A Reliability Study of Potential Risk Factors for Acute Traumatic Occupational Hand Injuries

David A. Lombardi, PhD,[1,2,3]* Gary S. Sorock, PhD,[1,2,3] Mary F. Lesch, PhD,[1] Russ Hauser, MD, ScD,[2] Ellen A. Eisen, ScD,[2,4] Robert F. Herrick, ScD,[2] and Murray A. Mittleman, MD, Dr.PH[1,5]

Background *A reliability study of transient workplace exposures was conducted within a case-crossover study of acute traumatic occupational hand injuries. Exposures examined included features of work equipment (unusual performance of equipment or materials, glove use), work practices (performing an uncommon work task, using an unusual work method), and worker-related factors (being rushed, distracted, ill, or working overtime).*

Methods *Interviewers administered a follow-up questionnaire about exposure up to 4 days after the initial interview. The reliability of the usual frequency of exposure in the past month, exposure reported at the time of the injury, and the degree to which subjects had confidence in the accuracy of their memory for exposure was assessed.*

Results *The test-retest reliability for the estimated number of hours of exposure during the past month was high for six of the eight potential risk factors (intraclass correlation coefficients (ICC) ranging from 0.84–0.99). ICC for the frequency per month and minutes per exposure, used to calculate the estimated number of hours per month, ranged from 0.12–0.99 and 0.49–0.99, respectively. Agreement of reported exposure at the time of the injury, evaluated using the kappa statistic, ranged from substantial agreement to perfect agreement (0.65–1.0). However, for some factors, there was little or no reported exposure. Most subjects (74–100%) reported being "very confident" or "certain" in their memory for each exposure.*

Conclusions *These findings suggest that the frequency and duration of unusual transient workplace exposures can be reliably recalled by subjects interviewed by telephone within 4 days of an injury.* Am. J. Ind. Med. 42:336–343, 2002. *© 2002 Wiley-Liss, Inc.*

KEY WORDS: reliability; test-retest; memory; exposure; occupational; case-crossover; hand; injury; epidemiology

INTRODUCTION

Occupational injury research often uses workers' retrospective recall of exposures prior to their injury. However, the reliability of these self-reported exposures may depend on the specific method of data collection in a given study. As outlined in the NIOSH report, Traumatic Occupational Injury Research Needs and Priorities [NIOSH, 1998], our current challenge is to move beyond descriptive information to the collection of potential risk factor data to develop specific strategies to prevent occupational traumatic injuries. The authors of that document note that "data are especially lacking on worker exposures to hazards."

[1]Liberty Mutual Research Center for Safety and Health, Hopkinton, Massachusetts
[2]Occupational Health Program, Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts
[3]Department of Biostatistics and Epidemiology, University of Massachusetts School of Public Health Sciences, Amherst, Massachusetts
[4]Department of Work Environment, University of Massachusetts, Lowell, Massachusetts
[5]Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts
Contract grant sponsor: National Institute for Occupational Safety and Health; Contract grant number: R01 OH-03763-03.
*Correspondence to: David A. Lombardi, Ph.D., Injury Epidemiologist-Epidemiology, Liberty Mutual Research Center for Safety and Health, 71 Frankland Road, Hopkinton, MA 01748. E-mail: david.lombardi@libertymutual.com

Accepted 25 June 2002
DOI 10.1002/ajim.10120. Published online in Wiley InterScience (www.interscience.wiley.com)

A limited number of studies have evaluated the quality of self-reported exposure data used in studying potential risk factors for occupational injury. Data collected using the Behavioral Risk Factor Surveillance System (BRFSS), a telephone administered survey, have been found to be relatively reliable in a test-retest study of injury related questions [Koziol-McLain et al., 2000]. Another study found that a single telephone interview is a reliable and valid method for collecting data on the duration and frequency of common work tasks [Kallio et al., 2000]. However, it was found that participants had greater difficulty recalling the frequency and duration for unusual tasks.

It is also important to examine the degree to which subjects have confidence in the accuracy of their memory for workplace exposures. This is based on the assumption that events and exposures that are recalled with a high degree of certainty may be more accurately reported. However, the relationship between confidence ratings and reliability of recall is unclear. Some studies have observed a positive correlation [e.g., Brigham, 1990], while others have suggested a low or negative correlation between validity of recall and self-rated confidence in the accuracy of recall [e.g., Chandler, 1994].

