METHODOLOGIC ISSUES

A comparison of information on motor vehicle crashes as reported by written or telephone interviews

A Alonso, S Laguna, M Seguí-Gómez

Injury Prevention 2006;**12**:117–120. doi: 10.1136/ip.2005.009134

See end of article for authors' affiliations

Correspondence to: Dr M Seguí-Gómez, Department of Preventive Medicine and Public Health, School of Medicine, University of Navarra, Irunlarrea, 1 31080 Pamplona, Spain; msegui@unav.es

Accepted 5 December 2005

Objective: To compare information about traffic crash injuries and kilometers driven reported in a written questionnaire with information reported in a telephone interview.
Design: Telephone and paper surveys.
Setting: The *Seguimiento Universidad de Navarra* (SUN, University of Navarra Follow-up) study, in Spain. The SUN study is an open enrollment cohort study with 17 000 enrolled graduates followed through biennial mailed questionnaires.
Subjects: A sample of 542 individuals from the SUN study participants.
Main outcome measure: Agreement on information about traffic crash injuries and mileage driven in a mailed questionnaire and a telephone survey.
Results: Participation was 90.4%. Considering the phone survey as the gold standard, data on traffic crash injuries in the mailed questionnaire had 83% sensitivity (95% CI 77% to 89%), 77% specificity (95% CI 71% to 82%), 74% positive predictive value (95% CI 67% to 80%), and 89% negative predictive value (95% CI 83% to 93%). Agreement beyond chance, measured by the kappa statistic, was 0.63 (95% CI 0.56 to 0.70). Correlation between questionnaire and telephone surveys and kilometers driven on average during a year assessed by the intraclass correlation coefficient was 0.64 (95% CI 0.57 to 0.70), $p<0.001$.
Conclusions: Information on sustained traffic crash injuries and traveled mileage over the previous two years as reported through mailed questionnaires in a highly educated population could be used in the study of associations between traffic crash injuries and a variety of risk factors.

Measurement error is an ever present problem in epidemiology, particularly important when relying on self reported data. Despite an emergent literature devoted to the study of injuries based in mainly telephone or mail surveys (used in national or regional health surveys and in cohort studies),[1–5] issues regarding validity or reproducibility of this data are unresolved.

In 2000, our group began a multi-purpose cohort study in Spain, the Seguimiento Universidad de Navarra (SUN).[6] One of its objectives was the assessment of risk factors associated with the occurrence of motor vehicle related injuries. The recruitment of participants and their follow up is through mailed questionnaires. There is scarcity of studies assessing the reliability of self reported information in motor vehicle related injuries,[7] and the reliability of self reported mileage driven has been seldom studied.[8] Moreover, validity or reliability of self reported motor vehicle related injuries has never been assessed in Spain.

The objective of this work was to evaluate the reliability of information on motor vehicle related injuries and mileage driven among SUN study participants.

METHODS

The SUN study

Detailed description of the SUN has been published elsewhere.[6] Briefly, SUN is an open enrollment cohort in Spain with currently more than 17 000 university graduates (from the University of Navarra and other Spanish universities, as well as members of different professional associations), recruited and followed up through biennial mailed questionnaires. Even though the main campus of the University of Navarra is located in Pamplona, Spain, participants in the SUN study are scattered all over Spain.

Among other objectives, this cohort is designed to estimate the incidence of motor vehicle related injuries requiring hospitalization and motor vehicle related injuries needing time out of work, as well as to assess risk factors for those injuries—most importantly, exposure to road traffic. Recruitment of participants began in 2000 and is still ongoing. Retention rate for the first follow up questionnaire is higher than 90%.

By July 2003, the first 5425 follow up questionnaires from our eligible participants (that is, those who had entered the cohort two years earlier) were recruited and processed.

Reliability study

Information about previous traffic crash injuries and average annual mileage (in kilometers) was gathered both at the baseline and at the first follow up questionnaires. Specifically, the injury questions in the follow up questionnaire were ''Since you answered the first questionnaire in this study, have you suffered one of these circumstances: (1) a traffic crash injury requiring hospitalization of at least 24 hours (yes/no, if yes, tell us month and year); (2) other traffic crash injury not requiring hospitalization (yes/no, if yes, tell us month and year)''. Mileage driven was assessed by the following question: ''How many kilometers do you travel on average per year?'' The possible categories for answering this question were: <1500 km, 1501–5000 km, 5001–10,000 km, 10 001–20 000 km, and >20,000 km.

