

# **SOII Research on Data Collection from Employees Literature Review**

## **Authors**

Cynthia Helba, Ph.D.  
Mariel Leonard  
Matthew Bernstein



**January 15, 2015**

Submitted to:  
U.S. Bureau of Labor Statistics  
Postal Square Building  
2 Massachusetts Avenue, NE  
Washington, DC 20212-0001

Prepared by:  
Westat  
*An Employee-Owned Research Corporation®*  
1600 Research Boulevard  
Rockville, Maryland 20850-3129  
(301) 251-1500

## Table of Contents

<u>Chapter</u>		<u>Page</u>
1	Introduction .....	1
2	Research Questions .....	2
	2.1 How Will the Employees Be Located and Selected to Participate? .....	2
	2.2 What Characteristics of the Employees to Be Surveyed Are Relevant to the Project?.....	4
	2.3 How Will the Survey Be Designed to Collect the Appropriate Data?.....	6
3	Conclusion.....	7
 <b><u>Appendix</u></b>		
A	Literature Summaries .....	A-1
	1 How Will the Employees Be Located and Selected to Participate? .....	A-1
	Module on Large Existing, Large Household Surveys .....	A-1
	New Household Study with Specialized Sampling for Rare Events .....	A-4
	Special Population Study – e.g., Union Study, Specialized Industry Studies .....	A-8
	Linked Employer and Employee Study.....	A-10
	2 What Characteristics of the Employees to Be Surveyed Are Relevant to the Project?.....	A-14
	All “In Scope” Workplaces.....	A-14
	“In Scope” Workers with Vulnerable Status.....	A-15
	Injury/Illness Status May Be Sensitive .....	A-19
	3 How Will the Survey Be Designed to Collect the Appropriate Data?.....	A-20
	Ensure Understanding of Reportable Events.....	A-20

<b><u>Appendix</u></b>		<b><u>Page</u></b>
	Ensure Questions Cue Accurate Reporting of Injuries and Illnesses.....	A-21
	Ensure Appropriate Data Collection Methods.....	A-25
B	Additional Relevant Studies .....	B-1
C	Literature Reviewed but Not Deemed Directly Relevant.....	C-1

The U.S. Bureau of Labor Statistics (BLS) Survey of Occupational Injuries and Illnesses (SOII) is a primary source of information on nonfatal injuries and illnesses that take place in the workplace in the United States. The SOII is collected yearly from a sample of employers who report information from their Occupational Safety and Health Administration logs and other materials. Since early in its inception, there has been concern about the completeness of the reporting of injury and illness using this method.<sup>1</sup> These concerns continue with many researchers, including those funded by BLS, examining the extent to which injuries and illnesses are undercounted.<sup>2</sup>

BLS has tasked Westat to explore the possibility of collecting information on workplace injuries and illness from employees rather than employers. Considering this avenue of collection is important because one of the “filters” or inhibitors to adequate collection of these injury and illness workplace data is employee reluctance to report health problems to supervisors because they may risk disciplinary action, stigmatization, harassment, or job loss. Alternatively, some safety reward systems actually reward employees who do not report injuries.<sup>3</sup> Employees may provide more complete information when not presented with actual or anticipated coercion from their employer.

This document is the result of the first of three activities that form this task, a review of the literature related to surveying employees about workplace injuries and illnesses. Guiding Westat’s review was one overarching requirement—identifying the best way to collect comparable data to the SOII. In addition, the review was organized by three research questions:

1. How will the employees be located and selected to participate?
2. What characteristics of the employees to be surveyed are relevant to the project?
3. How will the survey be designed to collect the appropriate data?

---

<sup>1</sup> Pollack, E. S., and Keimig, D. G. (1987). Counting injuries and illnesses in the workplace: proposals for a better system. National Academies.

<sup>2</sup> Wiatrowski, W. J. (2014), The BLS survey of occupational injuries and illnesses: A primer. *American Journal of Industrial Medicine*. doi: 10.1002/ajim.22312.

<sup>3</sup> Azaroff, L. S., Levenstein, C., and Wegman, D. H. (2002). Occupational injury and illness surveillance: conceptual filters explain underreporting. *American Journal of Public Health*, 92(9), 1421-1429.

The remainder of this report provides a summary of the themes that emerged as the literature was reviewed. Appendix A to this document provides brief summaries of the literature reviewed; Appendix B provides a list of additional surveys that may be assessed during the cost-benefit analysis; and Appendix C lists literature that was deemed slightly out of scope. A copy of any of this literature is available on request.

## Research Questions

# 2

The summary of themes that emerged from the literature is presented here by research question and by relevant topics.

### 2.1 How Will the Employees Be Located and Selected to Participate?

**Consider a module on existing large household surveys.** There are a number of large-scale surveys that could be used as a platform to study the general population of in-scope workers. All will require recall over at least a 1-year period. For example, the National Health Interview Survey (NHIS) is the one of most common sources of employee occupational injury and illness data besides workers compensation. NHIS data has included modules on employment and workplace injuries and illnesses. These data have been used frequently for secondary analyses. Other large-scale surveys considered in this review include the Behavioral Risk Factor Surveillance System, Current Population Survey, Medical Expenditure Panel Survey – Household Component, National Health and Nutrition Examination Survey, and National Longitudinal Survey of Youth.

**Consider a new household study with specialized sampling.** Workers with workplace injuries or illnesses are relatively rare in a general population of adults; therefore, a new study to assess prevalence of these injuries and illnesses would need a sampling design that would ensure an adequate number of respondents who experienced workplace injuries or illnesses. This review considered three strong presentations on the types of sampling that could provide an adequate number of injured or ill workers (Levy and Lemeshow, 2008; Kalton, 2009; Lohr, 2010). Although there was overlap in the presentations, each contributed some unique ideas about how rare

populations might be sampled. Some of the sampling methods these statisticians suggested are briefly described below. More detail can be found in Appendix A to this report.

- A dual sample or multiple-frame design of some type was suggested by all three of these discussions. One way this type of design can be used is when one frame has complete coverage but low prevalence (e.g., a population frame) and another has incomplete coverage but high prevalence (e.g., a register). By combining the sample estimates from the two frames, a population estimate can be made.
- Network sampling allows a sampled respondent to act as a proxy for others linked to them in a known way, such as family members. According to Lohr, if the issues under study are things the family members could reasonably have knowledge of; this method can reduce sampling variability of the prevalence estimates.
- Two-phase sampling is discussed by Kalton and Lohr. Two-phase sampling screens are used for the rare phenomena in the first phase followed by a second phase with a higher sampling fraction for those with the rare characteristic than those without.

Alternative sources of information for the sampling frame would strengthen the likelihood of finding individuals with workplace injuries and illnesses. Two studies were reviewed (Husberg, 2003; Husberg and Lincoln, 2006) that used the Alaska Trauma Registry to identify individuals with fishing injuries. The NIOSH Health and Safety Practices Survey of Health Care Workers (Steege, Boiano, and Sweeney, 2014) used professional association listings and Internet recruitment to ensure the collection of data concerning all the hazards included in the study. Kessler et al. (2012) used the membership of a health plan as the sampling frame of their study of injuries related to insomnia. Using a frame from health insurers would also have the benefit of linked administrative data that could be helpful in finding individuals with injuries.

**Consider special population study (e.g., union study, specialized industry studies).** Review of three union studies suggested that union members filed more claims or reported more injuries than nonunion workers. It was suggested that this might be because of union support for filing claims or employee support of union goals. While no strong literature describing unique illness and injury reporting characteristics of other industries was located during this initial review, what has been learned so far suggests that care should be taken to include both unionized and nonunionized employees in the survey and, it can be assumed, representatives of a variety of industries.

**Consider linked employers and employee surveys.** The review identified a number of examples of linked employer and employee surveys in which representatives of the employer establishment were first surveyed and then workers selected from that establishment were surveyed (Beatson, 2000; Greenen and Hamon-Cholet, 2000; Patek, Hidiroglou, and Lavellee, 2000; Probst and Estrada,

2010; Probst and Graso, 2013; Souza et al., 2014; Shuhui, Myers, and Layne, 2011). Each of these studies provided information on the kinds of methods that could be used to select and collect data from employees in a way related to their place of work. The Probst studies even suggest that data about injuries and failure to report them can be collected at the place of work itself as long as anonymity of the respondent is protected. This approach would provide access to the employees and, if desired, data from employing establishments. All of these studies raise a common concern. The size of the sample of employees is frequently small to allow for the collection of data from a broad spectrum of employing establishments. The likelihood of identifying workplace injuries would be limited without a shift to more employees and fewer establishments. The Medical Expenditure Panel Survey—Insurance Component might be a way to conduct a linked study using an existing vehicle for sampling establishments.

## **2.2 What Characteristics of the Employees to Be Surveyed Are Relevant to the Project?**

**Include employees from all “in scope” workplaces including those workplaces known to have a likelihood of undercount (e.g., multi-establishment firms, small firms).** One study of undercount (Boden et al., 2010) suggests that the SOII is relatively less likely to capture a case if it is in one of the following industries: transportation, communications, utilities, and wholesale or retail trade. Ensuring workers from industries or types of establishments known to be associated with the undercount are included is critical. Prior work on the undercount problem may need to be reviewed to develop the sampling plan or expert guidance secured.

Alternative methods for identifying employees from “in scope” workplaces were identified in the review. The 1995 BLS Survey of Employer-Provided Training (SEPT) (BLS, 1995) sampled from employee lists provided by participating employers. This approach focuses more on the population that may have workplace injuries or illnesses than does a general population survey thus making it easier to secure enough injured or ill workers to make a reasonable prevalence estimate. This approach, because it begins with the employer, would also help to ensure that workplaces known to have a strong probability of undercount are included.

The National Study of Organizations (Kalleberg et al., 1994) took an opposite approach to identifying employees and organizations. Using a special module on the General Social Survey, willing respondents provided information about their employers, creating a sample of organizations.

If a similar method were adopted for this study, such a sample of organizations could be supplemented if needed to ensure full coverage of employers but data would be readily available on the employees through a module on a national survey.

**Include all “in scope” workers especially vulnerable workers (e.g., non-English speaking, immigration status, right-to-work states).** Some unique characteristics of vulnerable workers will require attention in the design and administration of the survey. These include the following:

- Some vulnerable workers may not understand what is a reportable injury or illness (Forst et al., 2013). This can be tested during survey development and revisions made to the survey to ensure understanding without compromising comparability to the SOII.
- Some vulnerable workers may not understand vocabulary used to describe workplace injury or illness (Marucci-Wellman et al., 2013). Again, this can be tested during survey development.
- Special methods to collect from some vulnerable workers beyond written or audio translation might be useful in increasing the response rates (Pransky et al., 2002). This can be considered in a cost-benefit analysis.
- Some vulnerable workers in low-skill jobs may have low participation rates because of the difficulty of finding them; unwillingness to participate because of physical tiredness, fear of job loss, or fear of being involved with government agencies; and language barriers (Mobed et al., 1992). Ensuring that these more complex nonresponders are included will be a cost consideration for the study.
- Some vulnerable workers such as young workers may work part time making it difficult to determine if their injuries should be reported (e.g., they do not miss the next day’s work if they are not scheduled to work that day).
- Undocumented farm workers may differ in their knowledge of reportable injuries than those who are documented. Farm workers, in general, may not seek medical attention for any illness or injury (Villarejo et al., 2010).
- The New Immigrant Survey may be a way to capture information on a unique group of vulnerable workers if it is continued in future years.
- Low general and medical literacy among workers must be considered in creating the data collection materials.

To summarize, it may be possible to meet some of the identified needs of vulnerable workers through cognitive and pretesting during survey development. Other possible adaptations to ensure participation of vulnerable workers will be decisions primarily based on financial parameters.



