
Occupational Exposure Assessment in Case-Control Studies: Opportunities for Improvement [with Commentary]
Author(s): K. Teschke, A. F. Olshan, J. L. Daniels, A. J. De Roos, C. G. Parks, M. Schulz, T. L. Vaughan, H. Kromhout
Source: *Occupational and Environmental Medicine*, Vol. 59, No. 9 (Sep., 2002), pp. 575-594
Published by: BMJ Publishing Group
Stable URL: http://www.jstor.org/stable/27731761
Accessed: 28/07/2010 09:18

REVIEW

Occupational exposure assessment in case–control studies: opportunities for improvement

K Teschke, A F Olshan, J L Daniels, A J De Roos, C G Parks, M Schulz, T L Vaughan

Occup Environ Med 2002;**59**:575–594

Community based case–control studies are an efficient means to study disease aetiologies, and may be the only practical means to investigate rare diseases. However, exposure assessment remains problematic. We review the literature on the validity and reliability of common case–control exposure assessment methods: occupational histories, job–exposure matrices (JEMs), self reported exposures, and expert assessments. Given the variable quality of current exposure assessment techniques, we suggest methods to improve assessments, including the incorporation of hygiene measurements: using data from administrative exposure databases; using results of studies identifying determinants of exposure to develop questionnaires; and where reasonable given latency and biological half life considerations, directly measuring exposures of study subjects.

See end of article for authors' affiliations

Correspondence to:
Dr K Teschke, Department of Health Care and Epidemiology, University of British Columbia, Vancouver, BC, V6T 1Z3, Canada; teschke@interchange.ubc.ca

Accepted
13 February 2002

Community based case–control studies remain one of the most efficient epidemiological study designs, especially for investigating the aetiologies of rare diseases. For certain extremely low incidence outcomes, such as childhood cancers, case–control studies may be the only viable study method. In comparison to cohort studies, the other most common design used in occupational epidemiology, exposure assessment in case–control studies offers certain advantages, but also poses major challenges.

For exposures which occur in widely dispersed segments of the population, a population based case–control design theoretically allows examination of the broadest possible range of exposure levels, though the prevalence of exposure to most agents is likely to be low. When the exposed individuals are scattered in small worksites (for example, farmers), a case–control study centred in a geographical area where these workers reside may be logistically simpler than assembling a cohort. Perhaps most importantly, case–control studies offer the opportunity to enumerate multiple exposures, including occupational and residential exposures throughout a subject's lifetime, as well as medical and lifestyle factors that may confound or modify an exposure–disease association. Information on such a broad range of exposures is generally not available in industry based cohort studies.

Despite these advantages, exposure assessment remains one of the most problematic elements of case–control studies. Exposure data are usually gathered by interviewer administered questionnaires, or occasionally from mailed questionnaires, medical records, or vital statistics. Exposures ascertained using these sources are almost never quantitative measurements, but subject or proxy reported job histories, tasks, or recalled exposures to specific agents. On occasion, expert judgement is used to infer exposures from job histories, or to review and modify exposure self reports. The merit of these methods is therefore an essential consideration in the interpretation of study results.

The purpose of this paper is to review evidence about the validity and reliability of qualitative or semiquantitative exposure assessment techniques commonly used in case–control studies, with the aim of identifying means to optimise these methods. In addition, we will discuss some opportunities for greater quantification of exposures in case–control studies.

METHODS

A number of methods were used to gather the literature. The Medline database was searched from 1966 to April 2001 using the following terms: validity, reliability, sensitivity, specificity, agreement, kappa, intraclass, reproducibility of results, expert, subjective estimate, self-report, exposure estimate, semiquantitative estimate, qualitative estimate, or job–exposure matrix. Search results were limited using the following terms: occupation, hygiene, work, job, industry, or occupational exposure. All English and French abstracts and/or titles were reviewed for relevance.

There is little standardised terminology for identifying the literature on validity and reliability of exposure assessment methods. Therefore more manual search methods were also used, including a review of the citations in identified articles and the publications resulting from four international initiatives on assessment of occupational exposures in epidemiology: a conference in Woods Hole, USA in 1988 (reported in Rappaport SM, Smith TJ, eds, *Exposure Assessment for Epidemiology and Hazard Control*, Chelsea, MI: Lewis Publishers, 1991); a conference in Leesberg, USA in 1990 (reported in *Applied Occupational and Environmental Hygiene* 1991;**6**:417–558); a European concerted action (reported in *International Journal of Epidemiology* 1993;**22**(suppl 2):S1–S133); and a conference in Lyon, France in 1994 (reported in *Occupational Hygiene* 1996;**3**:1–208). Stewart and Dosemeci's bibliography of exposure assessment literature[1] was also consulted, as were two review articles on exposure assessment in case–control studies.[2 3]

This review does not include studies of the following issues: proxy reporting, questionnaire delivery methods, ergonomic or work organisation exposures, and industry specific job–exposure matrices developed for cohort studies or industry based nested case–control studies.

The paper begins with a review of the most common exposure assessment methods used in population based case–control studies: subject reported occupational histories; use of occupational histories to infer exposure (that is, job–exposure matrices); self reported exposures; and expert assessment of exposures. It then examines additional methods which should allow more quantification of exposure: using measurements from exposure databases; using determinants of exposure studies to design exposure questionnaires; and measuring exposures among study subjects. Some of the terminology of validity and reliability studies is briefly described in the appendix.

OCCUPATIONAL HISTORIES

Collection of data on each subject's employment history, including product manufactured or service provided, job title, and usual duties, has become a routine part of many population based case–control studies using questionnaires. Studies using medical records, birth or death certificates, or other administrative data sources also usually include information on at least one job, often the most recent or usual job. Data on occupation and industry, whether from medical records or questionnaires, are usually derived from self reports or, when a subject is dead or in some way incapable, reports by next of kin.

A number of studies[4–16] have examined the validity of self reported occupational histories by comparisons with company, pension, or union records; others have examined reliability by comparisons to previous self reports (table 1). Validity and reliability studies report rather consistent results, with levels of raw agreement for employer, job classification, person-years in a job, and start and termination dates generally in the range of 70–90% and with kappas from 0.65 to 0.82.[5 7–16] Some studies within single industries found lower agreement on the number of work area assignments (50.6%),[10] job title held longest (67%),[8] and starting date (62%),[11] perhaps because there may be minor distinctions between jobs within a company that are difficult to elicit by questionnaire.

The reliability and validity of occupational histories have also been tested by examining whether they can be used to accurately assign exposures. Rosenberg[17] examined the reliability of estimates of cumulative PCB exposure based on occupational histories taken first in 1976, then again in 1979. Average measured exposure in each job was cumulated using the two job histories and the results compared: the intraclass correlation was 0.94 for early jobs and 0.96 for jobs in the most recent 10 years. Birdsong and colleagues[18] assessed the validity of solvent exposure assignments based on self reported job histories by comparisons to those based on personnel records, and found that 99% of subjects were correctly classified as exposed or unexposed, but that the correlation between measures of exposure duration was only moderate (Pearson $r = 0.63$).

True validities of self reported jobs are likely to be somewhat higher than measured, as the reference standards are not likely to be true gold standards.[5] For example, personnel records may not reflect changes in the tasks an employee performs if the title or pay has not changed. Conversely, human resources personnel may record a change in job title, when the functional characteristics of a job may be unaltered. In addition, jobs may simply be labeled differently in administrative records than by employees. Reliability studies should avoid problems with job title terminology, because they test recall of a person's own way of describing a job.

Two reliability studies[12 14] raise the parallel issue of job coding: even if the job histories are well reported by study subjects, the way that research staff codes each job can affect their exposure group assignment. Wärneryd and colleagues[14] found the worst agreement for difficult to code clerical and administrative jobs. Kennedy and colleagues[19] found that errors in coding jobs were responsible for reducing an odds ratio for asthma of 1.5 to 1.0, because a job's potential for exposure to known allergens could not be properly classified when incorrectly coded.

Factors consistently found to reduce validity and reliability of occupational histories include increasing complexity of a subject's occupational history, shorter duration of a job, and longer period of recall.[5–7 10 11 13 14 17 18] Other factors, such as age, race, language, and education had either little or no association with recall.[5 7 8 13 18] Two studies were able to check for differences in validity of job reporting between cases and controls, and found no evidence to suggest recall bias.[5 10]

Given the reasonable quality of self reported occupational histories, epidemiological analyses by occupation and industry are likely to be useful initial steps towards the identification of hazardous exposures. Where exposures to complex mixtures are of interest, an industry or occupation may be an appropriate way to represent the combined exposures. The main shortcoming of analyses by occupation and industry is that they do not identify specific agents as risk factors. For example, painters may be exposed to solvents, but they also have varying potential for exposures to other agents, such as metals, pesticides, isocyanates, epoxies, wood dust, formaldehyde, and silica. In addition, although most painters use solvents, some may not. An increased risk in a job can only be suggestive of risks from particular agents. In addition, the lack of an association with a particular job may mask the effect of an agent to which only some individuals in the job are exposed.

EXPOSURE MATRICES: USING JOBS TO INFER EXPOSURES

In an effort to use the reasonably accurate recall by subjects of their occupational histories, but overcome the indirect connection to exposures, there was a movement in occupational epidemiology starting in the 1980s to develop job–exposure matrices (JEMs). JEMs list a wide range of occupations and/or industries on one axis, a wide range of exposure agents on the other, and the cells of the matrix indicate the presence, intensity, frequency, and/or probability of exposure to a specific agent in a specific job. In some JEMs, calendar period may form a third axis of the matrix. Industry based cohort studies have long used this matrix format for assigning exposures to cohort members' job histories within a company; the new idea was to create JEMs which could describe exposures across the range of jobs and industries that might be observed in a population based study.

