

National Data Set for Compensation-Based Statistics Edition 3.2

Explanatory notes

JULY 2025

Disclaimer

Safe Work Australia is an Australian Government statutory agency established in 2009. Safe Work Australia includes Members from the Commonwealth, and each state and territory, Members representing the interests of workers and Members representing the interests of employers.

Safe Work Australia works with the Commonwealth, state and territory governments to improve work health and safety and workers' compensation arrangements. Safe Work Australia is a national policy body, not a regulator of work health and safety. The Commonwealth, states and territories have responsibility for regulating and enforcing work health and safety laws in their jurisdiction.

Creative Commons

With the exception of the Safe Work Australia logo, this copyright work is licensed under a Creative Commons Attribution 4.0 International licence. To view a copy of this licence, visit creativecommons.org/licenses In essence, you are free to copy, communicate and adapt the work, even commercially, as long as you attribute the work to Safe Work Australia and abide by the other licence terms.

Contact information

Safe Work Australia | <mailto:info@swa.gov.au> | www.swa.gov.au

Explanatory Notes

National Data Set for Compensation-Based Statistics Edition 3.2

CONTENTS

1. Purpose	4
2. Glossary	5
3. Introduction.....	6
4. Objectives and use	7
4.1 Objectives	7
4.2 Analysis	7
4.3 Calculation of Measurement Rates	7
4.4 Jurisdictional Comparisons	8
4.5 Time Series Analysis.....	9
4.6 Industry and Occupational Analysis	9
4.7 Analysis of Causal Factors.....	9
4.8 Dissemination	9
5. Scope and coverage	10
6. Data compilation overview	11
7. Reporting and interpretation.....	12
7.1 Data suppression	12
7.2 Rates calculation.....	12
8. Data characteristics	13
8.1 Age of claimant	13
8.2 Time lost from work.....	13
8.3 Compensation paid	13
8.4 Time-series analyses	14
9. Data quality	15
9.1 Industry classification	15
9.2 Occupation classification.....	15
9.3 Details of injuries and diseases	15
9.4 Insufficiently coded data.....	16
9.5 Reliability of data.....	16
9.6 Other known issues.....	16

1. Purpose

Safe Work Australia (SWA) is responsible for the annual compilation of the National Data Set for Compensation-based Statistics (NDS) using workers' compensation claims data submitted by the various work health and safety (WHS) and workers compensation regulatory authorities in Australia

This document provides background information on the history of the NDS and sets out the objective, scope and methodology used in compiling the current edition (3.2) of the dataset.

Further, this document provides transparency with regard to the key strengths and limitations users should be aware of when working with NDS data.

2. Glossary

Term/Abbreviation	Definition
ABS	Australian Bureau of Statistics
ANZSCO	Australian and New Zealand Standard Classification of Occupations
ANZSIC	Australian and New Zealand Standard Industrial Classification
NDS	National Data Set for Compensation-based Statistics
NOHSC	National Occupational Health and Safety Commission
OMIE	Owner-managers of incorporated enterprises
RSE	Relative Standard Error. The RSE is the Standard Error (SE) expressed as a proportion of the estimate value (the SE is the square root of the variance).
Serious Claim	An accepted workers' compensation claim that involves one or more weeks of time lost (see below) and excludes all fatalities, and all injuries experienced while travelling to or from work or while on a break away from the workplace
SWA	Safe Work Australia
TOOCS	Type of Occurrence Classification System - A suite of four classifications to code the way an injury occurred, comprising: the Nature, Bodily location, Mechanism, and the Agency of injury/disease classification.
WHS	Work Health & Safety
WRMC	Workplace Relations Ministers' Council

3. Introduction

The National Data Set for Compensation-based Statistics (NDS) lists a standard set of data items, concepts and definitions for inclusion in workers' compensation systems operating in Australia. Its primary purpose is to enable the production of national and nationally comparable workers' compensation-based data.