**Recognize that those who have experienced workplace injuries and illnesses may consider these illnesses and injuries a sensitive topic.** Tourangeau and Yan (2007) provide a review of a large amount of the available literature, including meta-analyses, on surveys on sensitive topics. In addition in this document, they provide the results of three new meta-analyses conducted specifically for this publication. They conclude that when topics are considered sensitive by respondents they frequently misreport either because of embarrassment or fear of retribution should their response become known. Tourangeau and Yan provide evidence that several methodological factors can increase the likelihood of accurate reporting. Some of the factors they discuss are relevant to this future study. These factors include using a self-report technology, wording questions in a “forgiving” way, and adding assurances of confidentiality.

## **2.3 How Will the Survey Be Designed to Collect the Appropriate Data?**

**Ensure understanding of reportable events.** The literature identified for this area was mostly focused on why workers did not file workers’ compensation claims. It is included because some of these reasons might also hold for nonreporting and lack of understanding about what is reportable. Rosenman et al. (2000) provide a number of reasons that workers do not file claims that also provide insight into a possible lack of understanding about what is reportable. These reasons included injury not considered serious enough by the worker, the worker did not expect to miss work, and the worker expected to miss work but knew they would receive some form of sick leave or short-term disability from employer. This issue is also open to testing during survey development.

**Ensure questions cue accurate reporting of injuries and illnesses.** The review indicates that, in general, as time between injury and recall increases, reporting decreases (Landen and Hendricks, 1995; Warner et al., 2005). The evidence is mixed for the negative correlation between time and recall for very serious injuries (support by Landen and Hendricks and Warner et al., 2005; no support from Moshiro et al., 2005). Moshiro et al. (2005) report another problem with longer recall periods. They found that in the 12th month there was an increase in reported injuries compared to the 10th and 11th months. They speculated that this was due to telescoping of events from the past year.

With regard to ways to improve the quality of the data, Miller and Downes-Le Guin (1989) suggest some ways to improve the problem of memory decay that would fit a self-administered standard

survey including reducing the reference period and increasing the number of questions about events to allow more time for respondents to recall and form answers (also supported by Overpeck et al., no date). Smith et al. (2006) point out that the need for more careful definitions to ensure data quality. In their work with a module of the NHIS, the definition of “at work” was very complicated. For example, an injury in a bar may have been at work or during recreation. Volunteer firefighters may have not considered duty-related injuries work injuries even though they meet the standard definition of workplace injuries. The new study needs to provide cuing that will ensure that respondents understand the reporting rules so that the resulting data will be comparable to the SOII. See Wuellner and Bonauto (2014) for a comparison of workers’ compensation and SOII data that indicates differences in response based on cuing.

**Ensure appropriate data collection methods.** This review has noted a variety of interesting methods for identifying workers and workplaces. It has also noted some forms of data collection that while useful, might be cost prohibitive (e.g., house-to-house interviews). The review has also presented information supporting the use of self-administered data collection. In a final article, an interesting data collection method was presented that could be useful to consider. Huang et al. (2012) used 12-week diaries reported via telephone, Internet, or in writing. The use of interactive voice response or Internet data collection is becoming common in the survey field and should be considered for this survey.

## Conclusion

# 3

The basic survey designs identified in this review are a module on an existing national survey, a new study using a sampling strategy that focuses on estimating the prevalence of a rare event in the general population—workplace illnesses and injuries, a linked employer-employee survey, or perhaps, a set of industry-level surveys. Of course, designs that combine these basic features are also a possibility.

Beyond the basic characteristics possible for the final designs, the review has revealed that these designs must ensure coverage of industries where there is known undercount and vulnerable workers who may be harder to survey. The final designs must also address features of the administration and survey instrument that are known to be problematic, misreporting, lack of

understanding of what is reportable, and recall problems. These issues might be addressed through assurances of confidentiality, questions that cue the respondent to respond about any and all reportable injuries and illnesses, and addressing the recall difficulty issues with extra questions to set the timeframe or decreasing the reference period. The SOII is real-time data collection. It is clear that the data will collected for a 1-year period will not be comparable without a carefully crafted survey instrument that helps the respondent recall the reference period under question. It is also possible that a unique design may be needed that combines estimates for shorter reference periods to create a single estimate for the year.

The next step in this task is to conduct a cost-benefit analysis of the designs identified in this effort or facilitated by it. The cost-benefit analysis will extend beyond financial considerations. It will address, among other things, the coverage and recall issues described above. In the end, designs will be recommended to BLS for pretesting that achieve a desirable combination of cost, coverage, and data quality.

# Appendix A

## Literature Summaries

### 1. How Will the Employees Be Located and Selected to Participate?

#### *Module on Large Existing, Large Household Surveys*

Survey	<b>Behavioral Risk Factor Surveillance System (BRFSS)</b>
URL	<a href="http://www.cdc.gov/brfss">http://www.cdc.gov/brfss</a>
Summary Overview	This is an ongoing data collection beginning in 1984 conducted by the Centers for Disease Control and Prevention (CDC) and U.S. states and territories.
Method	Population: U.S. civilian noninstitutionalized adult (18+) population
	<p>Since 2011, BRFSS conducts both landline and cellphone surveys. For landline surveys, an adult in the household is randomly selected. The cellphone survey is used for respondents who receive 90 percent or more of their calls via cellphone and live in a private residence or college housing.</p> <p>Fifty-one states and territories use a disproportionate stratified sample design for landline samples. Telephone numbers are divided into high density and medium density strata which are expected to belong to households. BRFSS uses commercially available cellphone sampling frames.</p> <p>Each state implements BRFSS differently and may use different modules. Provides state-level estimates.</p>
Comments	All states use the core component of BRFSS that includes questions about demographics and current health-related perceptions, conditions, and behaviors.

Survey	<b>Current Population Survey (CPS)</b>
URL	<a href="http://www.census.gov/cps/">http://www.census.gov/cps/</a>
Author and Title	<b>Zbikowski, Andrew, and Antoinette Lubich.</b> "Design and Methodology: Current Population Survey." Technical Paper 66. US Department of Labor and US Department of Commerce. October 2006. Available online: <a href="http://www.census.gov/prod/2006pubs/tp-66.pdf">http://www.census.gov/prod/2006pubs/tp-66.pdf</a> .
Summary Overview	This is an ongoing data collection administered by the U.S. Census Bureau and supported the U.S. Bureau of Labor Statistics.
Method	Population: U.S. civilian noninstitutionalized population 15+ years
	<p>CPS uses a multistage stratified sample of approximately 72,000 assigned housing units from 824 sample areas.</p> <p>In the first stage of sampling, the United States is divided into primary sampling units (PSUs) encompassing a metropolitan area, a large county, or several smaller counties. All PSUs fall within the boundary of a state.</p> <p>In the second stage, a sample of housing units is drawn. Each month, interviewers collect data from sample housing units in in-person interviews. A housing unit is interviewed for 4 consecutive months, dropped out of the sample for the next 8 months, and interviewed again in the following 4 months. Overall, a housing unit is interviewed eight times in 16 months.</p>

	CPS collects demographic and labor force information on each member of the household. Labor force questions include employment status, hours worked, occupation, and industry.
Comments	The CPS collects data on self-employed, agricultural, and unpaid (as in a family business) workers, as well as the unemployed. CPS does not collect data on occupational injury or illness or other occupational health topics.

Survey	<b>Medical Expenditure Panel Survey – Household Component</b>
URL	<a href="http://meps.ahrq.gov/mepsweb/survey_comp/household.jsp">http://meps.ahrq.gov/mepsweb/survey_comp/household.jsp</a>
Summary Overview	MEPS has been conducted yearly since 1996 and is sponsored by the Agency for Healthcare Research and Quality.
Method	Population: U.S. civilian noninstitutionalized population
	Sample size: Approximately 15,000 households. For 2012, 14,763 households were sampled with a total of 37,182 individuals. For 2012, the response rate was 56.3 percent for the full year.
	The Household Component collects data from individuals within the household, which is supplemented by data from their medical providers. Each year a panel of sample households is selected from the previous year's National Health Interview Survey (NHIS) sample. Respondents participate in five rounds of interviews covering 2 full calendar years.
Comments	Relevant questions include demographics, health conditions, health status, use of medical services, access to care, income, and employment. Because individual responses are supplemented by medical providers, more technical and exact diagnoses may be available. Also, using a panel design reduces the length of the recall period, while still collecting 2 years' worth of data.

Survey	<b>National Health and Nutrition Examination Survey (NHANES)</b>
URL	<a href="http://www.cdc.gov/nchs/nhanes.htm">http://www.cdc.gov/nchs/nhanes.htm</a>
Summary Overview	NHANES has been conducted since the 1960s by the National Center for Health Statistics.
Method	Population: U.S. civilian noninstitutionalized population
	Approximately 12,000 individuals are asked to participate within a 2-year period. Approximately 10,000 participate in data collection.
	NHANES uses a multistage probability sampling design using counties (or groups of counties) as PSUs with blocks or neighborhoods as smaller units. Adults over 80 years, African Americans, Asians, Hispanics, and low-income whites are oversampled.  NHANES combines both in-person interviews and physical examination.  Each year of data collection creates nationally representative estimates; however, because only 15 PSUs are visited each year, data is released in 2-year cycles to reduce variance and protect respondent identity.
Comments	The interview includes demographic, socioeconomic, dietary, and health-related questions. Because each respondent is examined by a physician, more technical and exact diagnoses may be available.

Survey	<b>National Health Interview Survey (NHIS)</b>
URL	<a href="http://www.cdc.gov/nchs/nhis.htm">http://www.cdc.gov/nchs/nhis.htm</a>
Author and Title	<b>Luckhaupt, Sara E., and John P. Sestito.</b> "Examining National Trends in Worker Health with the National Health Interview Survey." <i>Journal of Occupational and Environmental Medicine</i> . Vol. 55, No. 12 Supplement. December 2013: S58-S62.
Summary Overview	This is an ongoing data collection conducted by the National Center for Health Statistics, supported by U.S. Census Bureau.
Method	Population: U.S. civilian noninstitutionalized population  NHIS uses a stratified multistage sample design. A new sample design is implemented after each decennial census. Since 1985, NHIS annual samples are subdivided into four panels, each considered a probability sample of the U.S. population. In this way, if sufficient funding to conduct all panels is not available, national estimates can still be produced. Additionally, NHIS assigns subsamples to 4 calendar quarters of 3 months, which allows the NHIS sample to obtain estimates for large population groups from a short period of data collection.  Data is collected in face-to-face interviews for each member of the household. If the interview cannot be completed in one visit, additional visits or telephone interviews are scheduled. In addition to the Household and Family Core components, one sample adult and one sample child are randomly selected for the Sample Adult and Sample Child components. Adults aged 65 and over who are African American, Hispanic, or Asian are oversampled.
Comments	The current design (2006-2015) is stratified by states, but is not designed to produce state-level estimates for every state.  NHIS is the one of most common sources of employee occupational injury and illness data besides workers compensation.  Data currently collected about employment and workplace injuries and illnesses have been used frequently for secondary analyses. Some examples are included in this review.

Survey	<b>National Longitudinal Survey of Youth (NLSY)</b>
URL	<a href="http://www.bls.gov/nls/">http://www.bls.gov/nls/</a>
Author and Title	<b>Dembe, A.E., J.B. Erickson, R.G. Delbos, and S.M. Banks.</b> "The Impact of Overtime and Long Work Hours on Occupational Injuries and Illnesses: New Evidence from the United States." <i>Occupational Environmental Medicine</i> . Vol. 62. 2005: 588-597.
Summary Overview	The authors reviewed data from the National Longitudinal Survey of Youth (NLSY).
Method	The NLSY is a longitudinal survey which followed individuals between 1987 and 2000. Followup interviews were conducted annually from 1979 to 1994, and biannually since 1996. Researchers made attempts to re-interview every remaining cohort member at each survey. Survey response rates ranged from 91.0 percent for the 1988 survey to a high of 92.5 percent in 1989 and a low of 83.4 percent in 2000 (excluding deceased respondents).  During the period of the survey, 10,793 respondents reported working in at least one job. The authors reviewed self-reported data on incidence of a work-related injury or illness by screening for an affirmative response to the following question:  "I would like to ask you a few questions about any injuries or illnesses you might have received or gotten while you were working on a job. Since [date of last interview] have you had an incident at any job that resulted in an injury or illness to you?"