Several such JEMs, using European or American occupation and industry coding systems, have been made publicly available (hereafter, these are called "generic" JEMs, in contrast to study specific JEMs). Some were created using expert judgement, usually aided by published literature and communication with industry personnel[19–24]; others were based on observations of potential exposure to hazardous agents in walkthrough surveys of a representative sample of US worksites[25]; a more recent Finnish JEM used a database of exposure measurements to aid expert assessments[26]; and a Swedish JEM of magnetic field exposures was created using measurement data.[27]

Table 2 lists studies[23 29–45] which have attempted to examine the validity of generic JEMs.[20–23 25 26 28] Only one of these used quantitative exposure measurements as the basis of evaluation. Tielemans and colleagues[43] compared the JEM of Hoar and colleagues [20] to urinary measurements of toluene, xylene,

Table 1 Validity and reliability of self reported occupational histories

Authors, year	Study population	Occupational history measure	Comparison measure	Method of assessing validity or reliability	Results
Samet *et al*, 1978[4]	70 shipyard insulation workers	Years of self reported employment as pipefitter in interviewer administered questionnaire in 1976	Years of self reported employment, by same method, 1 year later	- R^2 for regression of two measures of years of employment	- $R^2 = 0.95$
Baumgarten *et al*, 1983[5]	274 cancer cases and 23 controls from Montreal, Canada	Self reported employer, as indicated on a self administered questionnaire with an in-person interview follow up (1979–81)	Quebec Pension Plan records from 1966–1978 linked by social insurance number	- % agreement; agreement if start or end date within ± 1 year	- 82% of person-years in agreement - 87.9% had at least 6 of 13 years in agreement - 64.3% of work histories had at least 12 of 13 years in agreement
Koskela *et al*, 1984[6]	3000 living current and former employees of foundries, metal product manufacturers, and electrical devices manufacturers hired between 1950 and 1976 in Finland	Self reported occupational histories as reported on a questionnaire	Occupational histories based on employer personnel records	- % agreement	- 85–86% of person-years in agreement
Rosenberg *et al*, 1987[7]	288 workers employed at a US capacitor manufacturing plant between 1947 and 1976	Self reported work history, as reported in an interviewer assisted questionnaire in 1976	Company personnel records with jobs classified into 40 categories	- % of person months in which job classification agreed, computed for each individual	- 76.6% mean agreement for men (median >85%, range 20–99%) - 74.3% mean agreement for women (median >80%, range 15–100%)
Stewart *et al*, 1987[8]	229 controls from a cohort employed between 1940 and 1979 in 2 US shipyards	Self reported work at the shipyard, year of hire and termination, and job held longest, as reported in a telephone interview in 1983	Company personnel records of all jobs held for at least 1 month	- % agreement and kappa - sensitivity and specificity of reporting of specific jobs	- 95% agreement on shipyard employment - 81% on start year within ± 1 year, kappa = 0.78; 73% on termination year within ± 1 year, kappa = 0.71; 67% on job held longest, kappa = 0.65 - specificities for jobs from 0.95 to 1, sensitivities from 0 to 1, median = 0.71
Eskenazi and Pearson, 1988[9]	57 women working during pregnancy, recruited in a US prenatal clinic	Self reported primary occupation and industry reported in a self administered questionnaire	Self reported primary occupation and industry reported in a clinical interview an average of 3.3 weeks apart	- % agreement	- 86% perfect matches on primary occupation, 95% at least close matches - 84% perfect matches on primary industry, 93% at least close matches
Bond *et al*, 1988[10]	143 males employed for at least 1 year at a US chemical plant on or after 1940	Self reported work history at plant, as reported in a telephone interview in 1984	Company records of 400 department titles collapsed into 50 similar work areas	- % agreement	- 76.2% agreement on usual work area assignments - 50.6% agreement on number of work area assignments
Bourbonnais *et al*, 1988[11]	100 male employees with at least 5 years' experience in a Canadian shipyard	Self reported work history in shipyard, as reported in an in-person interview	Company records of job titles and starting dates for each job held ≥6 months	- % agreement	- 89% agreement on job held longest - 62% agreement on starting date within ± 1 year
Rona and Mosbech, 1989[12]	370 cancer cases from 8 European countries	Self reported work history, as reported in interviewer or self administered questionnaires	Self reported work history by same method, at least one month apart	- % agreement and kappa	- 84% agreement on most recent job, kappa = 0.82 - 78% on previous job, kappa = 0.76 - 74% on jobs held >10 years, kappa = 0.74
Brisson *et al*, 1991[13]	154 females employed for at least 5 years in 5 garment factories in Montreal, Canada	Self reported employer from 1955 to 1983, as reported in an in-person interview	Public pension and union records	- % agreement; agreement if start or end date within ± 1 year	- 81% of person-years in agreement 89% for recent employment vs. 74% for employment ≥12 years earlier
Wärneryd *et al*, 1991[14]	25,586 randomly selected members of the Swedish population aged 25 to 74	Self reported jobs held for at least 2 years throughout working life, as reported in an interview in 1977 or 1979–81	Self reported job during week of census in 1960, 1970, 1975, and 1980, coded to the Nordic Occupational Classification	- % agreement with occupational codes at 1-, 2-, and 3-digit level for each census year	- 81–88% agreement at 1-digit level - 66–78% at 2-digit level - 52–72% at 3-digit level
Booth-Jones *et al*, 1998[15]	49 US carpenters from a study examining musculoskeletal symptoms	Self reported history of any of 14 specific job duties	Self reported history of any of 14 specific job duties, 1 to 3 weeks later	- % agreement and kappa	- 87.9 % agreement - kappa = 0.73
Brower and Attfield, 1998[16]	480 coal miners enrolled in US National Study of Coal Workers' Pneumoconiosis	Self reported work history, as reported in interview in 1969–1971	Self reported work history, as reported in interview in 1977–1981	- intraclass correlation coefficient for job tenure	- ICC=0.95 - mean tenure 20.6 years based on 1st interview, 20.8 years based on 2nd

and chromium, and found only slight agreement and low specificities and sensitivities. Several studies examined agreement between two generic JEMs. Most found kappas to be slight to fair.[23 30 33 34 41] Other investigators have compared JEMs to self reports[23 30 33 34 36 45] or expert assessments.[31 32 35 37 38 40 42 45] Although neither self reports nor expert assessments can be considered gold standards, sensitivity and specificity of the JEMs against these assessment techniques were the usual comparison measures. Sensitivities were most often below 0.5, with specificities usually higher, above 0.85. Kappas for agreement tended to be low, similar to the JEM to JEM comparisons. Some studies compared odds ratios derived from generic JEMs and study specific expert exposure evaluations.[31 40] Although both methods produced increased odds ratios where expected, only the study specific expert assessments produced clear exposure–response trends. In McNamee's evaluation, a study specific JEM also performed better than a generic JEM.[40] Study specific "internal" JEMs are, in most instances, essentially the same as expert assessments; these are discussed later in the review.

Most authors investigating the properties of generic JEMs concluded that they were not sensitive, and in only slight to fair agreement with techniques in which they had more confidence. The often low sensitivities of generic JEMs are understandable given the number of cells for which exposures need to be evaluated, and the often unpredictable circumstances in which exposures may occur in industry. A major factor which contributes to the poor performance of generic JEMs is their inability to account for variability in exposures within jobs or, in most cases, across time.[19 31 35 36 41 45] In addition, generic JEMs may not be useful if the jobs or exposures under investigation are not included in the matrix, or are grouped in such a way as to obscure their impact. These limitations have tempered the early enthusiasm for generic JEMs and promoted study specific exposure assessment methods.

SELF REPORTED EXPOSURES

Questionnaires used in case–control studies commonly ask about more than a subject's occupational history, querying use of specific agents, trade name products, or classes of compounds. Over the past two decades, there have been numerous reports[4 9 10 42 43 46–72] examining the validity and reliability of this method of exposure assessment (table 3).

Many have compared self reported exposures to industrial hygiene measurements of exposure to one or a few agents. Most of these have found significant associations between the two measures, though the proportions of variance in exposure explained (R^2) by the self reports have varied widely, from as low as 0.03 to as high as 0.71, with a median of about 0.2.[47 49 51 53 58 67–69] Some of the problem is likely to lie with the gold standard. Self reported exposures are often elicited to represent "usual" exposures, whereas exposure measurements quantify exposure over individual shifts. Exposures are well known to be extremely variable over time and place,[47 49 51 53 58 68] so even a single worker may have measurements on different days that vary by orders of magnitude. This day to day variability can account for a large proportion of exposure variability, but is not meant to be explained by self reports.[47 51 68] When Kromhout and colleagues[47] restricted exposure variability to the between task variability estimated by the subjects, the median proportion of variance explained improved somewhat, from 0.14 to 0.23; though the range over all plants and substances became even wider (0 to 0.62).

Two studies summarised the validity of self reports against quantitative measurements in terms of sensitivity and specificity.[43 62] Both measures were extremely variable, ranging from 0 to 0.85 and from 0.34 to 1.0, respectively.

Other studies have compared self reported exposure estimates to estimates by experts (note that sometimes the experts used the self reported exposures or jobs as one of their data sources). In these studies, kappa for agreement was the most frequent measure of validity. Once again, a striking characteristic of the measures of agreement was their variability from study to study and within studies for different agents, with kappas varying from −0.05 to 0.94, median ~0.6.[9 56 61 63 64 70]

A few studies examined the reliability of self reported exposures estimated at different points in time. Kappas and intraclass correlation coefficients ranged from 0.36 to 0.84, median ~0.6.[50 54 56 59 65] Proportions of variance explained in continuous measures ranged from 0.16 to 0.84, median ~0.6.[4 71]

Two studies examined the characteristics of both generic JEMs and self reported exposures. Rybicki and colleagues[42] compared the two methods to an expert industrial hygiene review of exposures to copper, lead, and iron. They found that self reports had much higher sensitivities (0.65 to 0.84) than the JEM[25] (0 to 0.21), and slightly improved specificities (0.88 to 0.96 versus 0.86 to 0.93). Tielemans and others[43] used urinary measurements of chromium, toluene, and xylene as the basis for validity comparisons. Again sensitivities were higher using exposure self reports (0.41 for chromium and 0.85 for the solvents) than for the JEM[20] (0.26 and 0.6, respectively); however, specificities suffered as a result (0.68 for chromium and 0.34 for the solvents versus 0.79 and 0.63 respectively for the JEM), and therefore so did positive predictive values.

Given the variability in subjects' ability to accurately and reliably report their own exposures, it is worthwhile to consider whether there are characteristics that are consistently associated with improved reporting. Investigators have found that subjects were better able to estimate exposure to agents which they can easily sense, for example, solvents they can smell,[47 52] dusts with larger particle sizes,[68] and vibrations they can feel.[65 72] In a similar vein, they were more able to report exposures when queried in terms they recognised, for example, "oils and greases", "degreasers", or "stainless steel", rather than about specific chemical compounds, for example, "chromium" or "imidazoline".[55 62] Those involved in the purchasing or selection of chemicals were more likely to accurately recall exposures (for example, farmers or applicators using pesticides),[57 66] than labourers who were not involved in such tasks (for example, farmworkers harvesting crops).[73] Most investigators prompt recall with a list of exposure agents of interest. This method resulted in higher sensitivities than open ended questioning, without an equivalent loss in specificity.[57 62 74] Other characteristics of subjects, such as age, sex, duration of employment, socioeconomic status, education, disease symptoms, and language had little or no effect on the accuracy of reporting exposures.[42 55 59 65 68]

An important concern with exposure self reports is recall bias—that is, whether reporting is influenced by disease status. Most investigators who compared the responses of cases and controls found little or no difference in the validity or reliability of their exposure assessments.[55 57 63 70] Rodvall and colleagues[64] did find some variations in the accuracy of reporting between cases and controls; for some agents cases were better estimators, for some controls were better, but for most agents there was little substantive difference. A recent study indicated that exposures volunteered on open ended questioning were more likely to be subject to recall bias than exposures cited after probing with a list of agents.[75] There is also evidence that the potential for recall bias may be greater in studies which use subjective measures of both exposure and outcome (a design more commonly used in cross sectional studies).[58]

A difficulty that subjects face when deciding whether to report exposures is the lack of relative or objective benchmarks against which to judge their work conditions. For example, office workers whose building was sprayed with insecticides might consider themselves exposed, but might