Since its first release in 1987, the NDS has been implemented in workers' compensation-based collections administered by state, territory and Australian government agencies. Workers' compensation systems offer the best data source available in Australia for facilitating improvements in WHS in terms of scope and coverage considerations, ease of collection and cost.

Safe Work Australia is responsible for the annual production of the NDS.

Various minor adjustments have been made to the dataset over time, including modernising systems and expanding categorical coding coverage, culminating in the current NDS edition 3.2, which encompasses readily comparable data from the 2008-09 financial year onwards.

Throughout its history, the production of the NDS has required considerable co-operation from government, union, industry and other expert representatives involved. SWA wishes to express its appreciation to all these bodies for their valuable ongoing contributions to the development of the NDS.

4. Objectives and use

4.1 Objectives

The overall objectives of the NDS are to:

- assist in the goal to reduce work-related fatalities, injuries, and illnesses by providing a uniform national and nationally comparable evidence base of workers' compensation claims data, and
- assist in workers' compensation scheme management.

4.2 Analysis

The NDS can be used to support analysis on a variety of WHS topics, including:

- calculation of measurement rates
- jurisdictional comparisons
- time series analysis
- industry and occupational analysis
- analysis of causal factors.

4.3 Calculation of Measurement Rates

The value of the NDS is enhanced by the availability of suitable denominator data to enable the production of incidence and frequency rates for work-related injuries and illnesses. The incidence rate is defined as the number of claims per thousand in-scope jobs; the frequency rate is the number of claims per million hours worked in those in-scope jobs. These rates enable more precise analysis of NDS data by removing the effect of fluctuating and differing employment levels or hours worked across industries, occupations and cohorts of workers. This is particularly necessary in time series analysis and jurisdictional comparisons.

The Australian Bureau of Statistics (ABS) has developed labour force denominator data suitable for use in calculating rates using NDS-based data. Put simply, NDS denominators represent the working population covered by workers' compensation schemes. The methodology used to estimate the labour force denominators is periodically reviewed by the ABS to take account of the impact of changing working arrangements (for example, increasing prevalence of independent contracting working relationships) on the denominator data.

4.4 Jurisdictional Comparisons

The calculation of measurement rates enables comparisons of WHS performance and scheme management between the jurisdictions. The data are used as high-level indicators for the purposes of comparative performance monitoring.

4.5 Time Series Analysis

Time series analysis will be used mainly to monitor WHS outcomes across various cohorts. Time series analysis of WHS statistics for a given cohort can help in the evaluation of specific WHS policies.

4.6 Industry and Occupational Analysis

Analysis of industry and occupation data enables:

- identification of high-risk (and emerging high-risk) groups
- jurisdictions and employers to compare their performance.

4.7 Analysis of Causal Factors

Analysis of the type of occurrence data items, that is, nature, bodily location, mechanism, breakdown agency and agency of injury or disease, helps in the identification of the nature and causes of the problem and enables research, resources and risk reduction strategies to be better targeted.

4.8 Dissemination

Safe Work Australia will disseminate the NDS data through the [Safe Work Australia interactive data website](#).

5. Scope and coverage

Safe Work Australia compiles national workers' compensation statistics using data obtained from workers' compensation authorities in each state, territory and the Commonwealth Government. The data is collated into the National Data Set for Compensation-Based Statistics (NDS), which is SWA's primary source of information on work-related injuries and diseases¹.

The current edition of the NDS (3.2) provides a useful and versatile source of data on work-related injuries and illnesses, with detailed time series data available from 2008-09 onwards.

The data is very detailed, with information on the claimant's age, sex, industry and occupation of employment, type of injury or disease, mechanism of the incident, time away from work and amount of compensation paid. This allows for analysis by a range of variables to look at the driving factors behind changes in the rate of claims, and the impacts these injuries and illnesses have on the individuals affected.

