British Journal of Industrial Medicine 1988;**45**:29–32

Validity of self reported work history

R BOURBONNAIS,[1] F MEYER,[2] G THERIAULT[3]

From the Groupe de Recherche sur l'Organisation du Travail,[1] la Santé et la Sécurité, and Département de Médecine Sociale et Préventive,[2] Faculté de Médecine, Université Laval, Quebec G1K 7P4, and School of Occupational Health,[3] Montreal H3A 1A3, Canada

ABSTRACT Many epidemiological studies of the relation between work and disease use information on work history obtained by interview from the study subjects. A validation study was undertaken to evaluate the accuracy of this information collected from 100 workers in a shipbuilding industry. The information furnished by the workers was compared with that present in the company's registers. The work history (job titles and starting dates) was relatively accurate and the validity varied with the number of events to declare and with their duration; it also depended on the type of information and the precision required.

The validity of any information expresses the extent to which it measures what it purports to measure.[1] Validity is difficult to assess in occupational studies because a source of true information is often lacking. Company records are sometimes incomplete or inaccessible. When people have worked for several employers, as is often the case, it is difficult to link the records of these companies. For these reasons, many epidemiological studies on the relation between occupational exposure and disease rely on working histories reconstructed from interviews with workers. The validity of such information is not well documented. The few studies on the validity of job histories have compared workers' interview data with the corresponding information obtained by interviewing employers.[2 3] Only one study attempted to confirm the employers' name provided by employees using an objective source of information, a government pension plan registry.[4] In the present study workers were interviewed about their occupational history during their employment at a single company. The validity of these data was assessed using the information recorded in the company registers as a reference.

Methods

The study population consisted of 100 male employees in a shipbuilding yard. This company was chosen because the employer kept a detailed administrative work record for each employee and because workers were mobile within the company and could have held several jobs at various times during their employment. One hundred and fifty men with at least five years' employment were randomly selected from the company's 1750 active employees at the time of the study. As a secondary objective of the study was to assess the accuracy of occupational history obtained from next of kin seven workers with no respondent were excluded. Forty three subjects refused to participate. One hundred workers gave informed consent and constituted the final study population.

Trained interviewers met the subjects at home and obtained information on personal characteristics and occupational history at the shipyard. Workers were requested to identify all jobs held for at least six months, beginning chronologically with their present job, and to provide the starting and finishing dates of each working period. The detailed characteristics of each job were recorded in order to prevent miscoding job titles because of the different ways in which they are named. Subjects were unaware that the study entailed a check in the company's registers.

The corresponding information on job titles and dates was independently abstracted from the employer's administrative records. For each worker, all working periods of at least six months' duration were cumulated for each job title. The present report concerns the main job, defined as the one held by each employee for the longest cumulative period according to the employer's records.

One of the authors (RB) independently coded all the information from the workers' interviews and from the company records. The company's coding numbers were used to identify job titles in both interview and record data. Identification of the main job title by the workers was considered valid when the code number was identical to that in the company records. For dates, the information was judged valid

Accepted 9 March 1987

Table 1 *Number of workers and percentage of valid information on title and starting date of their main job according to personal and employment characteristics*

		No	Title	Starting date
Age	24–39	35	80	77
	40–49	30	93	70
	50–67	35	94	40**
Education (years)	1–7	46	87	50
	8–13	54	91	72*
No of jobs	1	40	100	85
	2	31	90	48
	3–5	29	72**	45**
No of years since starting date	2–9	28	86	89
	10–19	27	81	59
	20–29	36	97	50
	30–39	9	89	33**
Duration (years)	1–4	21	76	81
	5–9	20	80	60
	10–14	17	94	59
	15–36	42	98*	55

*p < 0·05; **p < 0·01.

when the difference in time between the date given by the subject and the corresponding date in the company records was 12 months or less.

Results concerning the title and the starting date of the main job were expressed as percentages of valid answers. These percentages were analysed according to workers' characteristics—age, education, number of jobs, time elapsed since the beginning of the main job, and duration of employment in the main job. Statistical significance was assessed by chi-square and Fisher exact tests.[5] The simultaneous effects of the potential determinants of the validity were examined by logistic analysis.[6] The analyses were supported by Statistical Analysis Systems Inc.