Table 2 Validity of generic job–exposure matrices

Authors, year	Study population	Generic JEM	Comparison measure	Method of assessing validity	Results
Hinds *et al*, 1985[29]	261 male lung cancer case and 444 population-based controls in Hawaii	Hoar *et al*., 1980[20]	Occupational cancer literature	- Consideration of odds ratios and dose-response for known carcinogens	Coal tar/pitch OR = 1.9 Petroleum pitch/tar OR = 2.0 Arsenic OR = 1.2 Chromium OR = 0.9 Asbestos OR = 12.6 Nickel OR = 1.6 Beryllium OR = 1.6
Linet *et al*, 1987[30]	342 chronic lymphocytic leukemia cases and 342 hospital-based controls without cancer, in the US	Hoar *et al*., 1980[20]; Sieber *et al*., 1991[25]	Self and surrogate reported exposures to butadiaene, asbestos, trichloroethylene, styrene, tetrachloroethylene, carbon tetrachloride, and benzene	- Sensitivity and specificity in comparison to self reports - Kappa comparing the two JEMs	- Sensitivities (benzene) from 0.10 to 0.24; specificities 0.85 to 0.91 - Sensitivities (asbestos) from 0.23 to 0.47; specificities 0.87 to 0.91 - Kappas from 0.01 to 0.60, median = 0.26
Cicioni *et al*, 1991[31]	143 mesothelioma cases and 35753 other cancer controls, excluding lung cancer cases, in the US	Sieber *et al*., 1991[25]	Expert classification of asbestos exposure based on occupation and industry of subject	- Comparison of odds ratios for asbestos exposure	- JEM ORs = 2.0 and 2.4 for low and high exposure, respectively - Expert ORs = 1.6 and 6.4 for low and high exposure, respectively
Kauppinen *et al*, 1992[32]	344 primary liver cancer and 861 controls with stomach cancer or myocardial infarction, in Finland	Pannett *et al*., 1985[21]	Expert (industrial hygienists) classification of exposure based on occupation and industry of subject	- Sensitivity and specificity in comparison to estimates of high exposure according to expert classification	- Sensitivities (for ~30 agents) from 0.02 to 0.90, median = 0.41 - Specificities 0.84 to 1.0, median = 1.0
Kromhout *et al*, 1992[33]	878 males from Zutphen, The Netherlands, followed up for lung cancer	Hoar *et al*., 1980[20]; Pannett *et al*., 1985[21]	Self reported exposures	- Sensitivity and specificity in comparison to self reports - Kappa comparing the two JEMs	- Sensitivities (12 agents) from 0 to 0.98, median ~0.45 - Specificities 0.17 to 1, median ~0.97 - Kappas from −0.07 to 0.87, median = 0.08
Ahrens *et al*, 1993[34]; Orlowski *et al*, 1993[23]	391 lung cancer cases and 391 population-based controls in Germany	Ferrario *et al*., 1988[22]; Orlowski *et al*., 1993[23]	Self reported exposures to asbestos in 19 job specific questionnaires	- Kappa comparing the two JEMs, and the JEMs and self reports	- Kappas from 0.44 to 0.67 for inter-JEM comparison, from 0.15 to 0.44 for self-report-JEM comparison
Luce *et al*, 1993[35]	616 subjects from a case-control study of sinonasal cancer in France	Ferrario *et al*., 1988[22]	Self reported exposure to formaldehyde and wood dust, with duration, intensity, and probability of exposure classified by expert (industrial hygienist) review	- Kappa comparing the JEM to the expert reviewed self reports	- Kappas (formaldehyde) from 0.17 to 0.24 - Kappas (wood dust) from 0.83 to 0.84
Roeleveld *et al*, 1993[36]	parents of 306 mentally retarded children and 322 children with other congenital handicaps (with known causes) from the Netherlands	Hoar *et al*., 1980[20]; Pannett *et al*., 1985[21]	Self reported exposures to 42 agents	- Sensitivity and specificity in comparison to self reports	- Sensitivities from 0.18 to 0.32 - Specificities 0.86 to 0.94
Stengel *et al*, 1993[37]	765 bladder cancer cases and 765 hospital-based controls; 298 cases with glomeronephritis and 298 hospital-based controls in France	Ferrario *et al*., 1988[22]	Self reported exposure to organic solvents, with review by experts	- Sensitivity and specificity in comparison to expert reviewed self report - Kappa comparing the JEM to the expert reviewed self reports	- Sensitivities from 0.23 to 0.63, median = 0.42 - Specificities from 0.87 to 0.98, median = 0.94 - Kappas from 0.29 to 0.45, median = 0.36
Stucker *et al*, 1993[38]	765 bladder cancer cases and 765 hospital-based controls, in France	Ferrario *et al*., 1988[22]	Expert classification of polycyclic aromatic hydrocarbon exposure based on occupation and industry of subject	- Sensitivity and specificity in comparison to expert classification	- Sensitivities from 0.13 to 0.96 and specificities from 0.57 to 0.99 depending on dichotomisation, specificity increased as sensitivity decreased
Le Moual *et al*, 1995[39]	10046 adults living in one of 7 French cities in 1975	Pannett *et al*., 1985[21]; Ferrario *et al*., 1988[22]	- Comparison of the two JEMs to each other and to a third "French" JEM developed for this study, considering broad exposure to "dusts, gases, and chemical fumes", and 28 more specific exposures	- Comparison of odds ratios and trend in exposure-lung function (FEV1) response of the three methods	- All three JEMs showed similar statistically significant decreasing trends in FEV1 with exposure to "dusts, gases, and chemical fumes" - Results for the 28 specific hazards were much more variable
McNamee, 1996[40]	102 chronic pancreatitis cases and 204 population-based controls from the UK	Cherry *et al*., 1992[28]	- Self reported exposures to hydrocarbons using job specific questionnaires, with expert review by hygienists and occupational physicians - Internal JEM using mean exposure scores from above method for each job	- Comparison of odds ratios between three methods	- Expert-reviewed self-reports ORs 1.9 and 3.7 for medium and high exposures respectively - Internal JEM ORs 1.7 and 2.2 - Generic JEM ORs 2.2 and 1.1

Table 2 Continued Validity of generic job–exposure matrices

Authors, year	Study population	Generic JEM	Comparison measure	Method of assessing validity	Results
Hawkes and Wilkins, 1997[41]	214 agents common to both JEMs	Hoar *et al.*, 1980[20]; Sieber *et al.*, 1991[25]	- Direct comparison of 2 JEMs, after conversion of all occupation codes to NIOSH-NOHS system, for 54 job groups in metal, paper and wood, and chemical industries	- Kappa comparing 2 JEMs	- Kappas from 0.02 to 0.27 in metal industry occupations - Kappas from –0.07 to 0.24 in paper and wood industry - Kappas from –0.12 to 0.14 in chemical industry
Rybicki *et al*, 1997[42]	188 subjects in a US case-control study of neurologic disease, all with some occupational history in manufacturing	Sieber *et al.*, 1991[25]	- Self reported exposures to copper, lead and iron, with expert review by an industrial hygienist	- Sensitivity and specificity in comparison to expert reviewed self report	- Mean sensitivities from 0 to 0.21 - Mean specificities from 0.86 to 0.93
Tielemans *et al*, 1999[43]	subjects of 2 case-control studies of male infertility in the Netherlands	Hoar *et al.*, 1980[20]	- Urine samples analysed for metabolites of toluene and xylene (n=267) and for chromium (n=156)	- Sensitivity and specificity in comparison to urine samples - Kappa in comparison to urine samples	- Sensitivities = 0.60 for toluene/xylene, 0.26 for chromium - Specificities = 0.63 for toluene/xylene, 0.79 for chromium - Kappas = 0.13 for toluene/xylene, 0.04 for chromium
Louik *et al*, 2000[44]	12188 mothers and 12017 fathers of children in a US case-control study of birth defects	Sieber *et al.*, 1991[25]	- Two experts (occupational hygienist and physician) assessed exposures to dichlorodifluoromethane, propylene glycol, and amorphous fused silica to ~200 industry/occupation combinations	- % agreement	- 20%, 3%, and 2% agreement between the JEM and at least one expert for dichlorodifluoromethane, propylene glycol, and amorphous fused silica, respectively
Benke *et al*, 2001[45]	838 subjects of a case-control study of glioma in Australia	Kauppinen *et al.*, 1998[26]	- Self reported exposures to 5 substances - Expert panel of 3 industrial hygienists, estimates of exposure to 5 substances	- Kappa	- Kappas from 0 to 0.62 (median = 0.07) in comparison to self-reports - Kappas from 0.07 to 0.46 (median = 0.28) in comparison to expert assessments

not give the same answer if asked to compare their exposure to that of pesticide applicators. In the study of Ising and colleagues,[67] subjects were able to categorise their noise exposure intensity very well; they were provided with examples of well known machines against which to gauge each noise category. In the studies of Kromhout and colleagues,[47] Hertzman and colleagues,[49] and Teschke and colleagues,[51] workers who rated only their own exposures tended to do so less well than workers or supervisors who ranked exposures in all jobs, illustrating that even relative comparisons help subjects put their exposures in context.

The variable quality of self reported exposure information indicates that although subjects can reliably and accurately report exposures in certain circumstances, it is also possible for subjects to provide exposure data of such low quality that true exposure–effect relations will be obscured or even reversed in direction.[76] It is incumbent on study designers to consider features which improve subjects' reporting accuracy, including prompted questions about agents they can sense, using familiar terms common in worksite discourse, and presenting guideposts which will help them to place their exposure in relation to that of others.

EXPERT ASSESSMENT OF EXPOSURES

There has been an increasing trend to use experts, such as occupational hygienists, chemists, engineers, and other professionals, to infer exposures from job histories or make exposure estimates based on review of subject reported information. Experts are expected to have a better vantage point than subjects: by training, they understand the mechanisms of occupational exposures and know where to find data about them; within the context of a study, they know the types of exposures considered relevant; and based on study data, they have an overview of the range of jobs whose exposures need to be estimated. But experts also bear some handicaps: they may not be familiar with many of the jobs and industries which appear in subjects' occupational histories; and unless they have detailed reports from subjects, they are certainly unlikely to be aware of conditions present in specific worksites of subjects. How these trade offs balance can be examined through studies of the validity and reliability of experts' exposure assessments (table 4).[43 47 51 77-93]

Because expert assessments have generally been considered the best possible exposure estimation method short of exposure measurements,[2] studies examining their validity have exclusively used comparisons to measurements. As in similar tests of subject's self reports, these validity studies have examined experts' estimates of exposure intensity for only a few agents. Many have reported results in such a way that the proportions of variance explained can be compared. As noted for self reported exposures, variability in the validity results is the most striking feature, with proportions of variance explained ranging from 0 to 0.86, with a median of about 0.3.[47 51 81 84 86 92] These results are slightly better overall than those of self reports. As with self reported exposures, it is likely that a portion of the unexplained variability is caused by day to day variation in measured exposure. The report of Kromhout and colleagues[47] excluded this variation, and found a considerable improvement in the median proportion of variance explained, from 0.25 to 0.45; though the range over all plants, substances, and hygienist estimators once again increased somewhat (0 to 0.63).