To evaluate the reproducibility (that is, reliability) of these variables, a telephone survey was conducted among 542 subjects, including all those who had responded affirmatively to the question of having had a motor vehicle crash with or without injuries requiring hospitalization of 24 hours or more and a 10% random sample of those who had not done so.

Three trained female interviewers blinded to the subjects' responses located the individuals and conducted telephone interviews during the summer of 2003. The complete telephone interview lasted an average of 15 minutes.

Table 1 Baseline characteristics of participants in the SUN study with two year follow up data by July 2003, participants in the reliability study, and differences between respondents and non-respondents to the reliability study

Variable	SUN study	Reliability study			
	n = 5425 (100%)	All, n = 542 (100%)	Respondents, n = 490 (100%)	Non-respondents, n = 52 (100%)	p Value*
Traffic crash injury requiring hospitalization for more than 24 hours	12 (0.2)	12 (2.2)	12 (2.4)	0 (0.0)	0.62
Traffic crash injury not requiring hospitalization	253 (4.7)	253 (46.7)	218 (44.5)	35 (67.3)	0.002
Age (years)					0.25
<40	3456 (63.7)	353 (65.2)	316 (64.5)	37 (72.5)	
⩾40	1969 (36.3)	188 (34.8)	174 (35.5)	14 (27.5)	
Females	3383 (62.4)	296 (54.6)	263 (53.7)	33 (63.5)	0.18
Marital status					0.27
Single	2879 (53.4)	310 (58.8)	274 (57.7)	36 (69.2)	
Married	2377 (44.1)	202 (38.3)	187 (39.4)	15 (28.8)	
Others	132 (2.5)	15 (2.9)	14 (2.9)	1 (2.0)	
Seatbelt use (yes)					0.30
Never	100 (1.8)	11 (2.1)	11 (2.3)	0 (0.0)	
Sometimes	1016 (18.8)	110 (20.7)	96 (20.0)	14 (26.9)	
Always	4282 (79.2)	410 (77.2)	372 (77.7)	38 (73.1)	
Airbag use (yes)	2160 (43.5)	242 (49.4)	212 (48.1)	30 (61.2)	0.08
Do you drive if you have had alcohol?					0.92
No, I don't have a driver license	522 (9.7)	39 (7.4)	34 (7.2)	5 (9.6)	
Yes, sometimes	1594 (29.6)	180 (34.2)	162 (34.1)	18 (34.6)	
Almost never	941 (17.5)	101 (19.2)	92 (19.4)	9 (17.3)	
Never	2328 (43.2)	207 (39.3)	187 (39.4)	20 (38.5)	

*Pearson's χ^2 test comparing respondents versus non-respondents.

Median time between the arrival of the two year follow up questionnaire and the telephone survey was 10 months. Questions on traffic crash injuries and kilometers driven in the telephone survey had the exact same wording as those in the mailed questionnaire. All questions were framed according to the period of time for which the written questionnaire was valid.

Statistical analysis

We assessed differences between respondents and non-respondents to the telephone interview using Pearson's χ^2. Then, for those who agreed to participate in the telephone interview (the respondents), we compared their answers in the written survey with the answers in the telephone survey. Given the low number of traffic crash injuries requiring hospitalization in the mailed questionnaire, we created a new variable combining them with injuries not requiring hospitalizations in all analyses. Reproducibility was assessed by computing percent agreement and Cohen's kappa statistic for injuries.[9] We also calculated sensitivity, specificity, and positive or negative predictive values for the mailed questionnaire, considering the telephone survey as the gold standard. Sensitivity was defined as the proportion of individuals who reported an injury in the telephone survey who also reported an injury in the mailed survey. Similarly, specificity was defined as the proportion of individuals who neither reported an injury in the telephone survey nor reported an injury in the mailed survey. On the other hand, positive predictive value was defined as the proportion of individuals who reported an injury in the mailed survey who also reported an injury in the telephone questionnaire. Negative predictive value was the proportion of individuals who neither reported an injury in the mailed survey nor reported an injury in the telephone questionnaire. Individuals with missing information in the mailed questionnaire were considered as false positives for the sensitivity computation and as false negatives for the specificity. In this way, our estimates were conservative. We computed rectangular 95% confidence intervals (CI) to take into account that diagnostic accuracy is described by sensitivity and specificity together (or both predictive values).[10] Rectangular 95% CIs correspond to standard 97.5% CIs. Additionally, we computed the area under the ROC curve as a global measure of the reliability of the information in the mailed questionnaire using a logistic regression model.

