.

## National Statistics status

The [United Kingdom Statistics Authority](#) has designated these statistics as National Statistics, in accordance with the Statistics and Registration Service Act 2007 and signifying compliance with the [Code of Practice for Statistics](#). National Statistics status means that statistics meet the highest standards of trustworthiness, quality and public value and it is our responsibility to maintain compliance with these standards.

All official statistics should comply with all aspects of the Code of Practice for Statistics. They are awarded National Statistics status following an assessment by the UK Statistics Authority's regulatory arm. The Authority considers whether the statistics meet the highest standards of Code compliance, including the value they add to public decisions and debate. The designation of these statistics as National Statistics was confirmed in March 2020 following a [review against the Code of Practice for Statistics](#).

Since the last review by the Office for Statistics Regulation, we have continued to comply with the Code of Practice for Statistics, and have made the following improvements:

- Published additional contextual information surrounding road collisions, such as the length of roads by speed limit
- Improved visuals by de-cluttering and standardising charts and tables

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

## **Administrative data quality assurance**

This release has been scored against the UK Statistics Authority's Administrative Data Quality Assurance matrix. The matrix is the UK Statistics Authority regulatory standard for the quality assurance of administrative data. The Standard recognises the increasing role that administrative data play in the production of official statistics and clarifies what producers of official statistics should do to assure themselves of the quality of these data. The toolkit that supports it provides helpful guidance to statistical producers about the practices they can adopt to assure the quality of the data they receive, and sets out the standards for assessing statistics against the Code of Practice for Official Statistics.

The criteria against which statistics are scored are as follows:

- Operational context and administrative data collection
- Communication with data supply partners
- Quality assurance principles, standards and checks applied by data suppliers
- Producer's quality assurance investigations and documentation

The matrix enables us to make an assessment about the risk profile of our statistics based on public interest and quality concerns. We have assessed these statistics as 'A2/A3' because we consider the statistics to be of medium/high data quality concern and of medium public interest. This implies that our assurance practices should be either 'Enhanced' (A2) or 'Comprehensive' (A3) depending on the relevance of the particular practices to our work.

Our initial assessment was that our existing processes met or exceed the requirements at the A2 and A3 levels of assurance in most cases, though there were also areas of the framework requiring improvement. Existing processes were strong in areas such as communication with suppliers, our internal quality assurance and our understanding of the operational context of the data collection. We identified some areas of relative weakness, for example providing information around sources of bias and error and understanding suppliers' quality assurance processes. We have since taken steps to improve our practices in these areas, including:

- Met with individual police forces to develop better understanding of their data collection and quality assurance processes. Helped forces to improve their processes and established

more collaborative approaches to QA by increasing the frequency and efficiency of data validations.

- Created a STATS19 data providers group with the police forces to share best practice, and stress the importance of data quality.
- Formalised agreements around roles and responsibilities in data collection, quality assurance and supply.
- Use a wider range of information to help verify data, including fatality reports, local authorities' intelligence and information presented to the Road Safety Wales group.
- Provide more detailed information about the sources of error in the data, with comparisons against related data sources and discussion around the likely scale of under-coverage.

Our work across these and other areas of the QAAD framework is ongoing and we continue to work with police forces and stakeholders to identify where improvements to data quality and statistical dissemination can be made.

## **Disclosure control and confidentiality**

The published statistics on road collisions do not include any personally or commercially sensitive information requiring action around disclosure control or protection of confidentiality. However, the source data do contain some personal sensitive fields, for example post codes, registration numbers and breath test results. All such data are held and transferred securely and they are not available in published outputs. Sensitive data are made available to external users in some circumstances, for example to support local authorities' road safety initiatives. In these cases, users conform through signed GDPR-UK data access agreements that they will meet the necessary data security requirements and destroy the data at a specified time in the future.

## **Assessment of statistical quality**

Police recorded road collision statistics adhere to the Welsh Government's Statistical Quality Management Strategy, and this is in line with the European Statistical System's dimensions of quality. Details of the dimensions, and how we adhere to them, are provided below:

### **Relevance**

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

A variety of organisations use the Welsh road traffic collision and casualty data. The Welsh Government uses road traffic collision and casualty data to help set road safety policy. It is also used for performance indicators, both for the Welsh Government's Transport Strategy and for some Health Performance indicators.

Other users include Highway Authorities, including the Welsh Government which is responsible for the motorway and trunk road network, and local authorities, which are responsible for other roads in Wales. Other bodies involved in road safety include Go Safe, Trunk Road Agents, Police &



Community Safety Partnerships, Road Safety Wales and the Royal Society for the Prevention of Collisions (ROSPA). Welsh Government also provides data to transport planning organisations to support road safety assessments.

There is some demand for information on non-injury road collisions, but these are not covered by the STATS19 process.

## **Accuracy**

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

As discussed in the administrative data quality assurance section, there are number of known issues relating to accuracy, such as an incomplete record of all incidents, and an underreporting in serious collisions.

In addition, changes in police recording practices may mean that the statistics are not directly comparable over time.

The figures are based on information available to the Welsh Government around 14 weeks after the end of the latest quarter. Published figures are subject to revisions if additional data subsequently become available.

There are likely to be some errors with collision details and geographical and demographic information, though these are minimised by the validations carried out by the Welsh Government, local authorities and police forces.

Ahead of the 2023 police recorded road collisions release, we undertook additional validation of the road speed limit data as recorded by police officers at the scene of the collision. This additional validation was undertaken due to the change in the default speed limit on unrestricted roads on 17 September 2023, and wanting to ensure that the speed limit data had been correctly recorded.