Sorock et al. [2001] recently reported the methods and initial findings of a case-crossover study of acute traumatic occupational hand injuries. The focus of that study was to assess the association between transient exposures in the workplace and the short-term risk of sustaining a hand injury. In that study, subjects were interviewed by telephone a median of 1.2 days after reporting a hand injury. The a priori hazard period (i.e., case period) was defined as 10 min before the injury. Two control periods were used. One was 60–70 min prior to the injury (matched-pair interval analysis) and the other was the average amount of work-time exposed in the previous month (usual frequency analysis). An assessment of the reliability of these self-reported exposures during both the case and control periods is essential to understand the validity of the study results.

The primary purpose of this study, which was conducted within the larger case-crossover study of traumatic hand injuries [Sorock et al., 2001], was to determine the degree to which reports of transient exposures are reproducible and the degree to which subjects have confidence in the accuracy of their memory for these exposures.

MATERIALS AND METHODS

Design

Between June and November 2000, hand-injured subjects were enrolled by 5 of the 23 clinics from the main study [Sorock et al., 2001]. Subject participation required both an initial telephone interview, conducted one to 2 days after the injury and a follow-up interview. The interviewer who conducted the initial interview re-administered portions of the original structured study questionnaire up to 4 days later. Both the initial and follow-up interviews utilized similar primary data collection instruments, training materials, and data management procedures.

Study Sample

Thirty-six subjects were recruited at occupational health clinics and gave written informed consent to participate in an initial telephone interview. Subjects also provided verbal consent to be contacted for the follow-up interview at the end of the first interview. At that time, the subjects were not told that the questions they just answered would be repeated, only that they would be asked about the state of their health and work conditions prior to the injury. The study was approved by the Harvard School of Public Health Human Subject Committee. Thirty-one subjects (86.1%) participated in the follow-up interview, two refused to participate, and three were unreachable. However, of those who participated in both interviews, two subjects were excluded from the primary analysis because their follow-up interviews occurred four or more days after the initial interview (4.2 and 6.9 days, respectively). Therefore, data from 29 subjects (80.6%) were analyzed.

Reliability Outcome Measures

In the main study, eight exposures were evaluated as potential risk factors for a job-related hand injury. Exposures assessed included: performing uncommon work tasks; using uncommon work methods; working with malfunctioning tools or equipment or working with materials which performed differently than usual; glove use; working overtime; being rushed; being distracted; and feeling ill.

To improve recall of the injury event, the interview began with an open-ended question asking the subject to describe the circumstances of the injury. The Work-Time-Before-Injury log [Sorock et al., 2001] was used to document the presence or absence and timing of exposures during the 90-min period before the injury and the frequency and duration of each transient exposure during the past month. The subject was also asked to re-state the injury time, his/her work tasks, and any work breaks during the 90-min period leading up to the injury.

To estimate the usual frequency of exposure in the past work-month, subjects were asked to estimate how often they were exposed to each potential risk factor in the past month, and the average duration (in minutes or hours) of each episode of exposure. Exposure to each potential risk factor was then calculated by multiplying the reported frequency per month by the average duration.

The reliability of subjects' reported exposure at the time of the injury and their reported usual frequency of exposure during the month before the injury were assessed. In the follow-up interview, subjects were also asked to rate their confidence in the accuracy of their memory for their usual frequency of exposure to each of the potential risk factors over the month before the injury. A five-point Likert scale ranging from 1 to 5 (1 = "Not at all confident," 2 = "Somewhat confident," 3 = "Average confidence," 4 = "Very confident," and 5 = "Certain") was used [Rosenthal and Rosnow, 1984].

Data Analysis

Intraclass correlation coefficients (ICC) were calculated to evaluate the reliability of the usual frequency of exposure to each of the potential risk factors over the past month [Shrout and Fleiss, 1979]. Cohen's kappa coefficient was calculated to assess the degree of reliability (agreement) between the first and second interview in the reported presence or absence of each transient exposure at the time of the injury [Fleiss, 1975, 1981]. A kappa greater than 0.80 represents almost perfect agreement over and beyond chance, a value between 0.60 and 0.80 represents substantial agreement, a value between 0.40 and 0.60 represents moderate agreement, and a kappa below 0.40 represents fair to poor agreement [Landis and Koch, 1977]. Continuous variables are presented using the mean $\pm$ SEM.