Results

The distribution of the subjects by personal and employment characteristics is presented in table 1 (first column). All workers provided information on title and starting and ending dates of their main job. Data pertaining to ending date are not presented as 80% of the subjects were still employed in their main job at the time of the study. The percentage of valid answers was 89% for the job title and 62% for the starting date of the main job. The proportion of valid information according to potential determinants of validity is shown in table 1. Job title was better identified by older workers and by those (often the same subjects) who had held their job for a longer period. Job starting date was more accurately reported by younger, more educated employees and by those (often the same subjects) who had started working more recently or had held their job for a shorter period. The validity of the information for both title and starting date decreased with the number of jobs held.

Multiple logistic regression was performed to untangle the influence of the above potential determinants of validity. The variables were introduced into the logistic model as categorised variables (with the categories presented in table 1) in a stepwise fashion. Only significant determinants of validity were kept in the final models. The results of these analyses are presented in table 2. The number of jobs was inversely related to the validity of the information on job title and starting date. Education was positively associated with the validity of more elaborate information (date) but did not influence the validity of the information pertaining to job title. Time elapsed since the beginning of the main job was related positively to the validity of data on job title but negatively to that on starting data. Neither age nor duration of employment in the job were significant determinants of validity.

Table 2 *Parameter estimates (β) and p values from logistic analysis of validity of information provided by 100 workers on title and starting date of their main job**

	Title		Starting date	
	β	p	β	p
Constant	3·31	0·003	0·89	0·430
No of jobs	−1·13	0·001	−0·58	0·009
Time since beginning	0·11	0·032	−0·06	0·025
Education	—	—	0·26	0·009

*Stepwise logistic regression analyses with entry and exit levels at p = 0·05.

Discussion

The validity of self reported work history assessed in this study depended on the complexity and remoteness of the information sought. Job titles were more accurately identified by workers than starting dates. The number of jobs held, the time elapsed since the beginning of the job to the report, and the level of education of the subjects were determinants of the validity of the occupational history.

The present study did not cover the occupational history of workers in its entirety since we had access only to the one company's records. The great mobility of employees within this company, however, permitted the collection of information on many jobs, which rendered the occupational history sought closer to a life working history. Data were thus collected on all the jobs held in the company and their starting and ending dates. Because of the wide range of the information gathered, we first restricted our analysis to a small portion of the data collected, the one concerning the main job. As the major concern in

occupational studies is about jobs held for the longest period, to assess a main exposure, the present findings concerning the validity of the information about the main job in the enterprise under study could be readily interpreted.

In a previous investigation the information obtained by interviewing 50 individuals was compared with that provided by their employers on a questionnaire.[2] The percentage of agreement between workers and employers was 90% for job duties, 76% for job titles, and 71% for duration of employment. By comparison, we obtained a higher validity for the information provided by the employees concerning their main job: 89% for job title, 76% for starting date.

Keating *et al* used employers' interview data to validate the information on job duties, job duration, and wages provided by 236 unemployed individuals.[3] Data from the two sources were strongly correlated. Although employers had access to company records to answer questions on employees' work history, their being an intermediary between the records and the researchers was a potential source of bias. In our study the risk of information bias was reduced by abstracting the work history directly from the company records on an independent sheet from the interview data, thus making the collection blind. As the information to validate (job titles and dates) was not controversial, as would be the level of exposure to a toxic agent, the data kept by the company for administrative purposes were considered to be accurate.

A recent study in Montreal validated the information obtained from 297 workers by interview, using the Quebec Pension Plan records as a reference.[4] The results, presented in person-years of concordance between the two sources, showed 82% agreement on employers' names declared for each of the 13 years of the study. The authors found a higher degree of agreement for the workers who had held two or more jobs. They found no significant effect of age, education, or family income on the accuracy of report. In our study a similar association was observed between the number of jobs held and the validity of the information on work history. Nevertheless, education was also identified as a significant determinant of the validity of the starting date which is a more elaborate piece of information to remember.

Many studies have outlined the importance of memory as a distorting factor in reporting past events.[7-10] The more distant the events are in time, the less valid is the information reported. The importance of the events in terms of personal or social significance, however, counterbalances the effect of time on memory or the capacity to report accurately these events.[11]

Keating *et al* compared the reports of occupational events between the immediate and more distant past, and concluded that validity remained as high for jobs held up to six years before the interview as for jobs held just before it.[3] Baumgarten *et al* evaluated the influence of the distance in time of the occupational events on the accuracy of their report.[4] No difference was observed in the validity of information of recent (2–8 years) and remote (9–15 years) events.