Two studies examined validity in terms of sensitivity and specificity.[43 88] The sensitivities were extremely variable, ranging from 0.21 to 0.79, median 0.35, but specificities were higher and more stable, from 0.91 to 0.98. In studies where exposure prevalence is low, as in most case–control studies, it is vital to maximise specificity to minimise attenuation of

Table 3 Validity and reliability of self reported exposures

Authors, year	Study population	Self report measure	Comparison measure	Method of assessing validity or reliability	Results
Samet *et al*, 1978[4]	70 shipyard insulation workers	Years of self reported exposure to asbestos, radiation, and fibreglass in interviewer administered questionnaire in 1976	Years of self reported exposures, by same method, 1 year later	- R^2 for regression of two measures of years of exposure	- R^2 = 0.84 for asbestos, = 0.82 for radiation, = 0.73 for fiberglass
doPico, 1982[46]	209 US grain workers	Self reported estimates of dust exposure, as heavy versus not heavy, and as less than average, average, or more than average	Measured levels of total dust concentration, one sample per person	- Comparisons of mean dust concentrations for self reported exposure categories	- Mean dust concentration for "heavy dust" = 10.1 mg/m^3 vs. 1.58 mg/m^3 for not heavy, p<0.001 - Mean dust concentration for "more than average" = 13.9 mg/m^3 vs. 4.21 mg/m^3 for "average", and 1.84 mg/m^3 for "less than average", p<0.01
Kromhout *et al*, 1987[47]	Employees of a paint factory (n=29), a food processing facility (n=58), a nonwoven materials factory (n=164), and 2 coach works (n=144) in the Netherlands	Self reported rankings of exposure to either dust or solvents, using a 4-point scale, for the task they were performing at the time of exposure measurement	58 solvent measurements in the paint factory; and 421 dust measurements in the other plants	- Proportion of variance in exposure explained (adjusted R^2) by the employees' rankings	- R^2 ranged from 0.03 to 0.23 for dusts, median = 0.14 - R^2 = 0.56 for solvents
Järvolm and Sandén, 1987[48]	951 males employed in Swedish shipyards	Self reported intensity of exposure to asbestos, in four ordinal categories (very low =1, low =2, heavy =3, very heavy =4)	Ratings by employee experts with long experience in the industry (4 production workers, 4 safety engineers, and 7 safety stewards), of each job's asbestos exposure intensity, in four ordinal categories	- Exposure-reponse relation between pleural plaques and estimated asbestos exposure intensity	- Pleural plaque prevalence with self-reported asbestos intensity: 1=33%, 2=34%, 3=41%, 4=48%; With expert estimated asbestos intensity: 1=40%, 2=37%, 3=35%, 4=42%;
Bond *et al*, 1988[10]	143 males employed for at least 1 year at a US chemical plant on or after 1940	Self reported exposures, based on open ended questioning, as reported in a telephone interview in 1984	171 agents catalogued by company hygienists	- % agreement	- Highest agreement for chlorine (11%) and asbestos (9%), lowest for sulfur dioxide and heat (<1%)
Eskenazi and Pearson, 1988[9]	57 women working during pregnancy, recruited in a US prenatal clinic	Self reported exposures to heat, cold, noise, poor ventilation, radiation, video display terminals, biological agents, fumes, gases, and dusts, as reported in a clinical interview	Industrial hygiene review of self reported exposures to heat, cold, noise, poor ventilation, radiation, video display terminals, biological agents, fumes, gases, and dusts, as reported in a self administered questionnaire	- Sensitivity and positive predictive value, using the industrial hygiene review as the gold standard; kappa	- Sensitivities ranged from 0.5 to 0.9, median = 0.67 - Positive predictive values ranged from 0.56 to 1.0, median = 0.75 - Kappas ranged from 0.42 to 0.94, median = 0.63
Hertzman *et al*, 1988[49]	172 Canadian sawmill workers	Self reported hours of exposure to chlorophenate fungicides, per year	Concentration of total chlorophenate in urine samples, measured in 150 workers in the summer, and 154 workers in the fall	- Pearson correlation coefficient	- Pearson r = 0.67 for summer samples; = 0.58 for the fall samples
Pron *et al*, 1988[50]	117 controls randomly selected from a municipal population in Canada	Self reported continuous exposure to second hand tobacco smoke, and number of worksites where exposed	Reinterview 6 months later	- Kappa	- Kappa = 0.46 for ever vs. never exposed to second-hand smoke at work; - Weighted kappa = 0.37 for number of worksites exposed
Teschke *et al*, 1989[51]	225 Canadian sawmill workers	Self reported hours of exposure to chlorophenate fungicides, per year	Concentration of total chlorophenate in one urine sample from each worker	- R^2, proportion of variance in urinary chlorophenate concentrations explained by self-reported hours of exposure	- R^2 = 0.15; with self-reported skin exposure also included in the model, R^2 = 0.17
Ahlborg, 1990[52]	From a cohort of Swedish drycleaners, 48 women with a pregnancy ending in a miscarriage requiring hospitalisation, perinatal death, low birthweight or malformed child, and 110 control women	Self reported presence of drycleaning operations in worksite and exposure to tetrachloroethylene	Employer reports on type of production and drycleaning agents used, including tetrachloroethylene, in various employment periods	- Sensitivity and specificity in comparison to employer reports	- Sensitivity and specificity of drycleaning operation for cases = 0.97 and 0.75; for controls = 0.96 and 0.69 respectively - Sensitivity and specificity of tetrachloroethylene exposure for cases = 1.0; for controls = 0.93 and 0.94 respectively
Bachmann and Myers, 1991[53]	224 South African grain mill employees	Self reported classification of their work as not dusty, slightly dusty or dustry, and very dusty	Investigators' classification of work areas into 4 dustiness scores: 0, 1, 3, 18; scores verified by exposure measurements	- R^2, proportion of variance in investigators scores explained by self-reported classifications - % agreement of dustiness classifications	- R^2 = 0.13 - % agreement = 54%

Table 3 Continued Validity and reliability of self reported exposures

Authors, year	Study population	Self report measure	Comparison measure	Method of assessing validity or reliability	Results
Holmes and Garshick, 1991[54]	116 US male veterans	Self reported work in a dusty job and exposure to asbestos, reported in a mailed questionnaire	Self reported work in a dusty job and exposure to asbestos, reported in a clinic based interview, an average of 7 months later	- % agreement	- 60% agreement on asbestos exposure - 71% agreement on dust exposure
Joffe, 1992[55]	420 employees of 5 factories in the printing or plastics industries in a study of fertility and miscarriage in the UK	Self reported exposures in the most recent job to imidazoline, carbon black, diazo dyes, resins, varnishes, oils and greases, solvents/degreasers, coloured inks	Department head and other company data on the use of these chemicals in each department, including changes over time	- Sensitivity and specificity using the company data as the the gold standard	- Sensitivities from 0.70 to 0.85 for oils and greases, solvents/degreasers, and colored inks, from 0.24 to 0.45 for remainder - Specificities at least 0.95 for imidazoline, carbon black, and diazo dyes, from 0.48 to 0.78 for remainder
Walter *et al*, 1992[56]	103 subjects of a case-control study of melanoma in Canada	Self reported exposure to fluorescent lights at work, as reported in an in-person interview	- Self reported exposure to fluorescent light, as reported on a mailed questionnaire, several weeks later - Employer reported exposures to fluorescent lights at work, for 25 jobs	- % agreement and kappa	- % agreement = 80%, kappa = 0.57 - % agreement overall = 68%; kappa for jobs where both employers and subjects able to classify fluorescent light exposure = 0.79 (n=19)
Blair and Zahm, 1993[57]	69 cases and 41 controls in a case-control study of soft tissue sarcoma and lymphoma in US farmers	Self reported use of herbicides and insecticides	Suppliers' reports of farmers' herbicide and insecticide use	- % agreement	- 59% agreement on use of both herbicides and insecticides
Fonn *et al*, 1993[58]	305 South African grain mill employees	Self reported dustiness in 4 ordinal categories (very high, high, medium, low), as reported in an interviewer administered questionnaire	Measured exposure to total dust, used to categorise work areas into the same 4 dustiness categories	- Kendall's tau - Contingency coefficient	- tau = 0.45 - cont. coef = 0.48
van der Gulden *et al*, 1993[59]	209 subjects of a case-control study of prostate cancer in the Netherlands	Self reported exposure to pesticides, fertilisers, iron and steel, non-ferrous metals, welding fumes, solvents, paints, and lubricating oils, as reported in a mailed questionnaire	Self reported exposure to the same substances, as reported in an telephone interview 3 to 5 weeks later	- % agreement and kappa	- % agreement from 75 to 88% for iron and steel, welding fumes, and fertilizers, and from 64 to 70% for remainder; median = 73% - Kappas from 0.55 to 0.70 for iron and steel, welding fumes, and fertilizers, and from 0.36 to 0.48 for remainder; median = 0.52
Halpin *et al*, 1994[60]	90 current sawmill workers, 14 former sawmill workers, and 58 light engineering factory employees, in the UK	Self reported dustiness, on an ordinal scale from 0 to 3	Personal dust samples of random sample of workers within certain mill/factory areas	- Comparisons of mean dust concentrations and median dustiness ratings for each area	- Mean dust concentrations for "0" dustiness areas = 0.24 mg/m^3; for "1" = 0.71 to 1.13 mg/m^3; for "2" = 1.32 to 6.25 mg/m^3
Savitz *et al*, 1994[61]	161 mothers who worked in the US textile industry, selected from subjects in a case-control study of miscarriage, preterm delivery and low birth weight	Self reported exposure (yes/no) to vibration, solvents, heat, and noise	Expert review of subjects' job histories, type of machinery, work methods, and environmental conditions, to assign exposure to each job as unlikely, possible, or likely	- Kappa	- Kappas for vibration = 0.08 for females, 0.23 for males; for solvents = 0.26 and 0.02; for narrow heat classification = 0.15 and 0.17; for narrow noise classification = 0.20 and 0.24.
Teschke *et al*, 1994[62]	78 sawfilers in 8 Canadian sawmills	Self reported exposures to 8 individual metal components of saws and 5 composite materials (coolant, babbitt, tungsten carbide, stellite, grinding dust), as reported in an interviewer administered questionnaire using either partial or detailed prompting	- Measured air concentrations of 8 specific metals above detection limits - Observations by hygienists of sawfilers' proximity to machines where the 5 composite materials	- Sensitivity and specificity using measured concentrations and observations of work as the gold standards	- Sensitivities for metals ranged from 0 to 0.58 (median = 0.22), specificities from 0.69 to 1.0 (median = 0.88) - Sensitivities for composite materials ranged from 0.80 to 1.0 (median = 0.83), specificities from 0.62 to 0.86 for all materials except grinding dust (0.18), (median = 0.78)
Fritschi *et al*, 1996[63]	1657 cases and 253 population based controls from a study among men with cancer at any of 19 tumor sites in Montreal, Canada	Self reported exposure to 11 groups of substances (fur or leather, wood products, glues, paints, pharmaceuticals, pesticides/fertilisers, insulation, oils/greases, fuels, solvents, plastics/rubber), as reported in a self administered questionnaire	Review by a team of industrial hygienists and chemists of subjects' job and exposure histories to assess exposure to 42 substances falling within these 11 groups of substances	- Sensitivity and specificity, using the industrial hygiene review as the gold standard; kappa	- Sensitivities ranged from 0.39 to 0.91 (median = 0.61) - Specificities ranged from 0.83 to 0.97 (median = 0.90) - Kappas ranged from 0.33 to 0.64 (median = 0.51)