The data also has strong coverage of the workforce. Approximately 85% of filled jobs in Australia are covered by a workers' compensation scheme².

There are some limitations with this data, which are listed below.

- The NDS does not cover all instances of work-related injuries and diseases.
 - While state, territory and Commonwealth Government workers' compensation legislation provides coverage for most employees, some specific groups of workers are covered under separate legislation. For example, workers' compensation claims lodged by police in Western Australia and military personnel of the Australian Defence Forces are not included in the NDS. For further information please see the [Comparison of Workers' Compensation Arrangements in Australia and New Zealand \(2023\)](#).
 - Work-related injuries and diseases of self-employed workers (or independent contractors) are under-represented, as workers' compensation schemes do not generally cover these workers.
 - Australian Bureau of Statistics data from the *Work-related injuries* survey show that approximately one-third of workers who experience a work-related injury or illness made a claim for compensation in 2021-22.
- Diseases are under-represented because many diseases result from long-term exposure to agents or have long latency periods, which makes the link between the work-related disease and the workplace difficult to establish. Further, workers with specific conditions such as mesothelioma may receive compensation through other mechanisms to workers' compensation.

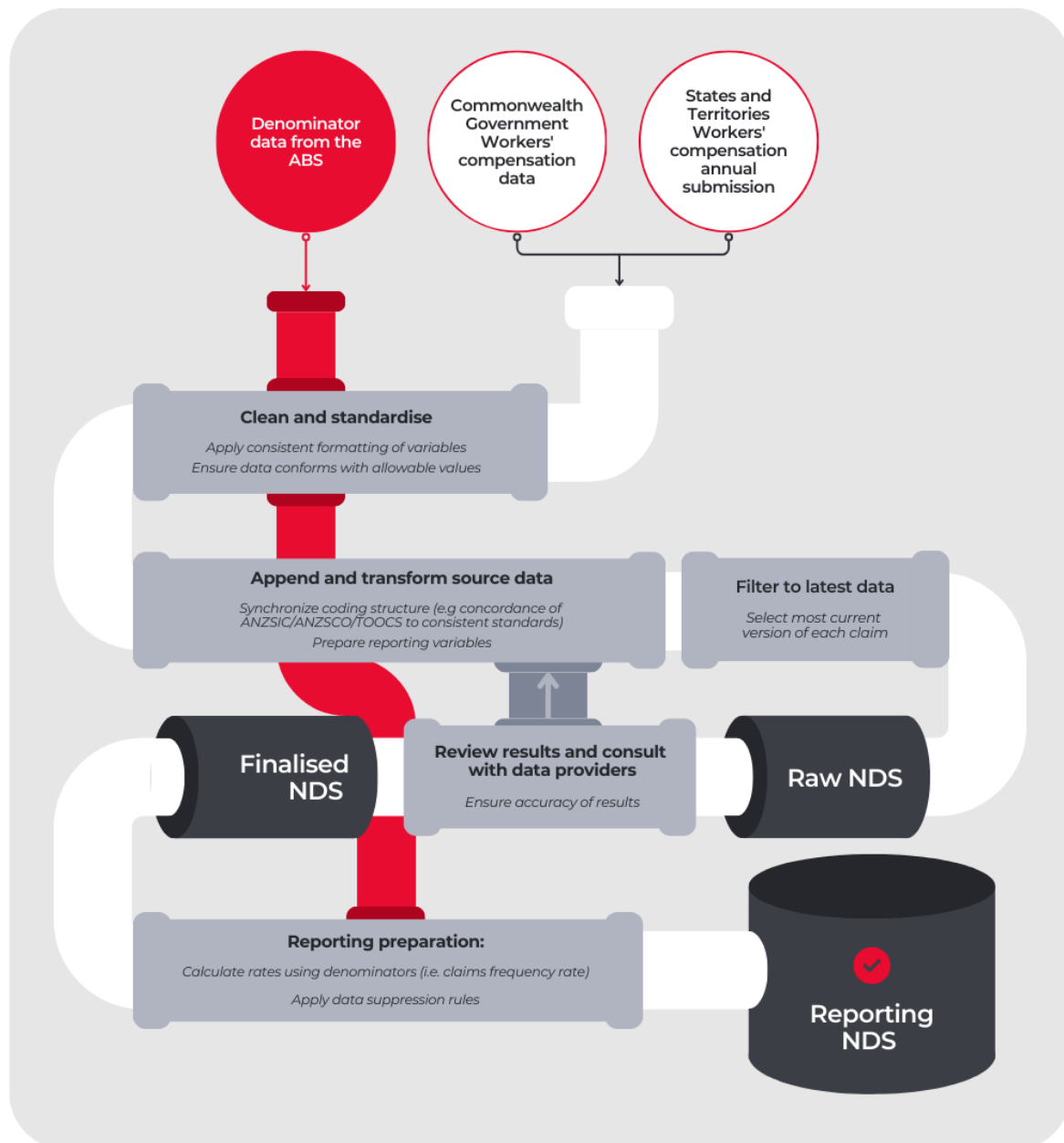
¹ NDS 3.2 Data specification sets out in detail the specific information sought from contributing jurisdictions for the purpose of compiling the NDS.