Time elapsed between the crash and the surveys could predict the probability of agreement between the surveys. To address this question, we first studied the association between an identical response in both surveys and time between them using a logistic regression analysis, where the outcome variable was concordance between both surveys and the main independent variable was time elapsed between them. Then, we did a second analysis with the same outcome variable, but where we considered as the main independent variable the time between the telephone survey and the self reported date of the injury. In both analyses, we included age and sex as potential confounders.

Agreement in the mileage driven between the mailed questionnaire and the telephone survey was computed using the intraclass correlation coefficient.[9] To estimate this statistic, we assigned the mean value for kilometers driven in each category of this variable.

Table 2 Agreement between telephone and mailed surveys for motor vehicle injury

		Telephone		
		Yes	No	Total
Mail	Yes	170	60	230
	No	28	219	247
	Blank	6	7	13
	Total	204	286	490

Table 3 Agreement between telephone and mail surveys regarding kilometers driven (car or motorbike) in the last year

		Telephone					
		<1500	1501–5000	5001–10000	10001–20000	>20000	Total
Mail	<1500	32	20	3	1	4	60
	1501–5000	19	38	20	16	8	101
	5001–10000	4	27	27	41	10	109
	10001–20000	5	6	17	44	41	113
	>20000	2	2	3	13	73	93
	Blank	2	2	1	2	7	14
	Total	64	95	71	117	143	490

RESULTS

A total of 490 individuals (90.4%) agreed to participate in the reliability study. Table 1 shows the characteristics of the participants in the SUN study and the selected sample for our study (by respondents and non-respondents). Non-respondents (mostly non-respondent because we could not locate them) were statistically significantly younger than respondents, but there were no other differences between the groups.

There were 265 participants who reported having suffered motor vehicle injuries in the follow up questionnaire (12 requiring hospitalization, 253 not requiring hospitalization). Thirty five of them (13.2%) did not participate in this reliability study (they were among the non-respondents), mostly because they could not be located.

For those who agreed to participate in the telephone survey, agreement beyond chance between written and telephone answers for this question measured by Cohen's kappa statistic was 0.63 (95% CI 0.56 to 0.70). Percent agreement was 83% (95% CI 78% to 85%). For information on traffic crash injuries, sensitivity was 83% (95% CI 77% to 89%), specificity was 77% (95% 71% to 82%), positive predictive value was 74% (95% CI 67% to 80%), and negative predictive value was 89% (95% CI 83% to 93%) (table 2). The area under the ROC curve was 0.82 (95% CI 0.78 to 0.86).

Table 3 shows the distribution of participants according to mileage driven reported in the mailed questionnaire and in the telephone survey. The intraclass correlation coefficient between written and telephone questions regarding mileage driven was 0.64 (95% CI 0.57 to 0.70, $p<0.001$).

There was no statistically significant association between time elapsed from the receipt of the two year follow up mailed questionnaire and the telephone interview and the probability of agreement (table 4). Using multiple logistic regression analysis, in which the dependent variable was the agreement between both questionnaires, none of the different variables considered (age, sex, and time between both surveys) was significantly associated with the probability of agreement.

When the effect of time between the telephone survey and the reported date of the crash related injury in the mailed questionnaire was the time related independent variable instead (table 4), time was significantly associated with the odds of agreement between both surveys ($p = 0.007$), with a longer time elapsed increasing the chance of disagreement.

Table 4 Mean time (months) elapsed between both surveys and between phone survey and reported date of injury in the mailed survey by agreement status between both surveys

	Concordant information, mean (SD)	Discordant information, mean (SD)	p Value*
Time elapsed between phone and mailed surveys	8.6 (4.1)	8.3 (4.1)	0.56
Time elapsed between phone survey and reported date of injury	20.3 (9.9)	24.0 (12.9)	0.007

*Obtained from a logistic regression model with agreement between both questionnaires as the outcome and time elapsed as the main independent variable, adjusting for age and sex.

DISCUSSION

We needed to evaluate whether to continue asking for motor vehicle related injuries and their possible risk factors in the SUN study as the four year follow up questionnaire had to be developed in the fall of 2003. Our analysis of the data available at the time suggested that, albeit imperfect, traffic crash injuries and mileage driven reported through mailed questionnaires in a highly educated Spanish population could be used in the epidemiological investigation of incidence and risk factors for traffic crash injuries.