Table 3 Continued Validity and reliability of self reported exposures

Authors, year	Study population	Self report measure	Comparison measure	Method of assessing validity or reliability	Results
Rodvall *et al*, 1996[64]	151 glioma cases and 343 population based controls in Sweden	Self reported exposures to pesticides; oil or coal products; paints, pigments or glues; plastic materials; radiation; and solvents, degreasers, or cleaning agents	Industrial hygienist's review of self reported exposures to classify probability and level of exposure in the corresponding job	- Kappa	- Kappas for pesticides = 0.88 and 0.46 for cases and controls respectively; for oil and coal products = 0.72 and 0.74; for paints, pigments and glues = 0.59 and 0.51; for plastic materials = 0.61 and 0.80; for radiation = 0.89 and 0.78; for solvents, degreasers, or cleaning agents = 0.69 and 0.58
Wiktorin *et al*, 1996[65]	343 Swedish workers in a study of musculoskeletal disorders	Self reported duration of exposure to vibrating floors, and vibrating hand tools, as reported in a self administered questionnaire	Self reported duration of exposure, as reported 2 weeks later	- Intraclass correlation coefficient	- ICC for vibrating floors = 0.70 - ICC for vibrating hand tools = 0.84
Calvert *et al*, 1997[66]	32 employees of 15 US structural fumigation companies	Self reported information on fumigation industry employment, % of jobs using methyl bromide and sulfuryl fluoride, and pounds of sulfuryl fluoride used in previous two weeks	Company personnel records and daily work records, fumigant use logs	- Pearson correlation coefficient	- Pearson r = 0.97 for years employed in structural fumigation; = 0.66 to 0.88 for percent of jobs using specific fumigants; = 0.68 for pounds of sulfuryl fluoride used on job
Ising *et al*, 1997[67]	80 employed men from a German population based case-control study of myocardial infarction	Self reported categorisation of noise in current workplace: 1=refrigerator; 2=typewriter; 3=electric lawnmower; 4=electric drill; 5=pneumatic drill	Measured one minute average noise level at each worksite	- Median noise levels for each self reported noise category - Spearman rank correlation coefficient between self reported ordinal category and measured noise level	- Median noise level for category "1" = 53 dBA; for category "2" = 53 dBA; for category "3" = 75 dBA; for category "4" = 88 dBA; for category "5" = 100dBA - Spearman r = 0.84
Nieuwenhuijsen *et al*, 1997[68]	104 workers from 10 US farms	Self reported dust exposure during a single sampling period of about 2 to 3 hours, on ordinal scale from "0" = "no dust exposure at all" to "10" = "dust exposure that severely restricted your view"	Measured inhalable and respirable dust concentrations	- Spearman rank correlation coefficient	-Spearman r for inhalable dust levels and self-reports = 0.67 -Spearman r for respirable dust levels and self-reports = 0.36
Rybicki *et al*, 1997[42]	188 subjects in a case-control study of neurologic disease, all with some occupational history in manufacturing in the US	Self reported exposures to copper, lead and iron, in 544 jobs reported in an interviewer-administered questionnaire	Expert review by an industrial hygienist of self reported exposures	- Sensitivity and specificity in comparison to expert review	- Mean sensitivities for iron = 0.65, for lead = 0.73, for copper = 0.84 - Mean specificities for iron = 0.88, for lead = 0.94, for copper = 0.96
Willemsen *et al*, 1997[69]	107 non-smokers from 36 offices in the Netherlands	Self reports of how often office mates smoke; how often bothered by the stench of tobacco smoke (both never, sometimes, regularly); and how much tobacco smoke there is on average in the office (7 categories from "no smoke" to "very much smoke")	Nicotine concentrations measured in each office for one full shift	- Pearson correlation coefficients between average self reports for each office and the measured nicotine levels in the office	- Pearson r = 0.41 for frequency of office mates smoking; = 0.33 for frequency of being bothered by stench; = 0.65 for average amount of tobacco smoke on average
Nordstrom *et al*, 1998[70]	28 carpal tunnel syndrome cases and 33 controls in a US case-control study	Self reported work in a noisy area where plugs or muffs used, and in cold temperatures in the winter	Observations of the subject's work for 1 hour by an ergonomist	- Kappa and Spearman rank correlation	- Work in a noisy area: for cases kappa = 0.44, Spearman r = 0.53; for controls kappa = 0.31, Spearman r = 0.40 - Work in a cold environment: for cases kappa = 0.31, Spearman r = 0.55; for controls kappa = 0.68, Spearman r = 0.74
Tielemans *et al*, 1999[43]	Subjects of 2 case-control studies of male infertility in the Netherlands	Self reported exposure to solvents (as indicated by indicating contact with any of the following: industrial cleaning products, degreasers, paint, glue, printing inks, paint removers, other solvents); or to chromium (as indicated by contact with welding fumes)	- Urine samples analysed for metabolites of toluene and xylene (n=267) and for chromium (n=156)	- Sensitivity and specificity in comparison to urine samples - Kappa in comparison to urine samples	- Sensitivities = 0.85 for toluene/xylene, 0.41 for chromium - Specificities = 0.34 for toluene/xylene, 0.68 for chromium - Kappas = 0.08 for toluene/xylene, 0.08 for chromium

Table 3 Continued Validity and reliability of self reported exposures

Authors, year	Study population	Self report measure	Comparison measure	Method of assessing validity or reliability	Results
Neale *et al*, 2000[71]	243 subjects of a colorectal cancer screening program among pattern and model makers in the US	Self reported exposure to 13 substances (cutting oils, epoxies, fiberglass, wood dusts, fibreglass, plaster dust, polyesters, solvents, welding fumes) for each job in their work history, as reported in a self administered questionnaire in 1985	Self reported exposure to same 13 substances as reported in a self administered questionnaire in 1988	- Pearson r for % of time exposed for job held in 1982 and 1985	- r for jobs held in 1982 ranged from 0.54 to 0.74, median = 0.60 - r for jobs held in 1982 ranged from 0.40 to 0.72, median = 0.57
Palmer *et al*, 2000[72]	179 workers in various jobs involving exposure to hand transmitted or whole body vibration in the UK	Self reported exposure to hand transmitted or whole body vibration, in 1-hour period, including name of vibrating equipment and duration of exposure	Observations of workers during one-hour period	- Sensitivity and specificity - Ratio of reported to observed durations of exposure	- Sensitivity = 0.96 for hand-transmitted vibration, 0.97 for whole body vibration - Specificity = 0.98 and 0.91, respectively - Median ratio of self-reported to observed exposure time = 2.5, range 1.2 to 6.3 for hand-transmitted vibration; Median = 1.1, range 1.0 to 1.2 for whole body vibration

effect estimates as a result of exposure misclassification[94]; therefore the high specificities are an encouraging result.

Studies examining agreement between experts' ratings have mainly compared exposure assessments of different experts, with kappas or intraclass correlation coefficients ranging from 0 to 1.0 with a median of about 0.6.[51 78 79 82 83 87 90 91] Two studies have examined repeatability of ratings by the same experts, with similar results (kappas from 0.26 to 0.77, median ~0.6).[89 91]

Three of the studies examining the validity of experts' assessments against exposure measurements similarly examined the validity of self reports, so provide a basis for comparison. Kromhout and colleagues[47] found slightly higher proportions of variance in solvent and dust measurements explained by hygienists' estimates, as did Teschke and colleagues[51] in a study of chlorophenate fungicide exposures. In the study by Tielemans and colleagues[43] of solvent and chromium exposure, sensitivities were higher for self reported exposures, but specificities and positive predictive values were higher for the experts' estimates.

Although expert assessments are often thought of as a single method, many different assessment structures and tools can be used by experts to assign exposures in case–control studies. One common structure involves using a subject's job description as the basis for assigning exposures, another is to have experts estimate exposures of jobs and/or industries, without subject supplied information. The data used to create exposure estimates are often published literature and judgement, as used in many of the first generic JEMs.[20–23 28] "Internal" JEMs differ from generic JEMs in that the exposures and jobs selected for assessment are study specific, and the assessors can be chosen for their particular expertise in these areas. Experts' estimates can be made subject specific, usually by providing experts with subjects' self reported exposure and job duty information. In a method developed by Gérin and colleagues[95] and elaborated for more jobs by Stewart and colleagues,[96] experts are guided by subjects' answers to detailed questions about tasks, materials, equipment, and control measures in occupation or industry specific modules. Finally, some expert assessment methods augment the above tools with whatever measurement data might be available, for example, measurements of similar jobs or industries from national exposure databases.[97]

Several studies have compared the validity and reliability of different levels of expert assessment. Stewart and colleagues[93] evaluated experts' assessments of formaldehyde exposure in manufacturing plants, starting with information on job title, then adding department, industry, date, and plant reports in stages. There was little difference in the quality of the assessments with the amount of data provided. Similarly, de Cock and colleagues[86] found little effect on experts' estimates of captan exposure among fruit growers between phases of assessment which started with a video about factors affecting exposure, then added information on pesticide application tasks, and finally information on pesticides. Segnan and colleagues[87] compared assessments by experts based on occupational histories to assessments based on industry specific modules (using as the gold standard, the same experts' estimates with additional product information and exposure measurements). They found little change in sensitivity using the industry specific modules, but median specificities increased from 0.52 to 0.77. Tielemans and colleagues[43] compared two very similar methods using urinary measurements of chromium, toluene, and xylene as the gold standard. Compared to using occupational histories alone, sensitivities increased slightly when industry specific questionnaires were used, specificities were nearly unchanged, and kappas increased.

Other investigators have examined the effect of offering industrial hygiene measurement data to the experts conducting the assessments. Hawkins and Evans[80] examined the ability of occupational hygienists to estimate toluene exposures of workers in the chemical industry, and found that initial estimates without data overestimated exposures by more than twofold, but that offering some limited measurement data allowed the hygienists to "calibrate" their estimates so they were less biased. Post and colleagues[81] examined hygienists' estimates of exposures to styrene and methylene chloride among polyester factory workers. Although the relative ranking of jobs did not seem to improve as the hygienists were provided with additional measurement data, the added data did improve their classification of jobs into quantitative exposure categories.