To assess the subjects' confidence in the accuracy of their memory for usual frequency of exposure during the past month, the percent of subjects at each level of confidence was calculated. The mean $\pm$ SEM for the five-point Likert scale for each exposure was also calculated. For each potential risk factor, the subjects' self-reported confidence in the accuracy of their memory was correlated with the absolute value of the difference in the usual frequency of exposure between the initial and follow-up interviews. Spearman rank correlation coefficients and *P*-values were calculated [Conover, 1980].

RESULTS

Characteristics

Nineteen of 29 subjects (66%) were men with an average age of 40 years (range 19–71) (Table I). The majority of subjects were white, non-Hispanic (72%), followed by Hispanic (21%), Black, non-Hispanic (3%), and Asian/Pacific Islander (3%). The majority of injuries were lacerations (87%), followed by avulsions (10%), and crushes (3%). Injuries were mild in severity with a mean total laceration length of 2.1 ± 0.2 cm requiring 4.0 ± 0.5 sutures. There were 29 finger injuries, the majority occurring to the distal (69%), followed by the medial (17%), and proximal (14%) portions of the finger.

For these subjects, the median time elapsed between the injury and the initial telephone interview was 1.4 days (range 0.2–11.2 days) and the average telephone interview lasted 22 min. The follow-up interview occurred a median of 1.2 days after the initial interview (range 1.0–3.9 days).

Reliability of the Usual Frequency of Exposure in the Prior Month

For five of the eight potential risk factors, at least 75% of the subjects reported that they had no exposure in the month before the injury. Table II shows that the reliability for the estimated number of hours (i.e., usual frequency) of exposure during the past month was high for six of the eight potential risk factors with ICC ranging from 0.84–0.99. Reliability was the highest for feeling ill, performing an unusual work task, using a different work method, and wearing gloves. Reliability was also moderately high (ICC > 0.80) for rushing and the use of equipment or materials, which performed differently than usual. However, the reliability was only moderate for working overtime, and lowest for being distracted (ICC = 0.55, 95% CI: 0.23–0.76). At the first interview, subjects reported that they were distracted an average of 5.3 ± 3.0 hr during the month before the injury, compared with an average of 6.9 ± 3.9 on the follow-up interview.

The analysis was repeated to include all study subjects who had both an initial and follow-up interview (n = 31) regardless of the amount of time, which passed after the initial interview (i.e., including the two subjects who were contacted 4.2 and 6.9 days after the initial telephone interview). The test-retest reliability for the estimated number of hours of exposure remained similar for the majority of factors. However for two factors, feeling ill and being distracted, ICC decreased (ICC = 0.47 and 0.20, respectively), possibly indicating better recall over a shorter time period.

Reliability of the Components of the Usual Frequency of Exposure

Frequency of exposure

Table III shows that the frequency of working overtime during the past month was the most reliable exposure, followed by the frequency of feeling ill, and using a different work method. The ICC was greater than 0.85 for these factors. Reliability was moderate for the reported frequency of performing a usual task, rushing, wearing gloves, and using equipment or materials, which performed differently than usual (ICC ranging from 0.62–0.76). Reliability was lowest for the reported frequency of being distracted (ICC = 0.12, 95% CI: −0.25–0.46).

Duration of exposure

The test-retest reliability was highest for duration per exposure (minutes) for the use of equipment or materials

