

Comparison Between OSHA ITA Data and BLS SOII Estimates

Background

The OSHA Injury Tracking Application (ITA) and the Bureau of Labor Statistics (BLS) Survey of Occupational Injuries and Illnesses (SOII) both provide information based on worker injuries and illnesses that were recorded on OSHA recordkeeping forms. While both data sources provide useful information about workplace injuries and illnesses, meaningful differences exist in these collections, including scope of coverage and methods; data collection processes; data cleaning, coding, and review procedures; data availability; and data uses.

Scope of Coverage and Methods

The scope of coverage varies between the OSHA ITA and BLS SOII collections. Some employers are included in both the OSHA ITA and the BLS SOII and some employers are only included in one dataset or the other.

OSHA's recordkeeping regulation at 29 CFR 1904 applies to all employers covered by the Occupational Safety and Health Act of 1970 (OSH Act), although some employers with ten or fewer employees or with establishments in certain low hazard industries are partially exempt from keeping injury and illness records. OSHA's recordkeeping regulation also provides that, upon written notification, employers who are partially exempt from recordkeeping must keep injury and illness records and submit their data to OSHA or to the BLS. Finally, OSHA's recordkeeping regulation requires establishments that meet certain size and industry criteria to electronically submit data to OSHA through the ITA annually, and more establishments are required to submit summary data than are required to submit detailed case data.

Most states with OSHA-approved State Plans have ITA requirements identical to OSHA. However, OSHA does not cover state and local government workers. When state and local government workers are covered by a State Plan, that State Plan dictates which establishments must submit recordkeeping data. State and local government workers not covered by a State Plan would not be required to submit data through the ITA. Thus, state and local government workers are not universally included in ITA data.

The completeness and accuracy of the ITA data is limited by how well employers record injuries and illnesses under the regulation, mediated by OSHA's compliance assistance and enforcement efforts. In 2024, OSHA collected summary data from over 370,000 establishments and details for over 850,000 work-related injuries and illnesses from approximately 90,000 establishments. However, OSHA does not collect ITA data through a scientifically selected probability sample. Therefore, OSHA does not calculate measures of reliability (such as relative standard errors, RSEs) and does not perform weighting or imputation to account for incomplete reporting. ITA data are not representative of the total worker population, given that only select establishments are covered by ITA reporting requirements.

BLS uses widely accepted statistical methods to select a sample of establishments each year to participate in the SOII. OSHA's recordkeeping regulations require employers to report injury and illness data to BLS when requested to do so for the SOII. The SOII covers private sector and state and local government workers in all states and in all industries, except for self-employed workers, agricultural establishments with 10 or fewer workers, private households, and postal workers, and includes in its estimates data for the rail and mining industries obtained from the Federal Railroad Administration (FRA) and Mine Safety and Health Administration (MSHA), respectively. SOII scope is not limited by establishment size (except small farms) or OSHA exemption status.

The BLS SOII uses a two-stage sampling process to reduce reporting burden for respondents while ensuring the ability to produce statistically valid estimates. First, BLS uses a frame of known business establishments to select a nationally representative sample across all in-scope industries to participate in the SOII. Second, for establishments responding to the SOII that have many cases, BLS further reduces respondent reporting burden by directing respondents to report details for a statistically random sample of cases. By selecting a sample from a known frame of all in-scope business establishments, BLS can produce statistically valid estimates by detailed industry that represent all in-scope establishments. Case-level statistics are similarly produced using details from a subsample of known cases involving days away from work and cases of job transfer or restriction. SOII generates statistically valid and nationally representative estimates for all in-scope industries from a sample of approximately 230,000 establishments. SOII estimates are available for the nation and for most individual states by detailed industry and for case circumstances (nature, event, source/secondary source of injury/illness and part of body affected) and worker demographics (occupation, age, gender, race and ethnicity). Because SOII estimates are based on a scientifically selected probability sample, relative standard errors (RSEs) are calculated to provide a measure of precision that can be used to make valid statistical comparisons among estimates.

Data Collection Processes

OSHA collects data through the ITA. OSHA reminds employers about their obligation to submit their ITA data through announcements on the OSHA website and other outreach methods. Data can be entered into the ITA in one of three ways: using an online data entry form, by uploading a Comma Separated Values (CSV) file, or through a direct connection to the database using an Application Programming Interface (API).

BLS notifies all sampled establishments by mail and/or email prior to the calendar year for which they are required to keep injury and illness records for the SOII. Following that calendar year, BLS again contacts sampled establishments by mail and/or email and asks to report their injury and illness data. Employers can meet their requirement to respond to SOII by submitting data through a web-based application, fax, telephone, or hard copy. BLS follows with employers who do not respond to the initial request to report data for the SOII and also with many respondents for which additional clarification and corrections may be needed to ensure the accuracy of reported data.