Other factors which might influence the validity and reliability of experts' assessments include the agents being assessed, and the expertise of the assessors. Segnan and colleagues[87] found higher intraclass correlations for insecticides, fungicides, nickel, copper, chromium, and aliphatics hydrocarbons than for specific pesticides, inorganic compounds, and halogenated organics. Sensitivities and specificities followed a similar pattern. Benke and colleagues[45] found that kappas for agreement were higher for cutting fluids,

Table 4 Validity and reliability of exposures estimated by experts

Authors, year	Study population	Expert measure	Comparison measure	Method of assessing validity or reliability	Results
Woitowitz *et al*, 1970[70]	Employees of a raw asbestos processing plant in Germany	Study personnel ranked employees' asbestos exposure on a 4-point scale, based on tours of the worksite, and consultations with department heads, trade union personnel, plant or government physician, firm safety officer, and shop committee	61 dust measurements	- Comparisons of mean dust concentrations of expert-assessed exposure categories	- Mean dust concentration for category 1 = 0.6 mg/m^3; for category 2 = 1.2 mg/m^3; for category 3 = 1.6 mg/m^3
Goldberg *et al*, 1986[78]	Subjects of a case-control study among Canadian men with cancer at any of 19 tumor sites	A rater (occupational hygienist, chemist, or engineer) assessed subjects' presence or absence of exposures to 172 to 275 substances, based on information provided by subjects in a detailed interview, their own knowledge, review of bibliographic materials, and consultations with other experts	Assessments of 1 to 2 other raters on the study team using the same methods; and assessments by an expert from certain industries external to the study	- % agreement between raters; kappa	- Average % agreement in 6 inter-rater agreement trials ranged from 95.5% to 98.5% - Average kappas ranged from 0.51 to 0.67
Kromhout *et al*, 1987[47]	Employees of a paint factory (n=29), a food processing facility (n=58), a nonwoven materials factory (n=164), and 2 coach works (n=144) in the Netherlands	Two occupational hygienists ranked employees' exposures to either dust or solvents using a 4-point scale	58 solvent measurements in the paint factory; and 421 dust measurements in the other plants	- Proportion of variance in exposure explained (adjusted R^2) by the hygienists' rankings	- R^2 ranged from 0.08 to 0.27 for dusts, median = 0.25 - R^2 were 0.37 and 0.58 for the solvents, for the two hygienists
Ciccone and Vineis, 1988[79]	88 soft tissue sarcoma cases and 157 population based controls from a rice growing region of Italy	An agricultural chemist with rice growing expertise assessed subjects' exposure to phenoxy herbicides using a 3-point scale, using information collected from the subjects on their job history, locations of farms, types of crops, characteristics of pesticide handling, and their subjective estimate of pesticide exposure	Assessments by a similarly expert agricultural chemist, using the same method	- % agreement between chemists; kappa	- 95.5% agreement - weighted kappa = 0.76
Hawkins and Evans, 1989[80]	12 chemical process workers in the US	24 occupational hygienists with expertise in chemical processing randomly selected from 2 professional organizations gave quantitative estimates of the employees' toluene exposures, first based on a qualitatitive description of the process and work environment, then after viewing limited historical measurement data	134 toluene exposure measurements	- Comparison of mean, median, and 90th percentiles of exposure estimates to measured exposures	- Mean measured exposure = 4.6 ppm; first estimate = 14 ppm; second estimate = 4.4 ppm - Median measured exposure = 0.37 ppm; first estimate = 8.9 ppm; second estimate = 3.0 ppm - 90th %ile of measured exposure = 16 ppm; first estimate = 41 ppm; second estimate = 21 ppm
Teschke *et al*, 1989[51]	225 Canadian sawmill workers	Three pairs of occupational hygienists estimated hours of exposure to chlorophenate fungicides per year, after walk through survey of sawmill	Concentration of total chlorophenate in urine samples, measured in 150 workers in the summer, and 154 workers in the fall	- Proportion of variance in urinary chlorophenate concentrations explained by hygienists' estimates of hours of exposure (R^2) - Intraclass correlation coefficients for agreement between members of each pair of hygienists	- Lumber industry hygienists: R^2 = 0.08; with skin exposure also included in the model, R^2 = 0.26; government hygienists: R^2 = 0.24; other industry hygienists: R^2 = 0.22 - Lumber industry hygienists: ICC = 0.68; government hygienists: ICC = 0.40; other industry hygienists: ICC = 0.50
Post *et al*, 1991[81]	Employees in nine jobs in a small polyester factory in the Netherlands	9 occupational hygienists each classified employees' exposures to styrene and methylene chloride into three absolute categories related to the Threshold Limit Value, using information about the process and jobs, and a visit to the plant; estimation was done two additional times after receipt of a small amount of measurement data	45 styrene and 28 methylene chloride exposure measurements	- Spearman rank correlation coefficient comparing relative ranking of jobs by each hygienist to measured ranking - % agreement between absolute classifications of jobs by each hygienist and measured classifications	- Spearman r ranged from 0.3 to 0.9 for methylene chloride, median = 0.65; and from –0.4 to 0.65 for styrene, median = 0.2 - % agreement ranged from 0.15 to 1.0 for methylene chloride, median = 50%; and from 0.1 to 1.0 for styrene, median = 55%
Dovan *et al*, 1993[82]	81 homes of cases and controls in a study of childhood cancer in the US	A study technician trained in a "wire coding" procedure classified the current configuration of the electrical transmission and distribution lines near each home into one of four ordinal categories in 1985	- Repeat of the wire coding in 1990	- % agreement in wire codes	- 90% agreement

Table 4 Continued Validity and reliability of exposures estimated by experts

Authors, year	Study population	Expert measure	Comparison measure	Method of assessing validity or reliability	Results
Macaluso *et al*, 1993[83]	29 paint department employees of a car assembly plant in the US	Occupational hygienist (university or consultant) classified exposures of 695 job-department-year combinations to 6 product use groups and 7 specific chemical groups, into six absolute concentration categories, based on historical measurement data	Assessments by 4 similar hygienists, using the same method	- Intraclass correlation coefficient for all job-department-year exposure scores and for cumulative exposure scores	- ICCs for all job-department-year exposure scores ranged from −0.05 to 0.64, median = 0.14; ICCs for cumulative exposures from 0 to 0.85, median = 0.16 - ICCs for all job-department-year exposure scores ranged from −0.07 to 0.33, median = 0.24; ICCs for cumulative exposures from 0 to 0.58, median = 0.33
Takahashi *et al*, 1994[84]	42 deceased Canadian men who had cancer at various sites, all subjects in a case-control study	A team of experts (occupational hygienists, chemists, or engineers) ranked subjects' exposures to asbestos into 3 ordinal categories based on information provided by subjects in a detailed interview, their own knowledge, review of bibliographic materials, and consultations with other experts	Measured asbestos fibre concentrations in 42 lung tissue samples taken at autopsy	- Comparisons of mean lung asbestos concentrations for expert assessed exposure categories - Proportion of variance in lung asbestos concentration explained (R^2) by experts' ratings	- Mean fibre concentration of subjects rated as having no asbestos exposure history = 0.09 f/ug dry lung; with low or moderate estimated exposure concentration = 0.14 f/ug; with high estimated exposure concentration = 8.7 f/ug - R^2 = 0.32, with age also in model, p<0.0006
Armstrong *et al*, 1996[85]	31 cases of lymphohaematopoetic cancer and 124 controls who were subjects of a nested case-control study of petroleum marketing and distribution workers in Canada	Study experts estimated exposures to total hydrocarbons and benzene, based on an algorithm which included mean measured exposure levels, and modifying information about the workplace, tasks performed, ambient environment, and products	15 measurements of total hydrocarbon exposure in the 6 job-years, and 51 measurements of benzene exposure in the 9 job-years; these data were withheld from the expert estimation process	- % difference in mean algorithm estimated exposures to withheld measurement means	- % difference for total hydrocarbons ranged from −49% to 220%, median = −35% - % difference for benzene ranged from −14% to 130%, median = 4.5%
de Cock *et al*, 1996[86]	15 fruit growing farms in the Netherlands	15 occupational hygienists, pesticide experts, and fruit growers ranked potential for pesticide exposure by dermal and inhalation routes, of 14 tasks and of 15 spraying activities; done in three phases: after viewing a video on factors affecting exposure, after viewing slides about pesticide application tasks, and after reading written information on tasks and pesticides	Measurements of airborne and dermal captan concentrations during spraying	- Spearman rank correlation coefficients comparing relative ranking of activities by each expert to measured ranking - Intraclass correlation coefficients for agreement between experts	- Spearman r ranged from −0.1 to 0.45 for inhalation exposure, median = 0.3; and from 0.03 to 0.9 for dermal exposure, median = 0.65 - ICCs ranged from 0.61 to 0.81 for inhalation exposure, median = 0.72; and from 0.53 to 0.71 for dermal exposure, median = 0.63
Segnan *et al*, 1996[87]	82 vineyard workers, 171 metal plating workers, and 158 leather goods workers in Italy	Using occupational histories, industry specific questionnaires, lists of products used, and where available, exposure measurement data (in separate stages) - 8 agronomists assessed exposure to 10 pesticides, 6 classes of chemicals, and to the broad groups "fungicides" and "insecticides" - 8 industrial hygienists assessed exposures to 20 chemicals and 5 classes of chemicals - 4 industrial hygienists assessed exposures to 20 solvents and 9 classes of chemicals	The same experts' ratings using the full set of data available to them	- Sensitivity and specificity using the ratings based on the full datatset as the gold standard - Intraclass correlation coefficients for the agreement between raters	- Sensitivities ranged from 0.13 to 0.99 (median = 0.78) using only the occupational histories, 0.05 to 1.0 (median = 0.76) using the industry-specific questionnaires, 0.87 to 1.0 (median = 0.97) using the product lists - Specificities ranged from 0.12 to 0.90 (median = 0.52) using only the occupational histories, 0.17 to 0.99 (median = 0.77) using the industry-specific questionnaires, 0.77 to 1.0 (median = 1.0) using the product lists - ICCs ranged from below 0 to 1.0 for all methods, with a median of 0.11 using only the occupational histories, 0.21 using the industry-specific questionnaires, 0.65 using the product lists, and 0.51 using monitoring data