Findings	The authors suggest that because the NLSY's primary objective is to evaluate participants' long term labor market transitions and wage history, it "avoids problems of information bias that typically plague attempts to ask injured workers about their working conditions and job exposures." The authors believe that unlike respondents to a survey specifically designed to collect data on occupational safety and health, "respondents to the NLSY will [not] intentionally or unintentionally be attempting to justify the legitimacy of a work-related disorder, establish its compensability under workers' compensation laws, or establish the employer's culpability for the injury. All of those issues are unrelated to the main concerns of NLSY and thus the data obtained presumably will be less susceptible to contamination by such considerations."
Comments	The conclusion drawn by these authors supports the placement of the SOII questions on an existing survey or on a new noninjury related survey.

### ***New Household Study with Specialized Sampling for Rare Events***

Author and Title	<b>Levy, Paul S., and Stanley Lemeshow.</b> "Chapter 14: Selected Topics in Sample Design and Estimation Methodology." <i>Sampling of Populations: Methods and Applications, Fourth Edition.</i> (Hoboken, NJ: Wiley, 2008).
Summary Overview	This chapter discusses a variety of different sampling and estimation techniques for special purposes including rare events.
Detailed Information	The relevant discussions in this chapter pertain to the study of rare events. Network and dual frame sampling are discussed. <ul style="list-style-type: none"> <li>■ <b>Network sampling</b> uses a modified counting rule. Typically, this type of sampling is used to gather information on the total number of cases of a rare disease without contacting all health care providers by allowing other health care providers who have treated the patient for other issues to report on the rare disease.</li> <li>■ <b>Dual-sample method</b> uses either two surveys or a survey and a registration system to better estimate a rare event. If the two data sources are independent, a population estimate of the prevalence of the rare event can be made based on events reported in both sources.</li> </ul>
Comments	Each of these sampling methods may offer appropriate methods of improving estimates of workplace illness and injury based on individual reports rather than workplace reports. Network sampling could be considered because it could possibly allow coworkers to report on illness and injuries. Dual samples could be a representative national sample and a sample of those who have reported workplace injury or illness. Using reports from these two independently selected samples would all and estimate of national levels of workplace illness and injury.

Author and Title	<b>Kalton, Graham.</b> "Sampling and Oversampling Rare Populations." <i>Bulletin of the International Statistical Institute, 57th Session, Durban, South Africa.</i> 2009: Available at: <a href="http://www.isi-web.org/publications-2/proceedings">http://www.isi-web.org/publications-2/proceedings</a> .
Summary Overview	This paper discusses a variety of different sampling and oversampling for rare subpopulations. While its focus is on methods for estimating the characteristics of members of rare domains, similar methods are often applicable for estimate the size of a rare population.
Detailed Information	This paper discusses seven types of sampling that might be used to sample and oversample rare populations. <ul style="list-style-type: none"> <li>■ <b>Screening</b> is useful when the relevant domains are easily identified with a few simple questions. One starts with an overall sample large enough to produce the required number of rare cases. Domain membership is determined for the whole</li> </ul>

	<p>sample and survey data are collected on subsamples of the entire sample. Costs for screening a large sample can be reduced using telephone, mail, IVR, or web. If using face-to-face methods, the use of compact clusters will cut cost. A form of network sampling, using neighbors as proxies can also cut costs as well as allowing more than one person in a household to be considered eligible. Risks in this approach included failure to achieve planned sample size of the rare domain, depressed response rate when rare domain members fail to respond to the screening or survey, noncoverage because of problems with the survey frame, and underrepresentation of the rare domain in the overall sample.</p> <ul style="list-style-type: none"> <li>■ <u>Disproportionate stratification</u> uses higher sampling fractions in strata where the prevalence of the rare population is higher. This approach will reduce the amount of screening needed. Risks to this approach include a loss in precision if the true prevalence rates of the rare population are mistaken for the strata.</li> <li>■ <u>Two-phase sampling</u> assumes that the accurate identification of the rare domain is expensive for some reason. The first phase is an inexpensive, imperfect screener. The second phase more carefully identifies the rare population. This form of sampling is often used when a rare medical condition is under study and only those with the condition are of interest.</li> <li>■ <u>Multiple frame designs</u> are used when a single sampling frame has incomplete coverage of those within the rare domain. A dual-frame design is often used when one frame has complete coverage and a low prevalence and another has incomplete coverage but high prevalence. One example of this would be an area frame and a register.</li> <li>■ <u>Multiplicity sampling designs</u> (or network designs) asks persons sampled for screening to serve as proxies for others who are linked to them in a known way. Family is often used as a basis of linkage. This allows one to identify members of the rare domain that are not in the original sampling frame. The informant must be able to provide contact information for the persons named. A risk in this form of sampling is that the proxy may not accurately report the domain status of the people he or she names.</li> <li>■ <u>Location sampling</u> is used for rare populations with no fixed location such as the homeless or for passengers at an airport, for example. Location sampling has been used to find very rare populations by sampling locations where this rare population congregates.</li> <li>■ <u>Accumulating or retaining samples</u> over time takes advantage of survey data collections that are repeated over time. If a weekly survey is accumulated over a year, there will be sufficient sample size of the rare population to do analysis. These provide period estimates rather than point-in-time estimates. A panel study can be used to identify members with non-static rare characteristics over time.</li> </ul>
Comments	<p>These types of samples all lend themselves to finding sufficient number of working people, working people in specific industries, or working people with workplace injuries or illnesses.</p>



Author and Title	<b>Lohr, Sharon.</b> “Chapter 14: Rare Populations and Small Area Estimation” <i>Sampling: Design and Analysis, Second Edition</i> . (Boston, MA: Brooks/Cole, 2010).
Summary Overview	This chapter provides brief summaries of seven types of sampling appropriate for estimating prevalence and characteristics of rare populations.
Detailed Information	<p>The sampling discussions included in Dr. Lohr’s text are described below.</p> <ul style="list-style-type: none"> <li>■ <b>Stratification with disproportional allocation</b> uses higher sampling fractions in strata where the prevalence of the rare population is higher. This approach will not work if an appropriate stratum for the rare population cannot be identified.</li> <li>■ <b>Two-phase sampling</b> screens in the phase 1 sample for the rare population and then uses a higher sampling fraction on those with the rare characteristic in the phase 2 sample. If the screening is perfect, the phase 1 sample can be used to estimate prevalence of the rare population and the phase 2 sample to estimate characteristics of the rare population. If the screening is not perfect, one solution is to sample both strata of members of the rare population (with a higher fraction) and the strata of nonmembers of the rare population.</li> <li>■ <b>Unequal probability sampling</b> is similar to stratification with disproportional allocation. In this design, probabilities of inclusions are estimated by a model based on characteristics related to the rare phenomena. These probabilities are then used to create strata or specific inclusion probabilities. In some cases geographic clusters can be used as the strata.</li> <li>■ <b>Multiple frame surveys</b> are used when there is no complete listing of members of the rare population. Incomplete frames can be combined and duplicates removed to construct a sampling frame for the population. Or, one can draw samples from each frame and combine the sample estimates for an overall population estimates.</li> <li>■ <b>Network or multiplicity sampling designs</b> links sampled households or individuals to other units in the population and asks the respondent to provide information on the units linked to him or her. For issues where the respondent is likely to be able to report on those linked to him, this method can reduce sampling variability of the estimated prevalence. Errors occur when the informant cannot provide accurate information on those linked to him.</li> <li>■ <b>Snowball sampling</b> assumes that members of the rare population know other members. Beginning with a few members of the rare population, the researchers asks them to identify other members of the population. In this way the sample grows exponentially. It is possible to get a fairly large sample this way but it is not a random sample in any way and there is no known probability of selection.</li> <li>■ <b>Sequential sampling</b> starts with one or a few observations and modifies the next phase of the sampling based on what is learned in the first. In this way, a researcher can learn from the first sample, what size an additional sample will need to be to achieve the desired precision. Adaptive cluster sampling is one form of sequential sampling that is based on the assumption that the rare population is geospatially aggregated.</li> </ul>
Comments	Similar to the designs described by Kalton, most of these types of samples all lend themselves to finding sufficient number of working people, working people in specific industries, or working people with workplace injuries or illnesses. However, snowball sampling does not yield a probability sample. In addition, any sampling requiring geographic clustering must be carefully researched to determine if household where there is an individual with workplace injuries and illness actually cluster geographically.

Author and Title	<b>Husberg, Bradley J.</b> "Surveillance for Nonfatal Work related Injuries in the Alaska Fishing Industry." <i>Proceedings of the International Fishing Industry Safety And Health Conference.</i> (Edited by J. Lincoln, D. Hudson, G. Conway, and R. Pescatore.) 2003: 353-358.
Summary Overview	Assess the number of non-fatal hospitalized commercial fishing injuries in Alaska between 1991 and 1998.
Method	Uses the Alaska Trauma Registry (ATR) to identify these injuries. The ATR is especially good for identifying injuries in Alaska because there are very few hospitals located nearby across the Alaska border; few injured in Alaska are taken from Alaska for treatment. It contains diagnosis codes that allow for the identification of type of injury and a narrative section that allows for identification of place of injury.
Findings	Reports that until 1997, commercial fishing had the highest number of hospitalized workplace injuries. In 1998, commercial fishing was overtaken by construction.
Comments	This study points out a registry of workplace injuries that could be added to additional data sources to provide a sampling frame.  A second study was presented by Husberg and Lincoln at the Second International Fishing Industry Safety and Health Conference using the same database to look at nonfatal injuries in the Bering Sea Crab Fishery. These proceedings were published in 2006.

Author and Title	<b>Kessler, Ronald C., Patricia A. Berglund, Catherine Coulouvrat, Timothy Fitzgerald, Goeran Hajak, Thomas Roth, Victoria Shahly, Alicia C. Shillington, Judith J. Stephenson, and James K. Walsh.</b> "Insomnia, Comorbidity, and Risk of Injury Among Insured Americans: Results from the American Insomnia Survey." <i>Sleep.</i> Vol.35. No.6. 2012: 825-834.
URL	<a href="http://www.hcp.med.harvard.edu/wmh/AIS_Study.php">http://www.hcp.med.harvard.edu/wmh/AIS_Study.php</a>
Summary Overview	Reports on survey of members of a health plan concerning insomnia, other health conditions, and injuries. Also uses claims data linked to individuals.
Method	Telephone survey with prenotification letter. Used established epidemiological tools for measuring insomnia and injury.
	Population: Adult members of a large health plan
	Sample: 10,094—drawn to be nationally representative. Only employed used for these analyses=7,428.
	Cooperation Rate was 65 percent.
Findings	Analyses for this study indicate a positive relationship between insomnia and workplace injuries.
Comments	This study demonstrates a different source for a sample universe—members of a health plan. Use of this type of frame would allow the use of administrative data as well. In addition, this raises the questions of using secondary analysis of nationally representative studies with relevant questions rather than a new data collection.