Incidentally, although the number of injuries requiring hospitalization may seem low, the cumulative incidence for this outcome is similar to that observed in other studies conducted in Spain.[11]

Our results on the reliability of the data concur with the findings of other studies. A reliability study conducted in Colorado, US, in 1998, using the Behavioral Risk Factor Surveillance System, reported a substantial agreement (kappa = 0.80) in a question regarding motor vehicle crashes requiring medical treatment.[7] Similarly, in the Youth Risk Behavior Survey Questionnaire, questions related to unintentional and intentional injury behaviors had moderate to substantial reliability in every query,[12] though that study did not explore health outcomes.

Other studies have assessed the effect of recall period in the estimation of injury rates, showing that the longer the period the lower the injury rates. For example, in the Child Health Supplement to the National Health Interview Survey, the injury rates declined from 24.4% for a one month recall period to 14.7% for a 12 month recall period.[2] A comparable recall bias, especially for minor injuries, was apparent in parents of children younger than 6 in a health maintenance organization in Washington state.[13] Similarly, a study using data from the Occupational Health Supplement in the National Health Interview Survey concluded that rates were 32% higher for at-work injuries when considering recall bias.[3] Comparable results were obtained in the Health and Retirement Study,[4] and in two studies conducted in developing countries.[14 15] For motor vehicle crashes, only one previous report has assessed the effect of recall period on injury rates.[16] In our study, time between the mailed questionnaire and the telephone survey was not associated with disagreement in the answers. However, time elapsed between the telephone survey and the reported date of crash

related injury was associated inversely with the probability of agreement between both surveys.

The results presented here have been obtained from a population with a high educational level. Thus, we cannot directly apply this information to the general population nor was this our objective. Our aim was to assess the quality of self reported information in our cohort study population. And considering the relatively high participation in the reliability study, we feel we can confidently extend the results to the whole cohort.

The present analysis on traffic crash injuries combines answers regarding injuries requiring hospitalization and injuries not requiring so. The latter is an open definition that allows for some variability in the participants' reporting of their experience, which can introduce some imprecision but, on the other side, allows us to be very sensitive and detect a higher number of injuries.

Finally, an obvious limitation of this study is that the question about validity of information in mailed questionnaires is not answered. However, reliability is a necessary condition to validity. Consequently, this study is a preliminary step in the inquiry on the validity of outcomes and exposures in our cohort study.

It is not easy to assess the validity of motor vehicle related injury. In one such study conducted among older drivers in the US, agreement between self reported and state recorded crash involvement was only moderate (kappa = 0.45), and there were personal variables both associated with correct reporting of crash involvement and the actual risk of crash involvement, raising the possibility of biased risk estimates.[17] In another study, the Dunedin Multidisciplinary Child Development Study, the sensitivity of self reporting of injuries requiring hospitalization varied depending on the type of injury and time elapsed since the event and ranged from 86% (unintentional injuries) to 14% (self inflicted injuries). In this same study, the sensitivity of self reported traffic crashes was only 49%.[5]

Even more challenging is the validation of mileage traveled in a motor vehicle. Although this is a very important exposure for motor vehicle related injuries, we have only found one piece of work in which the reliability of mileage driven was explored.[8] The author reported a Pearson's correlation coefficient of 0.5, lower than the reliability observed in our study.

The validity of information on other health outcomes and risk factors gathered through mailed questionnaires has been shown to be high;[18] SUN participants are all university graduates, and our results seem to align with these other findings.

In conclusion, we have shown that in a highly educated population in Spain, self reported information on motor vehicle related injuries and mileage driven does not have a great degree of error, in line with other similar studies. Although this fact does not necessarily confer validity, we believe the quality is high enough to continue to work with the data obtained through mailed questionnaires.

ACKNOWLEDGEMENTS

Dr Alonso was partially supported by a Fulbright Fellowship. The authors are indebted to all the SUN study participants for their continued cooperation and participation. The Department of Health of the Navarre Regional Government, the Spanish Ministry of Health (Fondo de Investigaciones Sanitarias, project 01/0619 and 03/0678), and Seguros Zurich are gratefully acknowledged for partially funding the present study. We would like to thank María Ibarrola-Vidaurre and Beatriz Pérez-Urbina for their work as interviewers and Cristina Rubio for comments to an earlier version of this manuscript.