² Safe Work Australia analysis of Australian Bureau of Statistics, Labour Account, September 2023

6. Data compilation overview

To prepare compensation claims data received from contributing jurisdictions for reporting, SWA undertakes extensive data manipulation and transformation processes annually. These processes are summarised in Figure 1 below.

Figure 1: NDS data compilation process overview



7. Reporting and interpretation

Safe Work Australia focuses much of its analysis on “serious claims”. A serious claim is an accepted workers’ compensation claim that involves one or more weeks of working time lost (see below) and excludes all fatalities, and all injuries experienced while travelling to or from work or while on a break away from the workplace.

In general, claim reporting is presented by the financial year of lodgement of the claim. A financial year begins on 1 July and ends on 30 June.

7.1 Data suppression

Safe Work Australia collects de-identified data from each of the jurisdictions. In addition, to further protect the privacy of claimants, claim numbers are not reported for groups with less than 5 claims.

Injury rates are also suppressed where the Relative Standard Error (RSE) of the denominator is at least 50%, as this represents a significant amount of uncertainty in the estimated injury rate.

7.2 Rates calculation

Estimates of the number of jobs and hours worked for each Australian workers’ compensation jurisdiction are supplied annually by the ABS. The data are primarily derived from the Labour Force Survey, which are adjusted to account for differences in scope between the Labour Force Survey and workers’ compensation coverage.

The ABS provides two sets of estimates for each jurisdiction—one is split by sex, age and three-digit industry and the second is split by sex, age and four-digit occupation. This restricts presentation of rates to the categories supported by the ABS data. Therefore, it is not possible to calculate rates for occupational groups within specific industries.

Because eligibility for workers’ compensation varies from jurisdiction to jurisdiction, further adjustments are necessary. The most significant adjustments are outlined below.

- Police in Western Australia, who are covered by a separate scheme, are excluded from the denominators.
- Under the Queensland legislation, owner-managers of incorporated enterprises (OMIEs), who are included in the standard definition of ‘employee’, have the option of purchasing workers’ compensation insurance for themselves. This population has been excluded.

In 2021-22, the Australian Bureau of Statistics (ABS) undertook a review of the methodology and data sources used for producing these denominators. A full analysis of the impact of the update on injury frequency rates following this review can be found in the [Denominator Impact Analysis report \(2023\)](#).

Injury rates are not available for the Seacare jurisdiction because the denominators are not supplied from the ABS.