**Special Population Study – e.g., Union Study, Specialized Industry Studies**

Author and Title	<b>Morse, Tim, Laura Punnett, Nicholas Warren, Charles Dillon, and Andrew Warren.</b> “The Relationship of Unions to Prevalence and Claim Filing for Work-Related Upper-Extremity Musculoskeletal Disorders.” <i>American Journal of Industrial Medicine</i> . Vol. 44. 2003: 83-93.
Summary Overview	Assesses the degree to which union membership influences higher filings of worker compensation claims regarding musculoskeletal disorder (MSD) cases.
Method	<i>Population(s)</i> : This study analyzed a 1996 telephone survey of Connecticut (CT) workers with work-related MSD.
	<i>Sample size</i> : There were two samples drawn, the first being a random population sample using a random-digit dial method of 3,200 working-age CT residents who were screened for significant MSD, and the second being a random sample of 323 cases drawn from a list of Connecticut Workers’ Compensation (WC) First Reports of Injury.
	For the random population sample, workers with significant MSD were classified as “likely work-related” based on a positive response to one of the following questions: <ul style="list-style-type: none"> <li>■ Was this pain or discomfort either due to or made worse by your work?</li> <li>■ Did you tell the medical person your problem was work related?</li> <li>■ Did they say that your pain or discomfort was caused by your job?</li> <li>■ Did they say that your pain or discomfort was made worse by your job?</li> <li>■ As your workday went on did the pain increase, decrease or stay the same?</li> <li>■ As your workweek went on did the pain increase, decrease or stay the same?</li> <li>■ When you were away from work did the pain increase, decrease or stay the same?</li> </ul> <p>After weighting, the researchers analyzed the data to assess the prevalence of union membership, and the relationship of union membership to MSD and to occupational characteristics. The second data sample was then compared to the first sample.</p>
Findings	Study was consistent previous studies that found that unionized workers were more likely to seek compensation. The effect of unionization is different for different industries: <ul style="list-style-type: none"> <li>■ Manufacturing had a high rate of WC claims for both unionized and non-unionized workplaces.</li> <li>■ Unionized workers were found to significantly file more WC claims in industries with less of a clear history of filing and with traditionally better alternative general health and disability programs, such as government and transportation/utilities.</li> <li>■ Construction had low rates of filing, which may be due to the chronic nature of MSD and a lack of a consistent employer.</li> <li>■ Findings did not support the possibility of MSD claim filing being higher due to higher levels of MSD in unionized workplaces.</li> </ul>
Comments	Authors conclude that unions may contribute to a workplace culture where filing claims is seen as a legitimate process as opposed to a deviant one, as unions provide protection from discrimination and retaliation. This suggests that conducting surveys of unionized workplaces or using lists of union workers might bias reports of workplace illness or injury upwards. It also may provide guidance to data collection and analysis in that union membership should be identified and considered in the analysis.

Author and Title	<b>Hirsch, Barry, Davald Macpherson, and J. Michael Dumond.</b> "Workers' Compensation Reciprocity in Union and Nonunion Workplaces." <i>Industrial and Labor Relations Review</i> . Vol. 50, No. 2. 1997: 213-236.
Summary Overview	This study examines union-nonunion differences in workers' compensation receipt.
Method	<i>Population(s):</i> Union and nonunion workers
	<i>Sample size:</i> 14 calendar years between 1977 and 1993, 109,913 total individuals who responded to the Current Population Survey (CPS) in the 14 calendar years between 1977 and 1993.
	Using matched panels of union and nonunion workers from the March 1977-1993 CPSs, the researchers controlled for job injury risk through inclusion of an industry injury rate variable, industry dummies, establishment size, and measures of occupational working conditions. They then estimated probit equations examining how union status affects the probability of a worker having a successful indemnity claim for workers' compensation during a year.
Findings	Unionization has a substantial effect on indemnity claims. <ul style="list-style-type: none"> <li>■ Union workers are far more likely than nonunion workers, other things being equal, to receive benefits from workers' compensation.</li> <li>■ Since union workers are more likely to receive benefits than nonunion workers, union workers are more likely to file a claim.</li> </ul>
Comments	This study once again points out that union member experiences of workplace injuries and illnesses are different than nonunion members and should be considered in design and analysis.

Author and Title	<b>Krause, Niklas, Teresa Scherzer, and Reiner Rugulies.</b> "Physical Workload, Work Intensification, and Prevalence of Pain in Low-Wage Workers: Results from a Participatory Research Project with Hotel Room Cleaners in Las Vegas." <i>American Journal of Industrial Medicine</i> . 2005: 1-12.
Summary Overview	Culinary Workers Union Local 226 (Hotel Employees and Restaurant Employees Union) in Las Vegas, the Labor Occupational Health Program (LOHP) at the University of California, Berkeley, and the Department of Medicine at the University of California, San Francisco initiated the survey due to concerns that increasing injury rates and health plan costs reflected changes in the work environment of hotel workers.
Method	<i>Population:</i> Unionized day-shift hotel room cleaners in Las Vegas
	<i>Sample:</i> Five union hotels with 1,276 eligible respondents. There were 941 responses (74%).
	University researchers and cleaners from non-participating hotels fluent in Spanish, Serbo-Croatian, or one or more Asian languages administered surveys in the union hall. Twelve-month prevalence of pain perceived as work-related was measured by the question "Have you had any pain or discomfort during the past 12 months that you feel might have been caused or made worse by your work as a hotel room cleaner?" This question used the medical-legal criteria used by physicians to determine whether reported pain is work-related, that is, (i) whether it was caused by work and (ii) whether it occurred in the course of conducting work duties; (iii) or whether these work duties aggravated a non-industrial pre-existing condition so that (iv) the aggravation resulted in disability or need for medical care. The latter two conditions were reflected in three follow-up questions: "If yes, have you visited a doctor about this pain or discomfort? (yes/no)"; "If yes, have you called in sick in the last 12 months because of this pain or discomfort you feel was caused by or made worse by your work as a hotel room cleaner? (yes/no)"; and "Have you taken any sick or vacation days off work in the last 12 months because of this pain or discomfort you feel was caused by or made worse your work as a hotel room cleaner? (yes/no)."

Findings	All but 10 respondents were women. The authors found “significantly” higher incidence rates of self-reported injuries, which they believe “suggests substantial under-reporting by either workers to employers or employers to OSHA (Occupational Safety and Health Administration).”
Comments	Both this study and an earlier study in San Francisco were used in contract negotiations. This may introduce bias in the form of employer reluctance to allow participation, or in employee over-reporting to improve union outcomes.

### ***Linked Employer and Employee Study***

Author and Title	<b>Beatson, Mark.</b> “Multiple Perspectives on Employment Relations: Experience From the 1998 Workplace Employee Relations Survey.” <i>Proceedings of the Survey Research Methods Section, American Statistical Association.</i> 2000:201-210.
Summary Overview	First in a series of three papers describing two stage establishment and employee surveys. This describes the 1998 Workplace Employee Relations Survey (WERS98) conducted in Great Britain. This survey system was designed to provide large scale representative data on linked employer and employee information.
Method	Data were collected from 3,000 workplaces and self-completion questions were obtained from nearly 30,000 employees. Using a two-stage sampling strategies, the first phase was a stratified sample of establishments with 10 or more employees. Data were collected from a representative of sampled establishments by interviewees. Respondents were asked for permission to survey employees (86% granted). The respondent provided a list of all employees and the interviewer randomly selected 25. Managers were given questionnaire packs to distribute to employees that included prepaid return envelopes. When there were fewer than 25 employees, all employees were surveyed.
Comments	While most of these survey address employee relations issues, this approach would provide data that could provide both SOII (like) data from the employer and employee reports. One issue to consider is the rarity of employee injuries. Increasing the survey of employees beyond 25 could be considered.

Author and Title	<b>Greenen, Hathalle, and S. Hamon-Cholet.</b> “C.O.I.: A Matched Employer/Employee Survey on Organizational Change and Computerization.” <i>Proceedings of the Survey Research Methods Section, American Statistical Association.</i> 2000: 211-222.
Summary Overview	Second in a series of three papers describing two stage establishment and employee surveys. This describes the 1997 C.O.I survey conducted in France. This survey system was designed to provide large-scale representative data on linked employer and employee information. It is actually four surveys: three workplace surveys depending on sector and an employee survey.
Method	Target population was Firms operating in the manufacturing, accounting, and DIY sectors. They were selected from a government file. Number of employees required for the firm to be included varied by sector. Employees were selected for another government listing that links employees with the firms in which they are employed. Two employees were selected for the survey from firms smaller than 500 employees and three from larger firms. The designers chose the small number of employees to balance the cost of a large number of firms. The sample sizes were 4,026 firms and 8,812 employees. The firm data were collected by mail. The employee data were collected by telephone or in person. The survey had response rates for the sectors ranging from 56 percent to 88 percent. The employee survey reached 89 percent of the selected employees with 71 percent of those agreeing to participate.

Comments	Should a design similar to this be selected, the details of the approach will be useful. However, the small number of employees selected for the employee component will not be sufficient to capture information on the rare workplace injury.
Author and Title	<b>Patek, Zdenek, M. Hidiroglou, and P. Lavallee.</b> "The Methodology of the Workplace and Employee Survey." <i>Proceedings of the Survey Research Methods Section, American Statistical Association</i> . 2000: 223-235.
Summary Overview	Third in a series of three papers describing two-stage establishment and employee surveys. This describes the 1999 Workplace and Employee Survey (WES) conducted in Canada. This survey system was designed to provide large scale representative data on linked employer and employee information. This was also designed to continue as a longitudinal study.
Method	Target population was all workplaces operating in Canada with paid employees with some exceptions. The population was grouped by industrial activity and geographic region and stratified by size. Interviews with workplace representatives were CAPI. Interviewers then drew a sample of three, six, nine, or 12 employees depending on the size of the workplace. Data were collected from these employees using CATI.
Comments	Should a design similar to this be selected, the details of the approach will be useful. However, the small number of employees selected for the employee component will not be sufficient to capture information on the rare workplace injury.

Author and Title	<b>Probst, Tahira M., and Armando X. Estrada.</b> "Accident Under-reporting among Employees: Testing the Moderating Influence of Psychological Safety Climate and Supervisor Enforcement of Safety Practices." <i>Accident Analysis and Prevention</i> . Vol. 42. 2010: 1438-1444.
Summary Overview	Reports on anonymous survey of workers from five industries with above average risk for employee injury.
Method	Paper and pencil survey implemented at worksites. Survey covered organizational safety climate, supervisor enforcement of safety regulations, and accident reporting behaviors.
	Population: Workers at five industries (light manufacturing firm, small heating and cooling facility, dental clinics, pulp and paper mill, and multiple restaurants.
	Response rate was high but difficult to calculate because total number of employees of organizations was not known. In total, 425 individuals of all mine employees.
Findings	Analyses for this study indicate that low perceived supervisor enforcement of safety policies is significantly correlated with accident under-reporting. A small qualitative component of the study asked about the non-reporting. The most frequently endorsed reasons were: the employee took care of it himself/herself, nothing would be done to fix the problem anyway, and desire to avoid negative consequences of reporting such as followup interviews or loss of injury free record.
Comments	This study also demonstrates that with anonymity, there is employee description of rationale for under-reporting and little social desirability bias in the reporting.

Author and Title	<b>Probst, Tahira M., and Maja Graso.</b> "Pressure to Produce=Pressure to Reduce Accident Reporting." <i>Accident Analysis and Prevention</i> . Vol. 59. 2013: 580-587.
Summary Overview	Reports on anonymous survey of coal mine workers from a single mine.
Method	Paper-and-pencil survey implemented at mine location. Survey covered perceived production pressure, reporting attitudes, negative consequences for reporting, and accident reporting behaviors.
	Population: Workers at mine who were directly engaged in mining activities.

	Response Rate was 42 percent (n=212) of all mine employees. Researchers were unable to get a precise count of those not engaged directly in mining activities.
Findings	Analyses for this study indicate a positive relationship between perceived production pressure and the number of accidents that occur and a negative relationship between perceived production pressure and the number of accidents reported.
Comments	This study demonstrates the use of a small place of employment to test the accuracy of reporting as well as the mechanisms that may drive underreporting. It is of importance to note that the guarantee of anonymity was apparently sufficient for participation in the workplace. Respondents described both reported and unreported injuries.