These results were partially presented at the 132nd Annual Meeting of the American Public Health Association, Washington, DC, USA, November 2004.

Key points

- There is a scarcity of studies assessing the reliability of self reported information regarding motor vehicle related injuries and mileage driven as a risk factor for motor vehicle crashes.
- In a university graduates cohort study in Spain, the degree of misclassification in self reported information about motor vehicle related injuries and mileage driven through mailed questionnaires was low enough for that information to be used in epidemiologic studies.
- Time elapsed between the telephone survey and the reported date of the crash injury was inversely associated with the agreement between both surveys.
- Additional studies are required to evaluate the validity of self reported information, although this task is difficult due to the absence of a perfect gold standard.

Authors' affiliations

A Alonso, S Laguna, M Seguí-Gómez, Department of Preventive Medicine and Public Health, School of Medicine, University of Navarra, Pamplona, Spain

A Alonso, Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA

M Seguí-Gómez, Department of Health Policy and Management and Center for Injury Research and Policy, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA

Competing interests: none.

REFERENCES

1 **Langley JD**, Cecchi JC, Williams SM. Recall of injury events by thirteen year olds. *Methods Inf Med* 1989;**28**:24–7.
2 **Harel Y**, Overpeck MD, Jones DH, *et al*. The effects of recall on estimating annual nonfatal injury rates for children and adolescents. *Am J Public Health* 1994;**84**:599–605.
3 **Landen DD**, Hendricks S. Effect of recall on reporting of at-work injuries. *Public Health Rep* 1995;**110**:350–4.
4 **Zwerling C**, Sprince NL, Wallace RB, *et al*. Effect of recall period on the reporting of occupational injuries among older workers in the Health and Retirement Study. *Am J Ind Med* 1995;**28**:583–90.
5 **Begg DJ**, Langley JD, Williams SM. Validity of self-reported crashes and injuries in a longitudinal study of young adults. *Inj Prev* 1999;**5**:142–4.
6 **Martínez-González MA**, Sánchez-Villegas A, de Irala-Estévez J, *et al*. Mediterranean diet and stroke: objectives and design of the SUN Project. *Nutr Neurosci* 2002;**5**:65–73.
7 **Koziol-McLain J**, Brand D, Morgan D, *et al*. Measuring injury risk factors: question reliability in a statewide sample. *Inj Prev*, 2000;6, 148–50.
8 **Hampel B**. Über die Zuverlässigkeit von Angaben zur jährlichen Fahrleistung. *Z Verkehrssich* 1978;**24**:166–70.
9 **Fleiss JL**. The measurement of interrater agreement. In: Fleiss JL, eds. *Statistical methods for rates and proportions*. New York: John Wiley and Sons, 1981:211–36.
10 **Pepe MS**. Measures of accuracy for binary tests. In: Pepe MS, eds. *The statistical evaluation of medical tests for classification and prediction*. New York: Oxford University Press, 2003:14–34.
11 **Plasencia A**, Borrell C, Anto JM. Emergency department and hospital admissions and deaths from traffic injuries in Barcelona, Spain. A one-year population-based study. *Accid Anal Prev* 1995;**27**:591–600.
12 **Brener ND**, Collins JL, Kann L, *et al*. Reliability of the Youth Risk Behavior Survey questionnaire. *Am J Epidemiol* 1995;**141**:575–80.
13 **Cummings P**, Rivara FP, Thompson RS, *et al*. Ability of parents to recall the injuries of their young children. *Inj Prev* 2005;**11**:43–7.
14 **Mock C**, Acheampong F, Adjei S, *et al*. The effect of recall on estimation of incidence rates for injury in Ghana. *Int J Epidemiol* 1999;**28**:750–5.
15 **Moshiro C**, Heuch I, Astrom AN, *et al*. Effect of recall on estimation of non-fatal injury rates: a community based study in Tanzania. *Inj Prev* 2005;**11**:48–52.
16 **Cash WS**, Moss AJ. Optimum recall period for reporting persons injured in motor vehicle accidents. *Vital Health Stat 1* 1972;**2**:1–33.
17 **McGwin G Jr**, Owsley C, Ball K. Identifying crash involvement among older drivers: agreement between self-report and state records. *Accid Anal Prev* 1998;**30**:781–91.
18 **Colditz GA**, Martin P, Stampfer MJ, *et al*. Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. *Am J Epidemiol* 1986;**123**:894–900.