8. Data characteristics

8.1 Age of claimant

The age of an employee is derived from their date of birth and the date on which the injury occurred, or the disease was first reported to the employer. Data by age is typically reported in groups: 15 to 24 years, 25 to 34 years, 35 to 44 years, 45 to 54 years, 55 to 64 years, and 65 years and older.

The '65 years and older' group includes claims from people who have retired (these people may have developed a work-related disease and made a claim years after retirement). Therefore, claims rates (which express claims relative to the size of the workforce) for those aged 65 years and over should be treated with caution.

8.2 Time lost from work

Time lost from work is measured in working weeks and excludes estimates of future absences. Time lost from work comprises the total period for which *weekly benefits* were paid. The reported time lost is not necessarily continuous and may occur over several separate periods. Where an employee returns to work on a part-time basis, they may continue to receive pro-rata weekly benefit payments in addition to the income they would ordinarily receive for the hours worked.

The median is used as the measure of central tendency because long-term claims that involve lengthy periods of time lost from work would significantly skew an arithmetic mean. Data from the preliminary year are excluded when reporting median time lost from work because some claims from the preliminary year are likely to be open and claimants may accrue more time off work in subsequent years.

The time lost data are updated for five years after the initial reporting period and records are more likely to be finalised within this period. However, the reported time lost may be lower than the actual compensation paid due to the claims that accrue additional time lost beyond the five years for which the record is updated.

8.3 Compensation paid

The "Total compensation paid" variable typically used for reporting is made up of distinct types of compensation specified by the different schemes. The different types of compensation include weekly benefits which covers the wages or salary lost because of the worker taking time off (time lost), medical and hospital expenses, and legal fees and compensation for negligence. For further information of what is included in the total compensation paid, see the [NDS 3.2 Data specification](#). Every claim is unique and so even similar injuries can have different circumstances and consequentially different compensation payment structures.

The "Median compensation paid" is used as the measure of central tendency because claims with large compensation amounts would significantly skew an arithmetic mean. The median calculation is based on non-zero amounts. Data from the preliminary year are excluded when reporting compensation payments because some claims from the preliminary year are likely to be open and claimants may accrue more compensation payments in subsequent years.

The compensation payments data are updated for five years after the initial reporting period and records are more likely to be finalised within this period. However, the reported compensation paid may be lower than the actual compensation paid due to the claims that

accrue additional compensation payments beyond the five years for which the record is updated.

8.4 Time-series analyses

Each year, when an additional year of data is added to the NDS, the prior five years of data are revised. Beyond this revision period, NDS records are not further updated. While the source data does not indicate whether a claim has been finalised, records are more likely to be finalised by the end of the revision period. Time lost and compensation paid are the main variables that are affected by revision, although any field could be updated, or claims can be added as more information is submitted to various jurisdictions.

When analysing trends over time, consideration needs to be given to legislative changes that may influence trends in workers' compensation data. Information on workers' compensation arrangements can be found in Safe Work Australia's [Comparison of Workers' Compensation Arrangements in Australia and New Zealand \(2023\)](#).

The 'Median compensation paid' and 'time lost from work' figures for 2021-22 were markedly lower in aggregate due to a large number of COVID-19 serious claims. These claims generally had lower compensation payments and time lost, which brought the median value down. This effect was most pronounced in NSW as this jurisdiction had the broadest presumptive occupational rights for COVID-19 compensation payments.

9. Data quality

9.1 Industry classification

Information about the industry of employment of the claimant is coded using the [Australian and New Zealand Standard Industrial Classification \(ANZSIC, 2006 revision 2.0\)](#).

For reporting, this is based on the industry of the business that employs the claimant (NDS data item C1). Hence, a claim made by a person employed under labour hire arrangements is coded to the Labour supply services industry class, which is in the Administrative and support services industry. Industry of employer will be different to industry of workplace (NDS data item D2) for some claims.