TABLE I. Demographic and Interview Characteristics

Characteristic	Men (n = 19)	Women (n = 10)	Total (n = 29)
Age (years)			
Mean ± SEM	41.2 ± 2.9	37.1 ±3.4	39.8 ± 2.2
Median	40.0	35.7	38.5
Range	19–71	21–50	19–71
Ethnicity N (%)			
White, non-Hispanic	16 (84.2)	5 (50.0)	21 (72.4)
Hispanic	2 (10.5)	4 (40.0)	6 (20.7)
Black, non-Hispanic	0 (0.0)	1 (10.0)	1 (3.4)
Asian or Pacific Islander	1 (5.3)	0 (0.0)	1 (3.4)
Type of injury[a] N (%)			
Laceration	16 (80.0)	10 (100.0)	26 (86.7)
Avulsion	3 (15.0)	0 (0.0)	3 (10.0)
Crush	1 (5.0)	0 (0.0)	1 (3.3)
Time from injury to initial interview (days)			
Mean±SEM	2.0 ± 0.4	2.7 ± 1.1	2.3 ± 0.4
Median	1.4	1.1	1.4
Range	0.4–4.3	0.2–11.2	0.2–11.2
Time from initial interview to follow-up (days)			
Mean±SEM	1.6 ± 0.2	1.7 ± 0.1	1.6 ± 0.1
Median	1.0	1.9	1.2
Range	1.0–3.9	1.0–2.1	1.0–3.9

[a]Subjects may have more than one type of injury and percentages are based on the total number (n = 30).

which performed differently than usual and using a different work method (Table III). However, for these two potential risk factors, the majority of subjects (>82%) reported no exposure in the past month, thus the number of minutes per exposure was equal to zero for these subjects. For those transient factors with the greatest exposure (e.g., wearing gloves, working overtime, and rushing) the test-retest reliability was still fairly high (ICC > 0.70). Reliability was the poorest for being distracted and for performing an unusual task (ICC of 0.49 and 0.53, respectively).

For both the frequency and duration of exposure, a simple linear regression of the data from the two interviews was conducted to examine the relationship between the magnitude of the response and the degree of reliability between baseline and follow-up measurements. Due to relatively small sample sizes, this exploratory analysis was

TABLE II. Intraclass Correlation Coefficient (ICC) and 95% CI for the Number of Hours of Exposure During the Past Month (n = 29)

	Reported hours of exposure		
Potential risk factor	Initial interview Mean ± SEM	Follow-up interview Mean ± SEM	ICC (95% CI)
Equipment/materials performed differently	0.79 ± 0.74	1.4 ± 1.4	0.84 (0.69–0.92)
Different work method	3.8 ± 3.6	5.2 ± 3.8	0.92 (0.84–0.96)
Unusual work task	0.49 ± 0.31	0.45 ± 0.25	0.93 (0.86–0.97)
Distracted	5.3 ± 3.0	6.9 ± 3.9	0.55 (0.23–0.76)
Rushing	29.2 ± 8.0	28.3 ± 7.6	0.85 (0.71–0.93)
Working overtime	21.5 ± 4.7	19.9 ± 3.6	0.70 (0.46–0.85)
Feeling ill[a]	0.43 ± 0.31	0.58 ± 0.33	0.99 (0.98–1.0)
Wearing gloves	62.6 ± 13.6	68.5 ± 14.1	0.91 (0.82–0.96)

[a]One subject had missing data.

TABLE III. Intraclass Correlation Coefficient (ICC) for the Frequency per Month and Minutes per Exposure Reported During the Initial and Follow-Up Interview (n=29)

	Frequency per month			Minutes per exposure		
Potential risk factor	Initial interview (# subjects without usual exposure) Mean ± SEM	Follow-up interview (# subjects without usual exposure) Mean ± SEM	ICC (95% CI)	Initial interview Mean ± SEM	Follow-up interview Mean ± SEM	ICC (95% CI)
Equipment/materials performed differently	(27) 1.0 ± 0.79	(26) 1.8 ± 1.5	0.76 (0.55–0.88)	2.4 ± 2.1	2.5 ± 1.9	0.99 (0.97–0.99)
Different work method	(23) 1.5 ± 0.68	(24) 1.6 ± 0.88	0.86 (0.73–0.93)	23.3 ± 16.9	29.3 ± 17.8	0.95 (0.90–0.98)
Unusual work task	(26) 0.36 ± 0.24	(24) 0.52 ± 0.26	0.62 (0.34–0.80)	13.1 ± 9.1	10.6 ± 5.4	0.49 (0.16–0.72)
Distracted	(22) 9.6 ± 5.3	(22) 3.5 ± 1.2	0.12 (−0.25–0.46)	13.8 ± 10.4	28.7 ± 16.9	0.53 (0.22–0.75)
Rushing	(7) 10.1 ± 1.9	(8) 8.9 ± 1.5	0.64 (0.37–0.81)	127.0 ± 27.7	163.9 ± 32.6	0.72 (0.48–0.86)
Working overtime	(8) 9.4 ± 2.0	(9) 9.4 ± 1.9	0.99 (0.98–1.0)	128.0 ± 25.3	136.5 ± 28.0	0.86 (0.72–0.93)
Feeling ill[a]	(26) 0.14 ± 0.08	(24) 0.18 ± 0.09	0.92 (0.83–0.96)	24.8 ± 18.2	46.6 ± 24.7	0.70 (0.45–0.84)
Wearing gloves	(9) 19.2 ± 5.8	(9) 28.2 ± 14.5	0.67 (0.42–0.83)	174.3 ± 37.1	183.4 ± 38.6	0.90 (0.80–0.95)