This highlighted several inconsistencies, particularly relating to collisions where the attending police officer had recorded the speed the road where the collision occurred as 30mph, but information provided by local authorities suggested the road was 20mph at the time of the collision.

As a result of this additional validation, around 120 collisions that occurred between 17 September and 31 December 2023 that were recorded by the attending police officer as occurring on either a 20mph or 30mph road were found to have the wrong speed limit recorded, and have manually been corrected following confirmation from local authorities. This represents around 25% of all collisions on either a 20mph or 30mph road between 17 September and 31 December 2023.

This additional validation also highlighted some historical inconsistencies together with potential reasons for the inconsistencies. These will not be amended but limitations will be communicated within publications.

Findings based on a small unrepresentative sample of cases suggest that the road speed limit recorded at the time of the collision is consistent with local authority held information in roughly two-thirds of cases. For cases where the information is inconsistent, this is likely to be due to:

- Police officers recording the observed speed limit as it was at the time of the collision rather than the permanent speed limit (which may differ due to temporary orders, roadworks, diversions or other causes).
- Inconsistencies between LA records and road limit signs 'on the ground'
- Errors in recalling the road speed limit (based on own knowledge) when completing STATS19 forms back in an office environment.

Based on the small sample, the road speed limit information is more likely to be consistent with LA data for 20, 30 and 60mph roads (where over 80% of collisions take place) than for 40, 50 and 70mph roads.

This issue affects the distribution of road collisions and casualties by speed limit. It does not affect the total number of collisions or casualties at a Wales, police force or local authority area. It does not affect the overall data on collision and casualty severity.

We do not propose to amend any other historic data, due to uncertainty about the quality of historic speed limit data and the scale of manual validation required. In addition, the speed limits where most collisions occur are typically more likely to be consistent with the speed limit recorded by LAs. We will communicate the quality limitations within our publications.

## **Timeliness and punctuality**

Timeliness refers to the lapse of time between the reference period and the publication of the statistics. Punctuality refers to the time lag between the actual and planned dates of publication.

Full calendar year estimates are usually published around 5 to 6 months into the following year.

This release includes the statistical bulletin, interactive dashboard, and StatsWales tables.

Provisional updates for quarters 1, 2 and 3 are also made to the dashboard and StatsWales during the year with a lag of approximately 6 months after the end of the reference period.

Delays with data provision and quality assurance have, on occasion, resulted in the postponement of the main statistical release.

Results for Great Britain are usually published by the Department for Transport 8 months after the end of the reference period .

## **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.

This Statistical Bulletin is pre-announced and published on the [Welsh Government's Statistics & Research website](#). All historical data are available on [StatsWales](#) and users are able to download and export bespoke cuts of data in open data formats.

We aim to inform known users about developments with the statistical outputs. For example we developed a user friendly interactive dashboard with which users can explore a range of location and demographic features of the data. We update and promote the dashboard on a quarterly basis. In our outputs, we aim to provide a balance of commentary, summary tables, charts and maps where relevant. The intention is to 'tell the story' in plain English, with a focus on certain topics of known interest to users.

Our statistical headlines and tweets are published in Welsh and English. All of our outputs adhere to the Welsh Government accessibility policy and they are subject to internal peer review.

Further information regarding the statistics can be obtained by contacting the responsible statistician named on the release or via [stats.transport@gov.wales](mailto:stats.transport@gov.wales).

## **Comparability and coherence**

The degree to which data can be compared over time and domain and the degree to which data that are derived from different sources or methods, but which refer to the same phenomenon, are similar.

Changes in police recording practices and our quality assurance procedures may mean that the statistics are not directly comparable over time. These can come about through changes to systems, protocols, software or personnel. When we are aware of such cases we try to establish and describe the likely impact on the statistics. There may also be other more gradual recording or behavioural changes which we do not become aware of. It is not possible to quantify or adjust for the impact of such things but we believe when they occur they are unlikely to be significant because, for example, comparisons against other sources (hospital and survey data) are relatively stable.

Conceptually, road collision statistics are fully comparable across Great Britain. However, there are some known differences in reporting systems used by police forces which are likely to result in variation in the comparability of certain information. For example, some police forces in England and Wales have moved to injury based reporting systems which automatically assign severity, and these systems are known to result in higher numbers of severe injuries compared with non-injury based reporting systems, which are based on police officers' judgements. As with the differences between forces in Wales mentioned above, there may also be other procedural differences between police forces across Great Britain which would mean differences in coverage and quality of data.

As discussed, evidence from other sources indicates that the police recorded road collisions data are only a subset of all personal injury collisions, with potentially large numbers of less serious incidents not being reported to, or recorded by, the police. This, and the fact that the STATS19 process excludes certain collisions (for example deliberate acts and collisions not on public highways) means that the STATS19 statistics are not coherent with estimates of road collisions or casualties from sources such as hospital episode statistics and insurance claims data.

Evidence indicates the inconsistencies between STATS19 data and other data sources covering road collisions mainly affect less serious incidents. For example, there is broad coherence, both in terms of levels and in terms of trends, between the STATS19 statistics on fatalities and the Office for National Statistics' (ONS') cause of death statistics for land transport collisions.

The DfT led STATS19 review will investigate whether it is possible to quantify the overall scale of under-reporting and what further work can be done to provide a better assessment of the coherence in trends from these statistics and other sources.

## **Evaluation**

We welcome feedback on any of our statistics. If you wish to contact us, please do so via:

[stats.transport@gov.wales](mailto:stats.transport@gov.wales)

Produced by Knowledge and Analytical Services, Welsh Government

Last updated: June 2024

## **Open Government Licence**



All content is available under the [Open Government Licence v3.0](#), except where otherwise stated.