Author and Title	<b>Souza, Kerry, Linda F. Cantley, Martin D. Slade, Ellen A. Eisen, David Christiani, and Mark R. Cullen.</b> "Individual-Level and Plant-Level Predictors of Acute, Traumatic Occupational Injuries in a Manufacturing Plant." <i>Occupational Environmental Medicine</i> . Published Online First: April 12, 2014.
Summary Overview	Reports on an anonymous employee satisfaction survey of 30,163 employees of a manufacturing company with 56 different locations. Information about work context and employee demographics at the facility level were linked to the company's database of injury reports to model the relationship among individual characteristics, location context as perceived by the employees, location characteristics, and workplace injury.
Method	Paper and online survey administered by HR research department of large manufacturing company.
	Population: 30, 163 employees at 56 locations of a single company
	Response rate: 70 percent
Findings	This study demonstrated the successful use of an employee survey to collect evaluative data on the work environment. The study found that work stress may increase the likelihood of injury and unionization may act as a protective factor against injury.
Comments	Because this study was able to collect evaluative work environment data through an employee survey with a high response rate, it can be assumed that questions about workplace injury could be added to such a survey with little impact on response rates. Ideally such a survey in the future could be structured so that employee's identities could be linked to administrative databases including injury report in the company without jeopardizing the employees in any way.

Author and Title	<b>Shuhui Wang, John R. Myers, and Larry A. Layne.</b> "Injuries to Hired Crop Workers in the United States—A Descriptive Analysis of a National Probability Study." <i>American Journal of Industrial Medicine</i> . 54:. 2011: 734-747.
URL	<a href="http://www.doleta.gov/agworker/news.cfm">http://www.doleta.gov/agworker/news.cfm</a>
Summary Overview	A report on the findings from an injury module added by NIOSH and BLS to the National Agricultural Workers Survey (NAWS) from 1999-2004. Combining these data, analyses on injuries were conducted on 13, 604 crop workers.
Method	The NAWS uses a multistage sampling strategy to select a national probability sample of crop workers. The stages are as follows: 1) Sample is stratified into 12 appropriate geographic regions; 2) From these 80 farm labor areas are selected which consist of similar farm usage; 3) Within the labor areas, counties are selected with probability proportional to the size of the farm expenses in that area; 4) Simple random selection is used to select growers within the counties; and 5) Workers are selected at each grower site. In each of the data collection years, the NAWS collected injury data on approximately 3,000 workers.

Comments	An extension of this multistage sampling strategy to workplaces throughout the US would yield a sample of workers from which estimates to compare to SOII estimates could be compared.
----------	--

Survey	<b>Medical Expenditure Panel Survey – Insurance Component</b>
URL	<a href="http://meps.ahrq.gov/mepsweb/survey_comp/Insurance.jsp">http://meps.ahrq.gov/mepsweb/survey_comp/Insurance.jsp</a>
Summary Overview	MEPS-IC has been conducted yearly since 1996 and is sponsored by the Agency for Healthcare Research and Quality.
Method	Population: Private and Public Sector Establishments
	<p>The MEPS Insurance Component is an annual survey of employers that collects information on employers' health insurance offerings. There are two distinct samples fielded in the MEPS IC survey: the List sample and the Household (link) sample (which is no longer in use, 1996-2001).</p> <p>The MEPS-IC List sample is an independently drawn, nationwide sample of establishments and state/local governments. The list sample is drawn from the Business Register, a confidential list of establishments in the United States maintained by the U.S. Census Bureau and from the Census of Governments, a frame maintained by the Census Bureau and updated once every 5 years. (Due to the confidentiality requirements surrounding any samples drawn from the Business Register, there can be no Public Use Files released with micro data.)</p> <p>“The MEPS-IC Household sample (also known as the Link sample or the Household Link sample) was a sample of employers that were identified by respondents in the MEPS-HC as their main employer or secondary employer that was the source of their health insurance. This sample was collected only from 1996 through 2001. This sample was directly linkable to the MEPS-Household Component survey and was specifically designed for that purpose. While the Household sample was not derived from a confidential frame, the promise of confidentiality given to respondents makes it impossible to issue a public use file with the microdata.”</p>
	<p>Sample size: Approximately 39,000 private sector establishments. For 2012, 23,877 single establishments and 15,157 multiunit establishments were sampled with a total of 39,216 establishments, plus an additional 2,912 state and local government units. The response rate was 79.9 percent for the full year (77.8% for single establishments; 83.17% for multiunit establishments). The response rate for state and local government units was 89.1 percent.</p> <p>The sample is drawn at the establishment level (particular workplace), not at the firm level (business entity); more than one establishment can be sampled from the same firm.</p>
Comments	<p>The response rates and samples from the state and local government units is not included in the full year rate/totals. The biggest issue with the Household Link Sample was confidentiality, which is why it was discarded</p> <p>This study might be a good source of establishment contact in a multi-stage approach to surveying employees.</p>



## 2. What Characteristics of the Employees to Be Surveyed Are Relevant to the Project?

### All “In Scope” Workplaces

Author and Title	<b>Boden, Leslie I., Nicole Nestorlak, and Brooks Pierce.</b> “Using Capture-Recapture Analysis to Identify Factors Associated with Differential Reporting of Workplace Injuries and Illnesses.” <i>JSM Section on Survey Research Methods</i> . October 2010.
Summary Overview	The authors compared data from the SOII to Boden and Ozonoff (2008)’s Wisconsin workers’ compensation data.
Method	The authors used a capture-recapture method to measure the undercount of workplace injuries and illnesses in SOII.
Findings	The authors found that “there is some evidence that the SOII is relatively less likely to capture a case if it is in Transportation, Communications and Utilities, or in Wholesale or Retail Trade than if it is in Agriculture, Construction, or Nondurables Manufacturing.”
Comments	Careful consideration needs to be given to covering all industries.

Author and Title	<b>Bureau of Labor Statistics.</b> “BLS Reports on the Amount of Formal and Informal Training Received by Employees. <i>1995 Survey of Employer-Provided Training: Employee Results</i> . December 19, 1996.
URL	<a href="http://www.bls.gov/ept/home.htm">http://www.bls.gov/ept/home.htm</a>
Summary Overview	This study uses employer lists as a frame for selecting employee participants.
Method	<i>Population:</i> Employees of private nonagricultural business establishments
	<i>Sample Size:</i> 2,214 eligible employees from the 1,062 establishments that participated in the employer survey.
	<i>Response rate:</i> 50.6 percent for the employee questionnaire and 47.7 percent for the employee log.
	<p>The sampling frame for the employee survey was a listing of employees supplied by the establishment respondent. Researchers randomly sampled two employees from all the employees in the establishment. If one or more of the employees was unavailable, researchers could generate up to six random numbers to try to secure the participation of two employees.</p> <p>Two survey instruments were used – an employee questionnaire and an employee training log. The employee questionnaire focused on employment and demographic characteristics, as well as general questions on types of training provided by the employer during the employee’s tenure and in the last 12 months and on the benefits of training. The employee log collected detailed information on all training and learning activities the employee participated in over a 10-day period.</p> <p>Each employee had a personal visit interview. During the interview, researchers administered a CAPI questionnaire to the respondent. Researchers also collected the employee log via paper and pencil for the 3-day period prior to the day of the interview and left behind a training log for the employee to complete over the next 7 days and mail back.</p>
Comments	Using this methodology, employee data could be linked or compared to employer data if higher numbers of employees were selected from each employer. In addition, it is one way of identifying employees for this new study.

Author and Title	<p><b>Kalleberg, Arne L., David Knoke, Peter V. Marsden, and Joe L. Spaeth.</b> "The National Organizations Study: An Introduction and Overview." <i>American Behavioral Scientist</i>. Vol. 37, No. 7. June 1994: 860-871.</p> <p><b>Spaeth, Joe L., and Diane P. O'Rourke.</b> "Designing and Implementing the National Organizations Study." <i>American Behavioral Scientist</i>. Vol. 37, No. 7. June 1994: 872-890.</p>
Summary Overview	This study uses the respondents of the work organization module of the General Social Survey (GSS) to develop a sample of organizations for studying.
Method	<i>Population:</i> Respondents and spouses from the 1991 GSS topical module of work organizations
	<i>Sample Size:</i> 1,427 nominations
	<i>Response Rate:</i> 64.5 percent completions, 23.1 percent refusals, and 12.4 percent pending
	<p>The 1991 GSS work organization module collected information on the names, addresses, and phone numbers at which GSS respondents work, which produced a multiplicity sample of work organizations of 900+ establishments selected with probability proportional to the size of their labor force. Establishments could be ruled ineligible if researchers were (1) unable to locate the establishment or it had gone out of business, or (2) it was a military establishment or other establishment for which informants may have been placed in jeopardy if they had answered.</p> <p>Interviewers conducted telephone interviews with informants from the establishments nominated by the GSS respondents. Of the 1,427 nominations, 909 were respondents and 518 were spouses. In case of duplicate nominations, informants were interviewed only once.</p>
Findings	Of the respondents (who were always the nominators), 86 percent were able to supply the establishment name, 3 percent did not know the name or did not answer, and 11 percent refused to provide this information. Although differences between respondents and spouses were significant at the 0.05 level, they were not large. Respondents were 5 percent more likely to provide an establishment name for themselves than for their spouses. Researchers found that "In general, inadequate information was a real but relatively minor problem. Ultimately, the refusal problem was twice as serious as the inadequate information problem.... Nearly 90 percent of those who refused to answer one item refused to answer all three."
Comments	This study provides information on strengths and weaknesses of using individual respondents to connect to work organizations.

### ***"In Scope" Workers with Vulnerable Status***

Author and Title	<p><b>Forst, Linda, Emily Ahonen, Joseph Zaroni, Alfreda Holloway-Beth, Michele Oschner, Louis Kimmel, Carmen Martino, Eric Rodriguez, Adam Kader, Elisa Ringholm, and Rosemary Sokas.</b> "More Than Training: Community-Based Participatory Research to Reduce Injuries Among Hispanic Construction Workers." <i>American Journal of Industrial Medicine</i>. Vol. 56. 2013: 827-837.</p>
Summary Overview	This study evaluates the effectiveness of a health and safety program designed for Hispanic construction workers.
Method	<i>Population(s):</i> Hispanic construction workers one-third of whom reported speaking English well and 61 percent had less than a high school education.
	<i>Sample size:</i> A total of 446 foreign-born Hispanic construction workers

	Eight different work centers around the country (in seven different cities) administered the Occupational Safety and Health Administration, 10-hour training sessions over a 3-year period to Hispanic construction workers.
Findings	Pre- and post-training surveys (in Spanish) were administered to measure the program's effectiveness. There was a significant gain in knowledge regarding preventing falls and knowledge regarding "grounding" on the risk of electrical shock. Of the participants, 84 also participated in a 3-month followup call. They reported having increased confidence to address hazards with supervisors.
Comments	This work suggests that this population may not know what is reportable and may be hesitant to report.

Author and Title	<b>Marucci-Wellman, Helen, David H. Wegman, Tom B. Leamon, Ta Thi Tuyet Binh, Nguyen Bich Diep, and David Kriebel.</b> "Work-Related Injury Surveillance in Vietnam: A National Reporting System Model." <i>American Journal of Public Health</i> . Vol. 103, No. 11. November 2013: 1989-1996.
Summary Overview	The authors field tested an active surveillance system as a field test for developing a national model of health surveillance.
Method	In 2006, the authors established an active surveillance system in Vietnam's Xuan Tien commune. They used community health treatment sites as the primary source of health reporting. The authors state they chose not to use employer reporting because while "larger state-owned and private enterprises have significant occupational medical care resources, including on-site nurses," the majority of employers are small- or medium-sized enterprises that "rely on the community health care system for occupational as well as all other injury and disease issues."
Findings	The authors emphasize the importance of defining the incident, episode or case in a "reasonably sensitive [but] not overly specific" way. For their survey in Vietnam, they struggled to find a word in Vietnamese that adequately captures the wide-range of events that the English word 'injury' entails; in Vietnamese, the word "injury" is mostly commonly understood to include only extreme injuries, which may reduce the count reported by hospitals in other studies.
Comments	Emphasizes the importance of considering colloquial terminology and translation of these questionnaires to communicate with non-English speaking workers.