Data Cleaning, Coding, and Review

Most ITA data are published on the OSHA website with minimal review, revision, or correction. For instance, addresses are not standardized, industry codes are not verified, company names are not validated, and information is instead published as reported by establishments. OSHA also does not validate the employee or injury and illness counts reported by establishments. Thus, there may be some errors or inaccuracies in the public facing database for some establishments.

OSHA advised employers not to submit certain personally identifiable information through the ITA, such as names and SSNs. OSHA intends to use automated information technology and some manual review to detect and remove information that could reasonably be expected to identify an individual directly before publicly posting data fields containing descriptions of what the employee was doing before the incident occurred, what happened, what the injury or illness was, and what object or substance directly harmed the employee. OSHA also uses the NIOSH Industry and Occupation Computerized Coding System (NIOCCS) to assign occupation codes to some ITA data.

BLS, working from a known establishment frame, can identify potential reporting errors and follows back with respondents to verify data and to make corrections where necessary. For example, reported

employment for an establishment that deviates considerably from the known expected employment may result in follow up with the respondent to verify that reported data are for the sampled establishment and to make corrections if necessary. Additionally, BLS use systematic checks combined with manual review to identify potential reporting discrepancies among summary and case details that may result in further contact with respondents to correct reporting errors. BLS reviews and revises SOII data and codes occupation (using the Standard Occupational Classification system, SOC), worker demographics, and the characteristics reflecting the circumstances of the workplace injury or illness using the Occupational Injury and Illness Classification System (OIICS). After SOII data have been reviewed, cleaned, and coded, BLS produces summary (industry-level) and case and demographic estimates. BLS uses systematic checks with manual review to further verify estimates prior to publication to ensure data quality meets expectations. For example, estimates with high relative standard errors (RSEs) may result in further review of underlying data and additional follow up to verify data with respondents. Ultimately, BLS may decide to not publish some SOII estimates that may be unreliable (e.g., high RSEs or data discrepancies that cannot be resolved).

Data Availability

ITA data is collected and published by both establishment and by incident. Variation may exist in how establishments record injuries and report their industry and company structure, which can sometimes make it difficult to compare across companies within an industry. Unlike SOII, OSHA publishes establishment information including company name and address, total hours worked, and total numbers of both injuries and illnesses by outcome (e.g., days away from work, job restriction or transfer) and category (respiratory disorder, eye injury, poisoning, skin condition, other illness, and injury,) for each establishment that submits data. OSHA also publishes injury and illness case details including the job of the worker, incident outcome, date of incident, number of days away from work, number of days in job transfer or restricted duty, the category of incident, the time the worker started work, and the time of incident. As noted above, OSHA intends to publish additional information from the Forms 300 Log and Form 301 Incident Report describing details about the injury or illness after OSHA has removed certain personally identifiable information (e.g., names, SSNs) that could reasonably be used to identify an individual.

Establishment-level data collected in the SOII are not publicly available. SOII estimates produced from collected data are available for the nation and for many individual states by detailed industry and occupation and for many case circumstances and demographic characteristics. Establishment level SOII data are protected under the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) and may not be published in identifiable form without written informed consent of the employer. In addition to fields from the OSHA forms, the SOII collects data on race and ethnicity that is not recorded on OSHA recordkeeping forms nor collected by ITA.

Data Uses

ITA data gives OSHA access to establishment-specific, case-specific injury and illness data that helps the agency and public identify establishments with specific hazards. This enables the agency to interact directly with these establishments to address hazards and improve worker safety and health. Public access to establishment-specific, case-specific ITA data allows employers, employees, potential employees, employee representatives, customers, potential customers, and the public to make more informed decisions about workplace safety and health at a given establishment. In addition, researchers are better able to identify patterns of injuries, illnesses, and hazardous conditions in workplaces. OSHA believes this access will ultimately result in the reduction of occupational injuries and illnesses.

ITA data includes all fields from the 300A Annual Summary, and most fields from the 300 Log and 301 Incident Report. By making most of the data it collects through ITA public, OSHA hopes to allow employers, employees, employee representatives, and researchers to identify and mitigate workplace hazards and thereby prevent worker injuries and illnesses by providing access to establishment and incident specific information.

The BLS SOII is the most comprehensive statistical survey on workplace safety and health in the U.S. Following data cleaning, coding, and review described above and using proven statistical methodologies to weight reported data to represent all in-scope establishments, BLS uses employer-reported hours data as the denominator for calculating SOII industry rates, and data from the BLS Current Population Survey (CPS) and Occupational Employment and Wage Statistics (OEWS) programs to calculate rates for occupation and worker demographics. SOII estimates of nonfatal workplace injuries and illnesses are available for the nation and for most states. These statistics are available by detailed industry and case type and by case circumstance and worker characteristics for cases that involved days away from work or job restriction or transfer. SOII statistics include relative standard errors that allow users to make statistically valid comparisons. See the [SOII Handbook of Methods](#) for additional details.