[a]One subject reported being ill 1 day per year, which was converted to 0.08 (rounded to 0.1) times per month, and one subject had missing data.

performed by pooling the data for all eight risk factors and creating separate regression models for the frequency per month and minutes per exposure. The R^2 for the duration per exposure (minutes) was 0.74, whereas for the frequency per month it was 0.57. An examination of the studentized residuals from each model showed that as the magnitude of the frequency and duration of exposures increased, so did the difference between baseline and follow-up responses.

Reliability of Exposure at the Time of the Injury

The number of subjects reporting exposure to each of the eight potential risk factors at the time of injury is reported in Table IV. The agreement was high for seven of the eight potential risk factors evaluated using the kappa statistic. For reports of rushing, glove use, and performing a different work method at the time of the injury, the kappa statistic ranged from 0.84–0.92 suggesting almost perfect agreement, whereas for the use of equipment or materials, which performed differently than usual, there was moderate agreement (kappa = 0.65). There was perfect or near perfect agreement among the 29 subjects for working overtime, feeling ill, or performing a usual task, however due to little or no exposure reported for these factors, kappa could not be estimated. There was also perfect agreement (kappa = 1.0) for being distracted, where 10% (3/29) of subjects reported being exposed at the time of the injury during both interviews.

Confidence in the Accuracy of Memory

Table V presents the degree to which the subjects had confidence in the accuracy of their memory for each of the potential risk factors for the month prior to the injury. Subjects had the highest confidence in their memory for the number of hours of wearing gloves, feeling ill, and working overtime and had the lowest confidence for doing an unusual work task when injured. However, 74% of subjects reported being either "very confident" or "certain" for this factor. Additionally, the mean confidence ratings for each potential risk factor ranged from 4.0 to 4.4, which indicates that subjects were on average "very confident" in the accuracy of their memory for these factors.

There were no consistent correlations between the absolute value of the difference in the subjects' usual frequency of exposure reported at the initial and follow-up interviews and the reported confidence in the accuracy of their memory of the respective exposures for the month prior to the injury. However, for performing an unusual work task, there was a statistically significant negative correlation ($R = -0.41$, P-value $= 0.03$), suggesting that workers with more reliable recall were more confident in their ability to recall and report this exposure.

DISCUSSION

The ability of workers to recall the frequency and duration of exposures that occur on the day of their injury, or in the recent past, is considered a potential limitation of retrospective study designs used to evaluate self-reported risk factors for occupational injuries. In this study, there was a high degree of reliability for workers' self-reports of exposure for all potential risk factors evaluated over the month before their injuries, except for being distracted. This was shown by high ICC. For self-reports of exposure at the time of injury, there was substantial or near perfect agreement, based upon

TABLE IV. Kappa Statistics for Exposure to Potential Risk Factors at the Time of the Injury

Potential risk factor	Exposed on: Initial interview (n=29)	Exposed on: Follow-up interview (n=29)	Kappa	95% CI of Kappa
Equipment/materials performed differently	2	1	0.65	0.02–1.0
Different work method	4	3	0.84	0.53–1.0
Unusual work task	1	0	No estimate	No estimate
Distracted	3	3	1.00	1.0–1.0
Rushing	10	9	0.92	0.77–1.0
Working overtime	0	0	No estimate	Perfect agreement (no discordant pairs)
Feeling ill[a]	0	0	No estimate	Perfect agreement (no discordant pairs)
Wearing gloves[a]	8	7	0.91	0.74–1.0

[a]One subject had missing data.