Industry source data from NT, and for NSW, QLD, WA, SA, TAS and ACT Private prior to 2013, is recorded using [ANZSIC, 1993](#). This coding has been converted to the updated 2006 classification for inclusion in the NDS using a probabilistic distribution methodology so that results are comparable. As a result, there may be a degree of error in industry coding where alignment of the 1993 and 2006 industry structures is poor.

9.2 Occupation classification

Information about the occupation of the claimant is coded using the [Australian and New Zealand Standard Classification of Occupations \(ANZSCO, 2013, Version 1.3\)](#).

For some jurisdictions' and previous lodgement periods occupation source data is recorded using previous versions of this coding structure including the [Australian Standard Classification of Occupations, Fourth Edition 1998](#). This coding has been converted to the ANZSCO classification system using a probabilistic distribution methodology. Hence, there may be a degree of error in occupation coding where alignment of the ASCO and ANZSCO structures is poor.

9.3 Details of injuries and diseases

Information about injuries and diseases of claimants is coded using the [Type of Occurrence Classification System \(TOOCS\) edition 3.2](#). The classification system is used to code the:

- nature of injury or disease
- bodily location of injury or disease
- mechanism of injury or disease
- breakdown agency of injury or disease
- agency of injury or disease.

Some jurisdictions' injury and disease source data is coded using earlier editions of the TOOCS or separate coding structures. For inclusion in the NDS, this coding has been converted to the TOOCS 3.2 classification system using a probabilistic distribution methodology. Hence, there may be a degree of error in coding injuries and diseases where alignment of the source data structure and TOOCS 3.2 is poor.

9.4 Insufficiently coded data

Some claims are not fully coded due to insufficient information being provided at the time of the claim. Where there is insufficient information, they may be coded to residual categories like 'other and unspecified' or 'not elsewhere classified'. These claims are included when totals are calculated.

9.5 Reliability of data

Data are subject to two types of errors—non-sampling errors and sampling errors.

Non-sampling errors

Non-sampling errors may occur in any statistical collection due to:

- incorrect inclusion or exclusion of respondents or cases
- non-response of respondents
- inaccurate information from respondents
- inaccurate recording of information by data collectors
- deficiencies in data collection materials and processes
- errors that occur during the entry, coding and editing of data.

Non-sampling errors may affect the numerator and denominator data. It is difficult to quantify non-sampling error.

Sampling errors

Sampling error is a measure of the variation that occurs when a sample, rather than an entire population, is surveyed. Standard error indicates the extent to which an estimate of a sample varies from the estimate of the population from which it is drawn. When a standard error is expressed as a percentage of the estimate to which it relates, it is known as a relative standard error (RSE). The denominator data used to calculate rates are subject to sampling error. Rates calculated using denominators with an RSE of 50% or more have been automatically suppressed as they are considered unreliable estimates for general use. Caution should also be taken for figures with an RSE of 25% or more.

9.6 Other known issues

NSW “Other Miscellaneous Labourers”

In the 2021-22 and 2022-23 data, an increase in the number of claims for “Other Miscellaneous Labourers” has occurred in NSW due to a default system coding rule for non-specified occupations being applied. Safe Work Australia is exploring options for future reporting and revision of the data to account for this issue.

Mental Stress Claims

Mental stress claims (mechanism 8) in TOOCS have been updated recently. Similar to the other classification systems, the adaptation to the latest version is not uniform across the different jurisdictions. For example, Victoria coded all mental stress claims only to one digit mechanism up until recently, which means that the data cannot be broken down to a two-digit mechanism in this jurisdiction, and as a result have been coded in “not further defined”. In 2022-23, Victoria began reporting using two-digit mental stress claims.

The remaining jurisdictions have used the two-digit classification for all years of data. However, the TOOCS classification system was updated in 2021 to include sexual harassment as a separate code. This two-digit code has been utilised by the jurisdictions from different periods. As a result, caution should be taken when interpreting trends or analysis that utilise two-digit mechanism of mental stress claims.