Author and Title	<b>Pransky, Glenn, Daniel Moshenberg, Katy Benjamin, Silvia Portillo, Jeffrey Lee Thackrey, and Carolyn Hill-Fotouhi.</b> "Occupational Risks and Injuries in Non-Agricultural Immigrant Latino Workers." <i>American Journal of Industrial Medicine</i> . Vol. 42. 2002: 117-123.
Summary Overview	Conducted between November 1997 and January 1998 by the Tenants and Workers Support Committee, this is survey of work injuries among Latino immigrant workers in the Washington, DC metro area.
Method	<i>Population:</i> Nonagricultural workforce Latino immigrants, aged 18+
	<i>Sample Size:</i> 516 questionnaires completed, 427 used for analysis. Estimated >85 percent, 7 percent refusals and 8 percent not available. All 72 interviews from one interviewer were eliminated due to quality issues.
	Eight community health educators from the neighborhood administered the questionnaire. A 20-block area in the central part of the community was divided into eight sections. Within their assigned sections, interviewers selected apartments consecutively by address number. At each apartment, all adults 18+ who had worked for pay within the last year were asked to participate. Those respondents who reported more than one work injury were instructed to provide detailed descriptions only for the most severe event.

Findings	The study found an injury rate for respondents was 12.22; “the average rate in the US in 1997 was 7.1 injuries/100 FTE workers.” A total of 84 percent of respondents spoke little or no English.
Comments	This work uses an unusual methodology to secure cooperation from a vulnerable population.
Author and Title	<b>Guillermina Jasso, Douglas Massey, Mark Rosenzweig, and James Smith.</b> “The New Immigrant Survey 2003 Round 1 (NIS-2003-1) Public Release Data.” March 2006, <a href="http://nis.princeton.edu">http://nis.princeton.edu</a>
URL	<a href="http://nis.princeton.edu/">http://nis.princeton.edu/</a>
Summary Overview	Supported by the National Institutes of Health National Institute of Child Health and Human Development (NICHD), National Institute on Aging Office of Behavioral and Social Science Research (OBSSR), the National Science Foundation, and the U.S. Immigration and Naturalization Service (INS), the NIS Pilot was conducted in 1996, the first full cohort (NIS-2003-1) sampled immigrants in the period May-November 2003, the baseline survey was conducted from June 2003 to June 2004, and the followup interview (NIS-2003-2) was conducted from June 2007 to December 2009. During the study adult respondents were asked about employment history in the United States and abroad as well as to report on health conditions and symptoms.
Method	<i>Population:</i> New legal immigrants to the United States, including both immigrants arriving in the United States with immigrant documents acquired abroad and immigrants who are already in the United States with a temporary nonimmigrant visa (or, in some cases, illegally) and adjust to lawful permanent residence. <i>Sample Size:</i> 12,500 adult (18+) respondents; 1,250 child respondents <i>Response Rate:</i> 68.6 percent for adults (8573 completed interviews), 64.8 percent for children (810 completed interviews)
Findings	NIS is multi-cohort prospective-retrospective panel study based on nationally representative samples of the administrative records, compiled by the U.S. INS, pertaining to immigrants newly admitted to permanent residence. Interviews were conducted in respondents’ preferred languages. In the NIS-P, 46 percent of the interviews with adult immigrants were conducted in English, 26 percent in Spanish, and the remaining 28 percent in 17 other languages.
Comments	Captures employment and health data on minority and non-English speaking respondents

Author and Title	<b>Mobed, Ketty, Ellen B. Gold, and Marc B. Schenker.</b> “Occupational Health Problems among Migrant and Seasonal Farm Workers.” <i>Cross-Cultural Medicine: A Decade Later [Special Issue] Western Journal of Medicine</i> . Vol. 157. September 1992: 367-373.
Summary Overview	The authors reviewed literature on migrant and seasonal farm workers to explore problems of surveying these populations on occupational and other health topics.
Method	The authors conducted a literature review of a number of studies of migrant farm workers to develop estimates of prevalence of various occupational injuries and illnesses
Findings	The authors note that data on migrant and seasonal farm worker health is of poor quality or missing altogether, for various reasons: <ul style="list-style-type: none"> <li>■ Difficulties in locating and identifying farm workers due to the seasonal and migratory nature of the work and large distances between camps or farms in rural, often remote, areas;</li> <li>■ Unwillingness to cooperate after a long workday;</li> <li>■ Underreporting of symptoms, either because of mild or short-lived symptoms, or because of fear of job loss;</li> </ul>

	<ul style="list-style-type: none"> <li>■ Language barriers;</li> <li>■ Undercounting of those workers who meet the legal definition of a migrant but who do not fit ethnic and demographic stereotypes or occupational classifications; and</li> <li>■ The desire of many immigrant workers to avoid contact with government agencies.</li> </ul>
Comments	Currently the only national reporting system that tracks farm worker health data is the Migrant Student Record Transfer System maintained by the Office of Migrant Education of the U.S. Department of Education.

Author and Title	<b>Villarejo, Don, Stephen A. McCurdy, Bonnie Bade, Steve Samuels, David Lighthall, and Daniel Williams III.</b> "The Health of California's Immigrant Hired Farmworkers." <i>American Journal of Industrial Medicine</i> . Vol. 53. 2010: 387-397.
URL	<a href="http://www.doleta.gov/agworker/naws.cfm">http://www.doleta.gov/agworker/naws.cfm</a> (California data only)
Summary Overview	The authors review data from the California Agricultural Workers Health Survey (CAWHS), a comprehensive health survey conducted in 1999
Method	<p>Researchers identified 1,174 households with eligible residents. Of these, 940 households with 1,121 men and 522 women agreed to participate. A total of 627 men and 343 women were randomly selected. Of these participants, 416 men and 238 women also received a physical examination; these 654 respondents are used as the basis for the report.</p> <p>CAWHS used a multistage sampling strategy with seven representative communities within all six of California's agricultural regions. Investigators enumerated all potential dwellings in the target areas, both formal dwellings and informal ones (such as campsites, sheds, garages, abandoned vehicles, run-down trailers, and jerry-rigged shacks). A random sample of dwellings was drawn in each area, and residents were contacted in-person. In dwellings where residents agreed to cooperate, the interviewer enumerated all eligible workers currently residing there. One or more residents was randomly selected with an overrepresentation of women. Individuals were eligible if they were 18+ years and had performed hired labor on a U.S. farm within the prior 12 months.</p> <p>CAWHS included a main interview of family and personal demographics, insurance status, use of health care services, use of traditional healers, use of home remedies, self-reported health conditions, clinically determined health conditions, work history, income and living conditions, workplace health conditions, experience with protective equipment and training, working with pesticides in the United States, field sanitation, workplace injuries, and immigration status.</p>
Findings	<p>Researchers found that the following:</p> <ul style="list-style-type: none"> <li>■ Recently arrived immigrant workers are more likely to be undocumented as compared with farm laborers who had a long history of employment in California agriculture.</li> <li>■ Both male and female workers who lacked authorization for U.S. employment were more likely to report that their employer was a farm labor contractor, not a grower, as compared with other workers.</li> <li>■ Among men who were documented, 70 percent said they were aware of the provision of California Workers' Compensation insurance that can provide payments to an individual who becomes sick or injured while working (indemnity payments). Among men who were undocumented, less than half (40%) said they knew about this feature.</li> <li>■ A total of 25 percent of the men and 13 percent of the women said they had never had a medical or clinical visit.</li> </ul> <p>Additionally, the authors found evidence that some workers experienced threats at the workplace as well as being afraid of disclosing workplace injuries</p>

Comments	These findings indicate that undocumented and documented workers differ in significant ways which may impact their sampling for and/or reporting of occupational injury and illness, including type of employer (farm labor contractor vs. grower), knowledge of reportable injuries and illnesses, fear of retaliation, and access to medical personnel. This last finding may necessitate using a broader definition of reportable injury or illness, as workers may not have received medical attention regardless of severity.
----------	--

Author and Title	<b>Davis, Letitia, and Beatriz Pazos Vautin.</b> "Tracking Work-Related Injuries among Young Workers: An Overview of Surveillance in the United States. <i>Health and Safety of Young Workers: Proceedings of a U.S. and Canadian Series of Symposia.</i> (edited by C.W. Runyan, J. Lewko, K. Rauscher, D. Castillo, and S. Brandspigel) 2013: 105-124.
Summary Overview	Reviews current state of surveillance of workplace injuries among young adults, especially those in their teens.
Method	Review of surveillance systems that provide information on workplace injuries in youth under 18 years of age.
Findings	Reports on important limitations of the SOII in capturing teen worker injuries. These include: <ul style="list-style-type: none"> <li>■ Most teens work part time. Time lost at work may be more difficult to assess because they are not scheduled to work consecutive days, for example. This may lead to a workplace injury not being reported because it seemingly doesn't meet the requirements</li> <li>■ In all but the most populous states, the sample size is too small to obtain detailed information on young workers' injuries.</li> </ul>
Comments	This study points out a special population group that should be considered in the study design.

### ***Injury/Illness Status May Be Sensitive***

Author and Title	<b>Tourangeau, Roger, and Ting Yan.</b> "Sensitive Questions in Surveys." <i>Psychological Bulletin.</i> Vol.133, No.5. 2007: 859-883.
Summary Overview	This article provides a review of the literature on sensitive questions in surveys and includes the results of three new meta-analyses.
Detailed Information	There is strong evidence that misreporting is quite common on survey questions about sensitive topics. Respondents misreport because they do not want to embarrass themselves (social desirability) or the fear repercussions of what they are reporting. In general, the bias in a single survey related to this misreporting is not random but in a single direction causing the estimates to be biased rather than simply more variable.  Methodological factors can increase the likelihood of accurate answers. Self-administered surveys appear to increase accurate reporting when compared to interviewer administered questionnaires. Techniques that further enhance responses include randomizing questions so that the interviewer or administrator will not know exactly which questions the respondent is answering, creating an environment where the respondent believes that the truth will ultimately be known to the survey administrator so they tell the truth rather than be embarrassed later (bogus pipeline), wording questions in ways that are more forgiving, adding assurances of confidentiality, and matching interviewer and respondent demographic characteristics.
Comments	These techniques should be considered in light of knowledge that there may be disincentives for workers to report injuries or illnesses in ways that can become known to the employer.

### 3. How Will the Survey Be Designed to Collect the Appropriate Data?

#### *Ensure Understanding of Reportable Events*

Author and Title	<b>Shannon, Harry, and Graham Lowe.</b> “How Many Injured Workers Do Not File Claims for Workers’ Compensation Benefits?” <i>American Journal of Industrial Medicine</i> . Vol 42. 2002: 467-473.
Summary Overview	This study looks at the under-reporting issue in Canada, and what factors influence workers to not file for an injury. A key screening question in the survey was if respondents had experienced a work injury during the previous 12 months.
Method	<i>Population(s): Canadian workers</i>
	<i>Sample size:</i> 143 workers who were found to be injured from work. The team drew an initial sample of 2,500 phone numbers throughout Canada to identify workers who had experienced work-related injury. To find these 143 workers, the researchers filtered first on those in paid employment, then on those who reported that their injury required modified work, medical treatment, or lost time.  Workers were then asked if they had experienced a work injury during the previous 12 months, and if so, if the injury required time off work, modified work, or medical treatment. They were then asked if the injury occurred on the job, and if a WC claim had been submitted.
Findings	<ul style="list-style-type: none"> <li>■ A total of 40 percent of the eligible workers in the study did not submit a claim.</li> <li>■ Those who were not members of a union were less likely to submit a claim than union members, but it was not significantly different.</li> <li>■ The less “serious” the injury, the less likely a claim was to be submitted.</li> <li>■ Single job holders were much less likely to file a claim.</li> </ul>
Comments	This study identified workers injured on the job for a survey about their experiences. It provides examples of filters for use in a household survey of workers.