Kappa, for all of the potential risk factors which were evaluated. These findings are consistent with the large percentage of subjects reporting either being ''very confident'' ''or''certain'' in their memory of the number of hours exposed to these factors during the past month.

There is a growing body of evidence which suggests a high degree of reliability of information obtained using telephone-based interviews examining injury related risks and behaviors [Koziol-McLain et al., 2000]. Data on the frequency and duration of usual work tasks obtained from telephone interviews 2 to 3 weeks apart have also been shown to be reliable. However, these same subjects had difficulty in remembering the frequency and duration for unusual work tasks [Kallio et al., 2000]. Recalling routine work tasks and rarer transient work related exposures might not be the same phenomena. Unusual or occasional work tasks occurring once per week or per month may not be as reliably recalled as unusual transient exposures associated with work tasks (e.g., unusual performing equipment or using an unusual work method to do a task). In the current study, even transient

TABLE V. Confidence in the Accuracy of Memory of Potential Risk Factors for the Number of Hours of Exposure During the Past Month

Potential risk factor	N (%) of subjects reporting very confident or certain in their memory	Confidence score (Mean ± SEM)	Correlation (*P*-value) with reliability of the subjects reported usual frequency of exposure[a]
Equipment/materials performed differently[b]	22 (81.5)	4.07 ± 0.13	0.18 (0.38)
Different work method[b]	23 (85.2)	4.00 ± 0.11	−0.02 (0.92)
Unusual work task[b]	20 (74.1)	3.96 ± 0.14	−0.41 (0.03)
Distracted[c]	25 (89.3)	4.18 ± 0.12	0.11 (0.58)
Rushing[c]	25 (89.3)	4.11 ± 0.13	0.16 (0.41)
Working overtime	29 (100.0)	4.27 ± 0.08	0.06 (0.76)
Feeling ill[b]	25 (92.6)	4.22 ± 0.11	−0.11 (0.58)
Wearing gloves[c]	28 (100.0)	4.43 ± 0.10	−0.01 (0.98)

Note: No subjects reported that they were not at all confident.
Confidence scores: 1 = not at all confident, 2 = somewhat confident, 3 = average confidence, 4 = very confident, and 5 = certain.
[a]Spearman rank correlation coefficient of memory score and the absolute value of the difference between the subjects reported usual frequency of exposure at the initial and follow-up interviews.
[b]Two subjects had missing data.
[c]One subject had missing data.

exposures that occur about once per month on average, could be recalled reliably within 4 days of the first interview (Table III).

There are studies in the literature, which suggest that expressing a high level of confidence in the accuracy of memory is associated with more reliable recall. Some [e.g., Brigham, 1990], but not all studies [e.g., Chandler, 1994], have reported a positive correlation between confidence ratings and accuracy. If a high degree of confidence in the accuracy of recall does correlate with reliable recall this would suggest a high level of accuracy in memory for transient exposures. In the present study, it was found that the subjects' degree of certainty in the accuracy of their recall did not often correlate with more reliable recall. For example, workers with less reliable recall of the amount of time they were distracted, were not more likely to report higher levels of confidence in their recall of this exposure. In this case, there was a slight positive correlation ($R = 0.11$; P-value $= 0.58$) between the absolute value of the difference in the usual frequency of exposure reported at the initial and follow-up interviews and the subjects reported confidence in their recall. However, a number of factors could explain this discrepancy including a shift in the criterion used to define "distraction." In contrast, for performing an unusual work task, workers with more reliable data were more likely to report a high degree of confidence in their recall. In this case, there was a significant negative correlation ($R = -0.41$, P-value $= 0.03$) between the absolute value of the difference in the amount of time that the subjects reported performing an unusual work task over the past month on the first and second interview and their degree of confidence in their recall. This finding would support the hypothesis that high levels of confidence in the accuracy of memory are associated with more reliable recall. Unfortunately, the sample size in this study is too small to fully evaluate these findings.