Table 4 Continued Validity and reliability of exposures estimated by experts

Authors, year	Study population	Expert measure	Comparison measure	Method of assessing validity or reliability	Results
Benke *et al*, 1997[88]	Jobs reported in a case-control study of glioma in Australia	3 industrial hygienists and 2 occupational physicians assessed the presence or absence of exposure to 21 chemicals in 199 jobs randomly selected from subjects' histories and 49 jobs with exposure measurements	Industrial hygiene reports for the 49 jobs from a database of surveys in the study region, over the period from 1978–1989	- Sensitivity and specificity for exposure in 49 jobs with exposure measurements - Kappas for pairwise inter-rater agreement for the 199 jobs of study subjects - Kappas for intra-rater agreement based on reassessment of 50 of the 199 jobs, at least 4 months later	- Sensitivities ranged from 0.48 to 0.79, median = 0.65; specificities ranged from 0.91 to 0.98, median = 0.94 - Kappas for inter-rater agreement ranged from 0 to 0.64, median = 0.19 - Kappas for intra-rater agreement ranged from 0.46 to 0.73, median = 0.60
McGuire *et al*, 1997[89]	179 job histories from population based case-control study of amyotrophic lateral sclerosis in the US	Panel of four industrial hygienists, by consensus, rated exposure to three groups of chemical agents based on job history information	Blinded repeat assessment by same panel	- % agreement and kappas	- % agreement 90% for metals, 82% for solvents, and 97% for agricultural chemicals - Kappas were 0.77 for metals, 0.64 for solvents, and 0.75 for agricultural chemicals
Semiatycki *et al*, 1997[90]	50 subjects from a case-control study among Canadian men with cancer at any of 19 tumour sites	Consensus review by two or more experts (industrial hygienists or chemists) of subjects' histories of their jobs, work environments, raw materials, products, and self reported exposures to assess presence and probability of exposure to 294 chemical and physical agents	Two experts reassessed exposure to 94 jobs by consensus, and reassessed exposure to 92 different jobs independently from one another	- Kappas	- In consensus reassessment, weighted kappa for 4 categories of probability of exposure = 0.80; kappas for presence of exposure ranged from 0.51 to 0.94 for the 18 highest-prevalence substances, median = 0.75 - In independent reassessments, weighted kappas for 4 categories of probability of exposure ranged from were 0.73 and 0.76 for the two experts
Rybicki *et al*, 1998[91]	Job histories of 60 and 64 of 608 subjects in a case-control study of neurologic disease in the US	Expert review by two industrial hygienists of self reported exposures to copper, lead and iron, of 60 and 64 study subjects	Prior expert review by one of the same industrial hygienists of self reported exposures of all 608 study subjects	- Kappas for intra-rater and inter-rater agreement	- Kappas = 0.26 for copper, 0.56 for lead, and 0.57 for iron, intra-rater - Kappas = 0.15 for copper, 0.29 for lead, and 0.49 for iron, inter-rater
Cherrie and Schneider, 1999[92]	17 jobs in brick manufacturing in the UK, 13 jobs in rubber and pigment coating in the UK, 14 jobs in fibre reinforced plastics in Denmark, 13 jobs in an asbestos contaminated warehouse in the US, and 6 jobs in man made mineral fibre manufacturing in the US	Two industrial hygienists' estimates of exposure concentrations of respirable dust, toluene, styrene, asbestos, or man made mineral fibre, based on descriptions of the jobs, tasks, work environments, and control measures, using a structured assessment method based on emissions, processing at the source, and controls	Measurements of airborne concentrations of respirable dust, toluene, styrene, asbestos, and man made mineral fibre	- Correlation coefficient comparing log transformed exposure measurements to the hygienists' estimates - Bias, as measured by ratio of the geometric mean exposure estimate to the geometric mean measured concentration	- Pearson r ranged from 0.31 to 0.93 for all agents except styrene, and from 0 to 0.31 for styrene, for which there was little variability in measured exposure, median = 0.39 - Bias ratio ranged from 0.47 (underestimate of exposure) to 2.86 (overestimate), median = 1.47
Tielemans *et al*, 1999[43]	Subjects of 2 case-control studies of male infertility in the Netherlands	Researchers used subject reported data from generic job histories and job specific questionnaires (additional data on tasks and substances used) to assess exposures to solvents and chromium	- Urine samples analysed for metabolites of toluene and xylene (n=267) and for chromium (n=156)	- Sensitivity and specificity in comparison to urine samples - Kappa in comparison to urine samples	- Sensitivities = 0.30 using the generic questionnaire and 0.40 using the job specific questionnaire for toluene/xylene, 0.21 and 0.28 respectively for chromium - Specificities = 0.92 and 0.93 for toluene/xylene, 0.94 and 0.93 for chromium - Kappas = 0.24 and 0.37 for toluene/xylene, 0.18 and 0.26 for chromium
Stewart *et al*, 2000[93]	30 jobs randomly selected from each of 10 formaldehyde and resin manufacturing plants participating in a cohort study in the US	Three industrial hygienists assessed exposure to formaldehyde into 4 ordinal categories, in 6 stages using increasing amounts of data, including job, department, industry, date, and plant reports	Original exposure estimates developed for the cohort study by 2 industrial hygienists using walkthrough surveys, historical documents, formaldehyde measurement data, and interviews with long-term employees, process flow and change data, exposure controls	- Relative bias and relative standard deviation, compared to mean exposure estimates - Intraclass correlation coefficients for inter-rater agreement	- Mean relative bias of –0.12 and relative standard deviation of 0.54 with the first stage of information - Mean relative bias less than ±0.05 and relative standard deviation less than 0.5 in the remaining 5 stages of data - Intraclass correlations ranged from 0.42 to 0.51, with no clear pattern according to amount of data provided

welding and soldering fumes, oils and greases, and solvents than for specific agents such as phenol, vinyl chloride, acrylonitrile, and toluene di-isocyanate. Post and colleagues[81] found that hygienists were able to rank exposures to methylene chloride better than styrene, perhaps because of differences in the odour thresholds. These studies suggest that experts are influenced by some of the same factors as subjects—that is, sensory perceptions affect judgements, and estimation is easier for broad classes of agents than for specific chemical compounds.

Some studies have examined the extent to which prior expertise affects assessments. In a study of fungicides in sawmills, Teschke and colleagues[51] found that lumber industry hygienists had higher inter-rater agreement, but the validity of their exposure estimates was very similar to that of hygienists from other industry sectors. In their study of pesticide use in fruit growing, de Cock and colleagues[86] did not find a consistent pattern for inter-rater agreement between their three groups of experts, but hygienists and pesticide experts gave more valid ratings than fruit growing experts, suggesting that the critical expertise is understanding the exposure rather than intimate knowledge of the work activity.

The evidence to date on expert assessments supports the belief that experts are better able to estimate exposures than study subjects, though this evidence is not as strong or consistent as epidemiologists might hope. Experts' estimates can be so poor that true exposure–effect relations are obscured or even reversed in direction,[76] indicating the value of testing reliability and validity for the most important exposures in a study, and ensuring that experts have access to information that may incrementally improve performance, such as subject reported exposures and work conditions, and measurement data.

QUANTITATIVE DATA

The above review of exposure assessment methods in common use in case–control studies indicates that there remains much room for improvement. Incorporation of quantitative exposure measurements into case–control studies has always seemed a quixotic goal, but developments in occupational hygiene data collection, management, and analysis suggest several means to systematically include measurements in exposure estimation for population based studies.

Exposure databases

Exposure databases are not new—data on ionizing radiation exposures have been collected on designated workers since 1950 in Canada[98] and elsewhere. The Mine Safety and Health Administration in the United States has been storing data on coal dust, silica dust, and other mining exposures since 1970,[99] and the German Institute for Occupational Safety began its comprehensive chemical exposure database a couple of years later.[100] However, the number of such databases[98–107] has increased substantially over the past two decades (see examples in table 5), with advances in computer technology. International conferences have been held to promote thoughtful data collection and compatibility between data sets.[108–110]

Administrative exposure data sets have only rarely been used in case–control studies, but they present many interesting possibilities. Databases such as the National Dose Registry in Canada offer the opportunity to assign cumulative radiation exposures over five decades to individual study subjects, since personal identifying information has been retained in the registry.[98] However, this level of detail is the exception.

Most exposure databases include job and industry information, but no data identifying individuals whose exposures were measured. This means that average exposures for an occupation and/or industry can be calculated and used to estimate exposures of subjects with those jobs. Of course, this method does not account for within job variations in exposure, and is not helpful where there are no measurements for a particular job–exposure combination. These problems might be addressed in part by using database information as only one component of exposure assessment. For example, Stewart and Stewart[97] proposed supplementing detailed occupational questionnaires and job specific modules with data from the US Occupational Safety and Health Administration Integrated Management Information System. The potential for tailoring database information to individual subjects depends on the supplementary data fields included in the database. For example, if information on tasks, control measures, raw materials, etc are included, as in the French COLCHIC system,[106] reports by subjects about these conditions in their own worksites could be used to adjust job based exposure estimates.

Given that exposure measurements in administrative databases are not likely to have arisen from subjects' workplaces, validity and reliability studies of estimates derived from databases should be conducted. There are other possible problems with administrative data. The original purpose of data collection (for example, complaint, compliance, research), changes in measurement techniques, and clustering of data in one or a few workplaces, all have the potential to bias exposure measurements. If information on these factors is included in the database, it may be possible to adjust for any biases using empirical modelling.[111]

Determinants of exposure studies

A method which holds promise for improving the validity of exposures assessed by questionnaires is to guide the formulation of questions and interpretation of responses using results of "determinants of exposure" studies. Such studies examine which characteristics (for example, workplace, process, employee) are associated with increased or decreased exposure levels. There is a growing body of literature on the determinants of exposure in a wide range of industries.[112] Factors which have been examined as potential exposure determinants are extremely varied, for example, type of facility, worksite construction materials, industrial processes, automation, raw materials and machinery used, geographical location, indoor versus outdoor work, ambient environmental conditions, tasks, work practices, training, ventilation, use of enclosures, skin contact, protective clothing, and cleaning facilities.

Translating these data into questions useful to assess exposures in case–control studies is not a simple process. Questions must be answerable by study subjects, therefore determinants such as tasks and equipment will be more feasible to query than technical ones such as air flow rates of ventilation systems. Given that determinants data are likely not to have been collected in the worksites or residences of the study subjects, it would also be necessary to consider the transferability of the information. Where determinants studies show consistent patterns and where there is greater variability between the determinants of interest than between worksites, it should be possible to develop useful questions to distinguish exposure levels.

Where sufficient information on exposure determinants is not available in existing scientific literature, researchers might consider designing their own determinants studies prior to embarking on an epidemiological investigation. There are some interesting examples of studies which have measured exposures in a large number of worksites to create predictive models for use in questionnaire based epidemiological research.[112 113]

Table 5 Examples of administrative exposure databases

Database name	Country/agency	Descriptions in scientific literature	Industries/agents	Start year/types of data
NDR National Dose Registry	Canada Radiation Protection Bureau, Health Canada	Ashmore et al, 1998[98]	80 occupations in 14 industry sectors ionizing radiation	From 1950 Types of data: subject ID, job, industry, date, sex, age
MIDAS Mine Inspection Data Analysis System	United States Mine Safety and Health Administration (MSHA) (some exposures measured by mine operators)	Watts and Parker, 1995[99]	Mining, milling coal dust, quartz dust, ~130 other substances, and noise	From 1970 Types of data: agent, exposure level, SIC code, date, occupation, mine location and identification, mine production level, mine type, mining method, ventilation code, number of employees
MEGA	Germany Berufsgenossenshcaftliches Institut für Arbeitssicherheit (BIA) (exposures measured by regional accident insurance institutes and private companies)	Vinzents et al, 1995[101]; Stamm, 2001[100]	Many industries 420 chemical agents	From 1972 Types of data: agent, exposure level, firm, industry, workplace, process, raw materials and products, work environment, measurement and analytic methods
IMIS Integrated Management Information System	United States federal Occupational Safety and Health Administration (OSHA), and some state plan enforcement agencies	Stewart and Rice, 1990[102]; Nelson et al, 1995[103]	All industries, except mining and agriculture >500 chemical and physical agents	From 1979 Types of data: agent, exposure level, inspection date, employer name and address, number of employees, SIC code, reason for inspection; job title, purpose of sampling
EXPO Exsponeringsregister	Norway National Institute of Occupational Health	Fjeldstad and Woldbaek, 1991[104]; Vinzents et al, 1995[101]	Many industries	From 1985 Types of data: agent, concentration in blood, urine, air, employee name, industry, job, substance, ISIC code
NEDB National Exposure Database	United Kingdom Health and Safety Executive (some exposures measured by industry)	Burns and Beaumont, 1989[105]; Vinzents et al, 1995[101]	All industries chemical agents	From 1986 Types of data: agent, exposure level, date, company and location, number of male and female employees, SIC code, job, process, monitoring method and duration, reason for visit, ventilation and personal protective equipment use, representativeness
COLCHIC Systéme de Collecte des Données Recueillies par les Laboratoires de Chimie de l'INRS et de CRAM	France Institut National de Recherche et de Sécurité (INRS) and Caisse Régionale d'Assurance Maladie (CRAM)	Vinzents et al, 1995[101]; Vincent and Jeandel, 2001[106]	All industries except mining, energy, rail, agriculture, and government ~600 chemical substances	From 1987 Types of data: agent, exposure level, sampling method and analysis, factory, industry, work operation, no. workers exposed, ventilation, use of protective equipment, temperature, representiveness
PHED Pesticide Handlers Exposure Database	Canada and United States Health Canada, US Environmental Protection Agency (EPA), National Agricultural Chemical Association (exposures measured by pesticide manufacturers)	Leighton and Nielsen, 1995[107]	Pesticide application pesticides, but active ingredient name not released, data reported by pesticide type and formulation type	From 1992 Types of data: dermal and inhalation exposure levels (by mass of unspecified "active ingredient") for pesticide loaders, applicators, mixers, and flaggers; site description, application method and rate, cab type, employee's experience, sampling duration