Author and Title	<b>Rosenman, K.D., J.C. Gardiner, J. Wang, J. Biddle, A. Hogan, M.J. Reilly, K. Roberts, and E. Welch.</b> “Why Most Workers with Occupational Repetitive Trauma Do Not File for Workers’ Compensation.” <i>Journal of Occupational and Environmental Medicine</i> . Vol. 42:1. 2000. 25-34.
Summary Overview	Michigan requires employers, clinics, hospitals, and physicians to report all known or suspected occupational diseases (but not injuries), with more than 20,000 reports produced every year, 90 percent coming from doctors. This study sought to determine why a worker did or did not file for workers’ compensation (WC) benefits.
Method	<i>Population(s): Michigan workers</i>
	<i>Sample size:</i> 1,598 workers with known occupational illnesses
	Researchers analyzed worker reports (from April through June 1996) looking to identify individuals with neck, back, or upper extremity disorders.  Workers in these reports received a letter 3-4 weeks after the report was received explaining the purpose of the study and inviting their participation. Depending on report volume/week, some weeks had all workers invited, while others had a random sample of 150 to adhere to realistic processing expectations.

	A telephone questionnaire was administered to willing participants (a total of 1,598 workers). The interviews asked about the event causing the disease report, demographics, health limitations from the trauma, and symptoms specific by body part. The severity of each type of trauma was assessed. The respondents were also asked if they filed a WC claim and if they were off work for 7 consecutive days.
Findings	<ul style="list-style-type: none"> <li>■ A total of 25% of the workers surveyed filed a WC claim.</li> <li>■ Main reasons for not filing a claim: <ul style="list-style-type: none"> <li>– Injury not serious enough (59%)</li> <li>– Did not expect to miss work (58.1%)</li> <li>– Expected to miss work but knew they would receive some form of sick leave or short-term disability from employer (28.3%)</li> <li>– Medical expenses covered by other insurance (35.9%)</li> <li>– Didn't think the injury was work-related (20.4%)</li> </ul> </li> <li>■ Part-time workers were found to be no less likely to file for a claim.</li> <li>■ Severity of injury was a strong predictor of filing.</li> <li>■ Low-income workers and unskilled workers were more likely to file a claim.</li> <li>■ Most significant predictor of filing for a WC claim was being off of work for 7 or more days.</li> <li>■ Workers who did not see their company physician were more likely to file a claim, with the odds of filing being greatest for those who saw a specialist or surgeon.</li> </ul>
Comments	Although this study is about workers' compensation, it provides insight into why workers may not report.

### ***Ensure Questions Cue Accurate Reporting of Injuries and Illnesses***

Author and Title	<b>Landen, Deborah D., and Scott Hendricks.</b> "Effect of Recall on Reporting of At-Work Injuries." <i>Public Health Reports</i> . Vol. 110, No. 3. May-June 1995: 350-354.
Summary Overview	The authors reviewed the literature on recall of workplace injuries and analyzed the recall of respondents to the National Health Interview Survey (NHIS).
Method	The authors used data from the 1988 Occupational Health Supplement (OHS) of the NHIS as well as reviewing the literature. In the 1988 OHS, an 'episode of injury' was defined as "any event causing an injury for which the respondent had (a) sought medical attention, (b) been unable to perform some work activities, (c) lost consciousness, or (d) transferred to another job."
Findings	<p>The authors state that "In the NHIS, the effect of recall on reporting of injuries can be assessed by examining the decrease in the number of injuries reported as time between reported date of injury occurrence and date of interview increases." This was achieved by asking respondents to report the date and characteristics of injuries. The authors calculated the recall period for each injury as the number of weeks between the date of the injury and the interview. For those injuries for which only the month of occurrence was available, they assigned the injury to the mid-point of the month.</p> <p>In reviewing the data, the authors also found that "The increase in the incidence estimate for non-lost-workday injuries (43 percent) was greater than that for lost-workday injuries (22.5 percent);" however, even the more severe lost-workday injuries "were greatly underreported among those in the 18-24-year group (44.6 percent)."</p>
Comments	Time of recall will be of critical concern to this study.



Author and Title	<b>Miller, Leslie A., and Theodore Downes-Le Guin.</b> “Improving Comprehension and Recall in the Consumer Expenditure Interview Survey: Discrepancies in Comprehension and Recall as a Source of Nonsampling Error.” <i>Proceedings of the Survey Research Methods Section</i> , American Statistical Association. 1989. Available at: <a href="http://www.amstat.org/sections/srms/proceedings/papers/1989_089.pdf">http://www.amstat.org/sections/srms/proceedings/papers/1989_089.pdf</a>
Summary Overview	The authors review the Consumer Expenditure Interview Survey (CEIS) and associated research in consideration of a survey redesign.
Method	The authors conducted literature reviews and cognitive testing and feasibility field tests as part of a discussion to redesign the CEIS and improve recall and comprehension among respondents.
Findings	<p>The authors report that “Currently the questions do not assist respondents in placing events in time by taking advantage of personal landmarks or cues respondents might use to anchor different reference periods.” Cognitive interview testing provided additional depth to this finding by revealing that “Most laboratory participants reported they would not use any temporal markers or cues to help them recall payments and reimbursements during the three month reference period;” however, respondents did use the individual for whom payments/reimbursements were made and the month the payment/reimbursement was made as mnemonic cues. The authors note that while the CEIS did not provide respondents “formal opportunities to use personal landmarks, the laboratory research suggests that many respondents attempt to create their own landmarks by remembering the 3-month history of specific [individuals].” They speculate that respondents may use personal events such as birthdays to mark the reference date. The authors consider the following methods for reducing memory decay:</p> <ol style="list-style-type: none"> <li>1. “Reducing the reference period;</li> <li>2. Increasing the number of questions about events to allow more time for respondents to recall and form questions;</li> <li>3. Giving respondents a chance to reconsider their answers by mailing a verification of responses or requesting the same information more than once in the survey;</li> <li>4. Using a bounding or reinterview procedure to give respondents an anchoring point and to sensitize them to the types of information that will be requested;</li> <li>5. Asking respondents to keep a diary to report events.”</li> </ol> <p>They conclude that not all methods are feasible for the CEIS, but ultimately recommend:</p> <ol style="list-style-type: none"> <li>1. Providing a sentence introducing and explaining the purpose of the data collection;</li> <li>2. Restructuring the interview to “better suit respondents’ scripts;”</li> <li>3. Providing recall cues and markers.</li> </ol>
Comments	These suggestions will be worth considering and further investigating in the development of the SOII employee data collection protocol.

Author and Title	<b>Moshiro, C., I. Heuch, A.N. Astrom, P. Setel, and G. Kvale.</b> “Effect of Recall on Estimation of Non-Fatal Injury Rates: A Community-Based Study in Tanzania.” <i>Injury Prevention</i> . Vol. 11. 2005: 48-52.
Summary Overview	The authors conducted a community-based survey within the Adult Morbidity and Mortality Project (AMMP) study areas in Tanzania to test the effect of recall on injury estimation. AMMP is a health and demographic surveillance system conducted in six districts of Tanzania, including Dar es Salaam city (an urban area) and Hai District (a rural area).
Method	<p>Researchers used two questionnaires:</p> <ul style="list-style-type: none"> <li>■ Questionnaire 1 recorded the information on whether an individual had an injury during the past one year, and included a list of injuries (broken bones, cuts or sprains, burns, dental, or other injuries) or injury events (transport accidents, falls,</li> </ul>

	<p>sports activities, snake or other animal bite, electric current, near drowning, struck by object, attempted suicide, or other injuries) read to the respondents.</p> <ul style="list-style-type: none"> <li>■ Questionnaire 2 recorded information concerning the description of the injury, the circumstances in which the injury occurred, and whether the injury was intended or not. Variables included were month and year, cause of the injury, place of occurrence, length of disability, and health facility use.</li> </ul> <p>Injuries were included if they had occurred in the past one year and resulted in losing one or more days of “normal” activity such as not being able to work or go to school. Only the month and year were reported for each injury, so researchers assumed the mid-point of the month when calculating the recall interval.</p> <p>For analytical purposes, each month covered by in the 12-month retrospective period was considered a separate recall period. This allowed researchers to examine drop-off in reporting overtime.</p>
Findings	<p>The authors found that an overall strong decline in reporting (81%) for minor injuries, while more severe injuries had “no consistent pattern in injury rates for the different recall periods.” However, “[for] a 1-month recall period, the rates were significantly higher in the rural compared to urban area and twice as high in males as in females.”</p> <p>Based on analysis, the researchers ruled out differences in levels of formal education, finding a similar decline found in both groups. The authors thus speculate that “basic cultural differences between population groups may account for the urban-rural contrast in memory decay.”</p> <p>The authors note an increase in estimated rates in the 12th month of recall in comparison with the 10th or 11th month of recall, which may be the result of events that happened more than a year ago being reported as occurring a year ago. Finally, the authors also report that the list of injuries and injury events was effective in aiding recall.</p>
Comments	<p>This article also point out recall considerations. Of new concern is the telescoping of previous year’s events into the last month of the year. The authors found differences in recall ability between respondents (male vs. female and urban vs. rural).</p>
Author and Title	<p><b>Overpeck, Mary, Ann C. Trumble, and Ruth A. Brenner.</b> “Population-Based Surveys of Sources of U.S. Injury Data and Special Methodological Problems.” No Date. Available online: <a href="http://www.cdc.gov/nchs/data/ice/ice95v1/c12.pdf">http://www.cdc.gov/nchs/data/ice/ice95v1/c12.pdf</a></p>
Summary Overview	<p>The authors discuss common methodological problems resulting from using population surveys to capture U.S. injury data.</p>
Method	<p>The authors conducted a literature review and expert analysis of population-based surveys including the National Health Interview Survey (NHIS) implemented in the United States.</p>
Findings	<p>“Analysis of NHIS data by injury type demonstrates the methodological strength of probing for injury episodes by asking about both medical attention and activity restrictions. Some types of injuries with high rates of medical attention do not result in high rates of activity restriction. Conversely, injuries serious enough to cause activity restriction do not always receive medical attention.”</p> <p>“Since some injuries have not received a medical diagnosis, lay terminology is needed to obtain an adequate description of the nature of the injury to facilitate coding of diagnostic categories. Even persons who received medical attention do not always understand the clinical terminology for the diagnosis or parts of the body affected. Probes about the part of the body affected, pictures of body parts, and alternate phrasing suggestions will help to identify the injury site.”</p>

Comments	This study reiterates the necessity of using lay or colloquial terminology and asking about occupational injuries and illnesses in a variety of questions to adequately cue respondents.
----------	--

Author and Title	<b>Smith, G.S., G.S. Sorock, H.M. Wellman, T.K. Courtney, and G.S. Pransky.</b> “Blurring the Distinctions Between On- and Off-the-Job Injuries: Similarities and Differences in Circumstance.” <i>Injury Prevention</i> . Vol. 12. 2006: 236-241.
Summary Overview	The authors use data from the redesigned National Health Interview Survey (NHIS) to compare circumstances of non-fatal work and non-work injuries.
Method	<p>The authors conducted secondary analysis from the 1997-1999 NHIS.</p> <p>Injury data was screened using the question “During the past 3 months, that is since (91 days before today’s date), (were/was) (you/anyone in your family) injured seriously enough that (you/they) got medical advice or treatment?” Interviewers captured verbatim text on injury, body part, and how the injury happened. Only injury descriptions codeable to the ICD-9CM nature of injury codes 800-999 (including acute musculoskeletal conditions but not gradual onset conditions such as tendonitis) were included in the injury file.</p> <p>Work-relatedness was determined by responses to: “What were you doing when the injury happened?” and “Where (were/was) (person) when the injury happened?” Up to two options were possible for each question which captured multiple activities such as driving and working. If “working at a paid job” was selected for either option the injury was considered work related.</p>
Findings	<p>The authors argue that “A major strength of the NHIS for occupational injury surveillance is that it is population-based and provides data on all medically treated injuries among people injured at work, regardless of workers’ compensation coverage, industry or employment status, or the severity of the injury. It also asks work-relatedness directly from the injured person or their proxy.”</p> <p>However, they note “conceptual difficulties” in determining appropriate definitions for occupational injuries. For example, in injury events such as an assault in a bar, one individual may be working while others are not. Respondents must consider themselves as “working for a paid job” in response to the question “What were you doing when the injury happened?” As a result, some workers, such as volunteer firefighters or those working in a family business or farm, may not report their injuries as work-related, even though they meet national criteria for work-relatedness.</p> <p>Additionally, the authors state it is possible that the work-relatedness of some reported injuries was not identified because work-relatedness was only asked as part of a general activity question with up to two responses allowed, instead of asking a separate work injury question.</p>
Comments	This study reiterates the necessity of using lay or colloquial terminology and asking about occupational injuries and illnesses in a variety of questions to adequately cue respondents.