Although our sample size was relatively small, it was sufficiently powered to evaluate the primary outcome variables, the reliability of the self-reports of the usual frequency of exposure to transient risk factors. However, the prevalence of four of the eight transient occupational exposures was low (e.g., percent exposed no greater than 10%) at the time of the injury. This limited our ability to further examine the reliability of these transient factors among different subgroups. For example, it would be of interest to compare the reliability of self-reports by job experience. In addition, the kappa statistic can easily be misinterpreted when exposures being compared are relatively rare [Maclure and Willett, 1987]. This is reflected by the relatively wide 95% CI for those exposures, which were most rare (e.g., the use of equipment or materials which performed differently than usual; 95% CI = 0.01–1.0) in this study. However, for exposures, which were more common, the confidence intervals were relatively narrow (e.g., rushing; 95% CI = 0.77–1.0).

The high degree of reliability of self-reported exposure to transient risk factors observed in this study may not be generalizable to studies of other occupational injuries using similar data collection instruments. The subjects in this study experienced only minor to moderately severe hand injuries, as measured by the type of injury, laceration length, and number of sutures. Thus, initial interviews with these subjects were conducted, for the most part, within 2 days of the injury. For more severe injuries, such as amputations or fractures, it may not be feasible to interview subjects as rapidly and the reliability of reports of exposures over longer time intervals is unknown. An extension of this study would be an examination of the reproducibility of potential risk factors for a period after 7 days (longer-term) after the initial interview. This would help to determine the feasibility of using a similar data collection instrument where the interview could not take place as close in time to the injury.

Our study suggests that transient workplace exposures, which occur around the time of an injury, can be reliably recalled up to 4 days after an initial interview. This is important since the case-crossover study design attempts to answer the question, "*what was different or unusual prior to the injury*?" [Maclure, 1991]. As most occupational injury research utilizes self-reported exposures, it is important to continue to develop methods to assess the reliability and validity of these data. This will ultimately allow for a better understanding of the etiology of workplace injuries and may lead to significant improvements in the design of strategies to prevent these injuries.

ACKNOWLEDGMENTS

The authors thank Dr. Philip Nasca, Dr. Kenneth Mundt, Dr. William Shaw, and Raymond McGorry who provided excellent comments on earlier drafts of this manuscript. The authors are grateful to Patti Boelsen for her help with the data entry and Kate Dwyer and Judith Ouelette for conducting the subject interviews. The preliminary results of this study were presented as a work in progress at the Consortium for Injury and Disability Research (CIDR). This work was supported in part by NIOSH Grant R01-OH-03763-03.

REFERENCES

Brigham J. 1990. Target person distinctiveness and attractiveness as moderator variables in the confidence-accuracy relationship in eyewitness identifications. Basic Appl Soc Psychol 11:101–115.

Chandler C. 1994. Studying related pictures can reduce accuracy, but increase confidence, in a modified recognition text. Mem Cognit 22: 273–380.

Conover WJ. 1980. Practical nonparametric statistics. 2nd ed. New York: John Wiley & Sons, Inc.

Fleiss JL. 1975. Measuring agreement between two judges on the presence or absence of a trait. Biometrics 31:651–659.

Fleiss JL. 1981. Statistical methods for rates and proportions. 2nd ed. New York: John Wiley & Sons, Inc.

Kallio M, Viikari-Juntura E, Hakkanen M, Takala EP. 2000. Assessment of duration and frequency of work tasks by telephone interview: Reproducibility and validity. Ergonomics 43:610–621.

Koziol-McLain J, Brand D, Morgan D, Leff M, Lowenstein R. 2000. Measuring injury risk factors: Question reliability in a statewide sample. Injury Prevention 6:148–150.

Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. Biometrics 33:159–174.

Maclure M. 1991. The case-crossover design: A method for studying transient effects on the risk of acute events. Am J Epidemiol 133:144–153.

Maclure M, Willett WC. 1987. Misinterpretation and misuse of the kappa statistic. Am J Epidemiol 126:161–169.

NIOSH National Occupational Research Agenda (NORA). 1998. Report on traumatic occupational injury research needs and priorities. US Department of Health and Human Services, Public Health Service, CDC. DHHS (NIOSH) Publication No. 98–134.

Rosenthal R, Rosnow RL. 1984. Essentials of behavioral research: Methods and data analysis. New York: McGraw-Hill.

Shrout PE, Fleiss JL. 1979. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 86:420–428.

Sorock GS, Lombardi DA, Hauser R, Eisen E, Herrick R, Mittleman MA. 2001. A case-crossover study of occupational traumatic hand injuries: Methods and initial findings. Am J Ind Med 39:171–179.