In our study distance in time makes a significant contribution after controlling for the number of jobs held and education. This contribution affects the validity of the data in a different manner for job titles and for starting date. It is related positively to the validity of job title and inversely for the starting date. It seems that the importance of the event for the individual has a greater effect on the capacity to remember less detailed information such as job title or employers' names rather than more complex details such as dates.

Assessment of occupational exposure of workers in retrospective studies is difficult, often complicated by the unavailability of company records. The present study suggests that the workers themselves can provide valid information, especially when it pertains to job titles and time events related to their main job. Nevertheless, the validity of data obtained from employees on their occupational history could be improved by providing them (at time of interview) with a historical account of the principal events in the company, together with dates of technological or organisational changes. These time references would probably facilitate the recall of events and related dates. Such historical accounts could be reconstructed from various sources: company and union records and interviews with management and workers.

We thank Mrs S Gingras, département de médecine sociale et préventive, Université Laval, for her help with the data processing; Mr P M Bernard for his statistical advice; and the workers and management of the shipbuilding enterprise for their participation in this study.

The data presented in this paper formed part of an MSC dissertation submitted to Université Laval by RB and supervised by GT. It was supported by a grant from l'Institut de Santé et de Sécurité du Travail du Québec.

References

1 Last JM. A dictionary of epidemiology. New York: Oxford University Press, 1983.
2 Weiss DJ, Davis RV. An objective validation of factual interview data. *J Appl Psychol* 1960;**44**:381–5.
3 Keating E, Paterson DG, Stone CH. Validity of work histories obtained by interview. *J Appl Psychol* 1950;**34**:6–11.
4 Baumgarten M, Siemiatycki J, Gibbs G. Validity of work histories obtained by interview for epidemiologic purposes. *Am J*

Epidemiol 1983;**118**:583–91.
5 Fleiss JL. *Statistical methods for rates and proportions*. New York: John Wiley, 1973.
6 Kleinbaum DG, Kupper LL. *Applied regression analysis and other multivariate methods*. Boston: Duxbury Press, 1978.
7 Cannel CF, Marquis KH, Laurent A. *A summary of studies of interviewing methodology*. (Data evaluation and methods research, series 2, No 69.) Washington: US GPO, 1977. (PHS publ No 77-1343.)
8 Parry HJ, Crossley HM. Validity of responses to survey questions. *Public Opinion Quarterly* 1950:61–80.
9 American Statistical Association. Studies of underreporting of health events in the household interview. *Proceedings of the American Statistical Association* 1980;**69**:4–17.
10 Sudman S, Bradburn NM. *Response effects in surveys*. Chicago: Aldine Publishing Co, 1974.
11 Ridley JC, Bachrach CA, Dawson DA. Recall and reliability of interview data from older women. *J Gerontol* 1979;**34**:99–105.

Vancouver style

All manuscripts submitted to the *Br J Ind Med* should conform to the uniform requirements for manuscripts submitted to biomedical journals (known as the Vancouver style)

The *Br J Ind Med*, together with many other international biomedical journals, has agreed to accept articles prepared in accordance with the Vancouver style. The style (described in full in *Br Med J*, 24 February 1979, p 532) is intended to standardise requirements for authors.

References should be numbered consecutively in the order in which they are first mentioned in the text by Arabic numerals above the line on each occasion the reference is cited (Manson[1] confirmed other reports[2–5]...). In future references to papers submitted to the *Br J Ind Med* should include: the names of all authors if there are six or less or, if there are more, the first three followed by *et al*; the title of journal articles or book chapters; the titles of journals abbreviated according to the style of *Index Medicus*; and the first and final page numbers of the article or chapter.

Examples of common forms of references are:

1 International Steering Committee of Medical Editors. Uniform requirements for manuscripts submitted to biomedical journals. *Br Med J* 1979;**1**:532-5.
2 Soter NA, Wasserman SI, Austen KF. Cold urticaria: release into the circulation of histamine and eosino-phil chemotactic factor of anaphylaxis during cold challenge. *N Engl J Med* 1976;**294**:687-90.
3 Weinstein L, Swartz MN. Pathogenic properties of invading micro-organisms. In: Sodeman WA Jr, Sodeman WA, eds. *Pathologic physiology: mechanisms of disease*. Philadelphia: W B Saunders, 1974:457-72.