Author and Title	<b>Warner, M., N. Schenker, M.A. Heinen, and L.A. Fingerhut.</b> “The Effect of Recall on Reporting Injury and Poisoning Episodes in the National Health Interview Survey.” <i>Injury Prevention</i> . Vol. 11. 2005: 282-287.
Summary Overview	The authors studied the effect of recall in the National Health Interview Surveys (NHIS) following the discussion to change the recall period.
Method	The authors use data from the 1997-1999 NHIS surveys. They used the question “During the past 3 months, that is since [91 days before today’s date], [were/was] [you/anyone

	in the family] [injured/poisoned] seriously enough that [you/they] got medical advice or treatment?" to screen for cases.
Findings	<p>The authors found evidence of both memory decay and telescoping. They found that episodes earlier in the recall period (further in the past) were more subject to recall bias; however, the “decay appeared to vary by severity of the episodes.” Increasing the recall period increases the sample size, particularly of severe injuries, but also increases the likelihood of reduced recall of less severe incidents.</p> <p>The authors recommend using a 5-week or 1-month recall period:          “For estimating the annual total number of injuries or the annual number of less severe injuries based on the NHIS data, limiting analyses to episodes with an elapsed 5 cumulative weeks has statistical, intuitive, and analytic appeal. The lowest estimated MSEs were attained when the annualized estimates were based on 3-6 cumulative weeks elapsing between injury and interview. Moreover, for comparison of NHIS data with other data, 5 weeks would make the estimate analogous to an estimate with a one month recall reference period.”</p> <p>However, they do note that:          “For injuries that are less likely to be forgotten, such as those requiring hospitalization or that are more severe (for example, fractures), episodes with a 3-month period can be used. The longer time period between injury and interview increases the number of recorded events. The resultant larger sample size allows for more detailed analyses and greater stability of estimates, which is particularly beneficial for studies of rarer events.”</p>
Comments	Should be considered in creating survey instrument for data collection. Reports of workplace injury or illness are likely to occur very close to the time of the injury or illness. This timely reporting would not be available to survey respondents.

Author and Title	<b>Wuellner, Sara E., and David K. Bonauto.</b> “Injury Classification Agreement in Linked Bureau of Labor Statistics and Workers’ Compensation Data.” <i>American Journal of Industrial Medicine</i> . Vol. 57. 2014: 1100-1109
Summary Overview	The authors compared SOII and workers’ compensation data to measure the undercount of workplace injuries and illnesses in SOII.
Method	The authors linked cases from the SOII to Washington State workers’ compensation (WC) claims.
Findings	In addition to the expected discrepancies due to reporting methods, they found differences in terminology and codes used: WC reports were more technical medically and the codes contained “higher portions of systemic diseases and multiple injuries... compared with SOII codes.” Additionally, they warn that “[the] examples provided in the SOII questions may lead the survey respondent’s description of the incident and injury or illness to conform to the examples provided. Indeed, among matched cases, a greater portion was coded in SOII as ‘strained back’ and ‘carpal tunnel syndrome’ while WC codes reflected a greater variety of injury types.”
Comments	While the new study will not involve workers’ compensation, the new study needs to ensure that the questions provide the same level of cueing as the SOII data form does.

***Ensure Appropriate Data Collection Methods***

Author and Title	<b>Huang, Yueng-Hsiang, Santosh K. Verma, Wen-Ruey Chang, Theodore K. Courtney, David A. Lombardi, Melanye J. Brennan, and Melissa J. Perry.</b> “Supervisor vs. Employee Safety
------------------	--

	Perceptions and Association with Future Injury in U.S. Limited-Service Restaurant Workers." <i>Accident Analysis and Prevention</i> . Vol. 47. 2012: 45-51.
Summary Overview	This study examined the association between supervisor and employee perceptions of management commitment to safety and the rate of future injuries in limited-service restaurant workers.
Method	<i>Population(s)</i> : Restaurant employees and supervisors
	<i>Sample size</i> : 453 employees of 34 limited service restaurants located in six U.S. states, belonging to three major chains.
	<i>Data Collection</i> : The employees of these restaurants were asked (for 12 weeks) to report their weekly injury experiences and the number of hours they worked during the week. They were able to report via telephone, Internet, or in writing. The surveys were available in English, Spanish, and Portuguese.
Findings	Supervisor perceptions were not a significant predictor of future employee injury rates. Only employee individual perceptions of management commitment to safety were significantly associated with the rate of injury.
Comments	This study provides examples of interesting and easy reporting methods for employees. See also: Verma, S.K., Chang, W., Courtney, T.K., and Lombardi, D.A. (2010). Workers experience of slipping in the US limited service restaurants. <i>Journal of Occupational &amp; Environmental Hygiene</i> , 7 (9) 491-500.

Author and Title	<b>Silverstein, Barbara A., Diana S. Stetson, W. Monroe Keyserling, and Lawrence J. Fine.</b> "Work-Related Musculoskeletal Disorders: Comparison of Data Sources for Surveillance." <i>American Journal of Industrial Medicine</i> . Vol. 31. 1997: 600-608.
Summary Overview	Thmontae authors drew on data from workers' compensation, personal medical benefits, mandated reporting, employer records, self-reports, and in-person assessments of workers to compare the availability and quality of data, and estimated prevalence of work-related musculoskeletal disorders.
Method	<i>Population(s)</i> : Active workers in four automotive plants
	<i>Sample size</i> : 626 workers (1987 baseline), 579 (1989 followup). A total of 416 workers participated in both the baseline and the follow-p.
	<i>Response rate</i> : 67 percent for baseline
	Self-administered questionnaires were distributed and collected by plant workers trained as ergonomic monitors during between May 1987 and May 1988. Followup questionnaires were completed in one plant in May 1989. In-person assessments were conducted by researchers and physicians beginning in May 1987 and 1-1.5 years later.
Findings	"[The] way in which a symptoms questionnaire is administered has an effect not only on response rate but also on prevalence rate. When [the Dickinson et al., 1997] questionnaire was administered by researchers, there was a higher response rate but lower prevalence rate than when self-administered. They conclude that self-administered questionnaires tend to be returned primarily by those who have a problem. In the current study, it appeared that those with more problems (identified on university examination) may have been less willing to indicate problems on the symptoms questionnaire. This difference may again relate to concerns about confidentiality. In the Dickinson study, 'assurances of confidentiality' were made in a cover letter and each self-administered questionnaire was returned to the researchers in a sealed envelope. This is in contrast to our study in which the university evaluations were stored and processed at the university while self-administered questionnaires were collected and processed by non-university personnel and stored at the plants."
Comments	Supports the use of a self-administered questionnaire.

## Appendix B Additional Relevant Studies

These studies were not reviewed for this document but may be reviewed during the cost-benefit analysis phase of this task.

1. Workers' Rights-Access, Assertion, and Knowledge (WRAAK) study:  
<http://www.dol.gov/asp/evaluation/CompletedStudies.htm>
2. Family and Medical Leave Act study:  
<http://www.dol.gov/asp/evaluation/CompletedStudies.htm>
3. Quality of Work life Questionnaire:  
<http://www.cdc.gov/niosh/topics/stress/qwlquest.html>
4. National Survey of Mining Population:  
<http://www.cdc.gov/niosh/mining/works/coversheet776.html>
5. National Electronic Injury Surveillance System (NEISS) follow-back investigations:  
<http://www.cpsc.gov/en/Safety-Education/Safety-Guides/General-Information/National-Electronic-Injury-Surveillance-System-NEISS/>
6. *An Experimental Study Using Opt-in Internet Panel Surveys for Behavioral Health Surveillance:*  
<http://stacks.cdc.gov/view/cdc/19117>.
7. *Evaluating Workplace Injury and Illness Records; Testing a Procedure:*  
<http://www.bls.gov/opub/mlr/1988/04/rpt2full.pdf>
8. Health and Retirement Study (University of Michigan): <http://hrsonline.isr.umich.edu/>

## Appendix C

# Literature Reviewed but Not Deemed Directly Relevant

**Boden, Leslie, and Emily Spieler.** “The relationship between workplace injuries and workers’ compensation claims: the importance of system design.” In: Victor RA, Carrubba LL, eds., *Workers’ Compensation: Where Have We Come From? Where Are We Going?* Cambridge, Massachusetts: Workers’ Compensation Research Institute. 2010.

**BLUE-ETS Conference.** *Proceedings from the BLUE-ETS Conference on Burden and Motivation in Official Business Surveys.* Held March 22-23, 2011 in Heerlen, The Netherlands. (Reviewed 25 abstracts and presentations.)

**Givens, Jimmie D., and James T. Massey.** “Comparison of Statistics from a Cross-Sectional and a Panel Survey.” National Center for Health Statistics.No Date. 640-645.

**Huang, Yueng-Hsiang, Michael Ho, Gordon S. Smith, and Peter Y. Chen.** “Safety climate and self-reported injury: Assessing the mediating role of employee safety control.” *Accident Analysis and Prevention.* Vol. 38. 2006: 425-433.

**Lee, Jin, Yueng-Hsiang Huang, Michelle M. Robertson, Lauren A. Murphy, Angela Garabet, and Wen-Ruey Chang.** “External Validity of a Generic Safety Climate Scale for Lone Workers Across Different Industries and Companies.” *Accident Analysis and Prevention.* Vol. 63. 2014: 138-145.

**Leigh, J. Paul, James P. Marcin, and Ted R. Miller.** “An Estimate of the U.S. Government’s Undercount of Nonfatal Occupational Injuries.” *Journal of Occupational Environmental Medicine.* Vol. 46. 2004: 10-18.

**MacEachen, Ellen, and Agnieszka Kosny.** “The “Toxic Dose” of System Problems: Why Some Injured Workers Don’t Return as Expected.” *Journal of Occupational Rehabilitation.* Vol. 20. 2010: 349-366.

**Nestoriak, Nicole and Brooks Pierce.** “Comparing Injury Data from Administrative and Survey Sources: Methodological Issues.” *Proceedings of the Survey Research Methods Section, American Statistical Association.* 2012: 4605-4617. .

**Oleske, Denise, Robert Brewer, Pamela Doan, and Jerome Hahn.** “An Epidemiologic Evaluation of the Injury Experience of a Cohort of Automotive Parts Workers: A Model for Surveillance in Small Industries.” *Journal of Occupational Accidents.* Vol. 10. 1989: 239-253.

**Smith, Gordon, Yueng-Hsiang Huang, Michael Ho, and Peter Chen.** “The Relationship Between Safety Climate and Injury Rates Across Industries: The Need to Adjust for Injury Hazards.” *Accident Analysis and Prevention.* Vol. 38. 2006: 556-562.

**Sorock, Gary, Gordon S. Smith, Gordon R. Reeve, John Dement, Nancy Stout, Larry Layne, and Susan T. Pastula.** “Three Perspectives on Work-Related Injury Surveillance Systems.” *American Journal of Industrial Medicine.* Vol 32. 1997: 116-128.