Subject specific exposure measurements

An avenue for exposure assessment which has only rarely been used in case–control studies is direct exposure measurements of the study subjects. For outcomes with short induction and latency periods, measurements of current exposures may serve as reasonable surrogates for exposures in the disease induction period. Measurements of exogenous agents in biological tissues assess the body burden at the time the sample was taken, but can provide information on historical exposures in a limited set of circumstances—that is, where the chemical of interest has a sufficiently long biological half life, and the body burden is not affected by the disease or its treatment.[114]

There are a number of case–control studies which have used exposure measurements. For example, Floderus and colleagues,[115] in a case–control study of brain cancer and leukaemia, made 924 magnetic field measurements of 169 jobs (those held longest) in the workplaces of study subjects. Veulemans and colleagues[116] measured urinary metabolites of methoxy and ethoxy acetic acid in 1019 infertile men and 475 controls. Tielemans and colleagues[43] measured levels of industrial solvents in the urine of 99 cases with reduced semen quality and 27 controls. Caldwell and colleagues[117], and Scheele and colleagues[118 119] measured pesticide levels in bone marrow and serum in adult and childhood cancer cases and controls.

One of the great difficulties of measuring exposures in case–control studies is the potentially wide geographical dispersion of study subjects. This logistical difficulty might be possible to overcome with advances in sample collection and preservation methods. For example, urine and semen samples can be collected by study participants in their homes and shipped to the study site. Blood samples can be collected by a family physician or local clinic and forwarded to the appropriate laboratory for analysis. Advances in occupational hygiene monitoring equipment over the past several decades also make it reasonable to consider mailing simple sampling equipment, such as passive dosimeters or electronic data loggers, to study subjects for exposure assessment. As an example, Kromhout and colleagues[120] mailed magnetic field dosimeters to subjects of a cohort study in geographically dispersed locations in the United States.

If these more quantitative methods of exposure assessment are adopted in case–control studies, the issues involved will be similar to those faced by researchers using measured exposure data in cohort or cross sectional studies—that is, sampling strategy issues such as how many measurements to take, and epidemiological analysis issues such as whether and how to group subjects.[120–124]

DISCUSSION

This review illustrates that exposure assessment methods typically used in case–control studies, though often thought of as distinct from each other, are inter-related and interdependent. Generic job–exposure matrices have most often been based on experts' judgements. Some JEMs use self reports to provide estimates of the proportions of exposed individuals in each job.[40 125] Assessments by experts almost always rely on self reports as the starting point, using job history data at a minimum, but often utilising subjects' exposure reports and sometimes information on work tasks and conditions. Self reports themselves are answers to questions formulated by experts. Not surprisingly then, the results of validity and reliability studies of these estimation methods show similarities. Foremost is the conclusion that questionnaire based methods commonly used in case–control studies do not produce consistently valid and reliable results, underscoring the importance of continued development and testing of methods.

Evidence to date also reveals a number of strategies which can optimise these exposure estimation methods. Self reported exposure estimates may be improved by using terms familiar to workers, by asking about exposures that can be smelled, seen, or felt by subjects, and by presenting benchmarks against which exposures can be gauged. Instead of asking about exposures themselves, subjects can be asked about factors related to exposures, but more likely to be known and accurately recalled (for example, tasks, raw materials, equipment, processes); empirical models can be used to relate these factors to exposures. Experts find it easier to make estimates for commonly used agents and classes of chemicals, rather than arcane individual agents. In addition, experts' assessments may be improved by providing experts with exposure measurement data, information about the properties of the agents, and data reported by subjects about their work conditions and exposures. Occupational history taking would benefit from techniques such as chronicling of major life events to enhance recall,[126] particularly where the job history is complex, for example, multiple short term jobs or jobs in the distant past.

There are a number of issues important to exposure estimation methods which have not yet received much attention. Although studies have investigated the effect of time since a job was held on the quality of an occupational history,[7 10 11 13 14 17 18] the effect of the duration of elapsed time on the validity of subjects' or experts' exposure estimates has not been examined. In many epidemiological investigations using experts, more than one expert is used, but the optimum number of experts and the value of independent versus consensus estimates has rarely been tested.[88]

Although many studies examining the validity of exposure estimation methods indicated rather disappointing performance, it is important to remember that gold standards are never perfect. This was particularly extreme for studies of generic job exposure matrices; all comparisons, except one, were to self reported or expert estimates of exposure. Studies of self reports and expert assessments more frequently used measured exposure levels as the basis for evaluation, usually using one of two techniques. Where continuous exposure estimates were made, proportions of variance explained or correlations were calculated. In almost every case, exposure estimates assigned to a study subject were compared to measurements of exposure taken on individual days, thus requiring the estimation method to predict not only subject to subject variability in exposure, but also day to day variability within subject. Short term variations in exposure are not thought to be related to body burden or disease development, except where biological half lives are very short.[127] Therefore, for studies of chronic diseases, it would be more reasonable to test whether an estimation method is related to the long term average exposure level. In studies where only the presence or absence of exposure was estimated, sensitivity, specificity, and/or positive predictive value were used as the measures of validity. The issue of individual daily measurements of exposure versus long term average exposure is also a consideration here. But in addition, calculations of sensitivity and specificity require that the gold standard measurements be dichotomised. The definition of a value above which exposure "exists" is difficult and often arbitrary, for example, the analytical detection limit has often been used. Ideally the cut point would be set at a level above which there is disease potential, but case–control studies are often conducted at the initial stages of aetiological research, before such knowledge has accumulated. Another consideration in defining what constitutes exposure is that in most case–control studies exposure prevalence is low, so specificity is more important than sensitivity for minimising attenuation in exposure–response relations.[94] Therefore it is usually better to use a stringent definition of exposure (for example, only highly exposed subjects considered exposed) in epidemiological analyses.

There is room for an increase in the sophistication of validation studies. In cohort and cross sectional studies, where quantitative measurements are usually made, the major methodological developments in exposure assessment in the past decade have focused on the benefit of grouping study subjects for analysis, based on similarities in exposure. By assigning subjects the mean exposure of their group, the precision of the exposure estimate is increased, and the error structure approximates the Berkson error model. The advantage is a reduction in misclassification bias that can attenuate the observed association in exposure–response analyses.[121 123 128] Since the advantage of grouping was recognised, methodological research on quantitative exposure measurements for epidemiology has been directed at finding the best ways to group study subjects.[120 122 124] It seems reasonable that validity testing of experts' or subjects' estimates should incorporate these methods. Thus in validity studies, instead of comparing exposure estimates for individual subjects to individual exposure measurements, the exposure estimation method could be used to group subjects and these groups compared to optimal groupings based on exposure measurements. This idea is an extension of that of Kromhout and colleagues,[47] who examined the proportion of between group exposure variability explained by exposure estimates for individual subjects, as a way to exclude day to day variations in measured exposures. The proposed approach will provide a more reasonable (and likely less stringent) test of the validity of estimation methods.

In summary, among the exposure estimation methods in common use today, expert assessment is usually the best approach. All exposure estimation methods, whether by subjects or experts, can have low validity and reliability; they therefore need to be carefully designed using evidence about techniques which improve performance and, where possible, tested. A new generation of case–control studies could evolve if methods which incorporate exposure measurements are adopted. Direct measurements of study subjects, if the science and logistics permit, would be ideal. A more frequently feasible method would be to combine questionnaires and measurements—that is, subjects can be asked about factors shown to be related to exposures in determinants of exposure models, and the models used to predict exposure levels. If quantitative methods are embraced, many of the methodological developments in exposure assessment for cohort and cross sectional studies could be applied directly to case–control studies. In addition, the inclusion of exposure measurement data would extend the utility of results of case–control studies—in risk assessments and exposure standard setting.

ACKNOWLEDGEMENTS

This work was supported in part by an Izaak Walton Killam Memorial Fellowship. The authors greatly appreciate the helpful comments of

Main messages

- The main techniques currently used for exposure assessment in population based case–control studies include generic job–exposure matrices (JEMs), exposure self reports by study subjects, and assessment of exposures by experts.
- An extensive literature is now available with which to evaluate the validity and reliability of these methods.
- Most generic JEMs do not perform well, no matter how they are evaluated. Self reported exposures are usually better than generic JEMs, but vary greatly in validity and reliability. The accuracy of self reports is improved by using terms familiar to employees, asking about agents that can be sensed, and providing relative or absolute benchmarks against which to gauge exposures. Expert assessments are usually somewhat better than self reports, though validity and reliability are also variable. Experts are aided in their assessments by subject reported data on exposures and work conditions, and measurement data. Careful design and evaluation are required for all exposure estimation techniques.
- Exposure assessment methods which incorporate quantitative measurements are difficult in population based studies, but increasingly possible with improvements in measurement techniques and administrative databases. These methods offer the possibility of a new generation of exposure assessment in case–control studies.

Jane Schroeder, Hugh Davies, the reviewers, and University of North Carolina Epidemiology Department students, staff, and faculty who participated in a seminar series on this subject.

APPENDIX: SELECTED TERMS USED IN VALIDITY AND RELIABILITY STUDIES

The following is a brief and simplified overview of some terminology used in the validity and reliability studies reviewed in this paper. For a full understanding, it is best to consult the methodological literature, some of which is cited below.

Note that although the following discussion separates terminology according to whether the measures are usually used in validity versus reliability studies, the measures are sometimes used in either type of study.

Common measures of *validity* when using a *dichotomous* classification of exposure—that is, exposed versus unexposed

- Sensitivity—proportion of those truly exposed who are classified as exposed by the assessment method being evaluated (values between 0 and 1).
- Specificity—proportion of those truly not exposed who are classified as unexposed by the assessment method being evaluated (values between 0 and 1).
- Positive predictive value—proportion of those classified as exposed who are truly exposed (values between 0 and 1). This proportion depends on the sensitivity and specificity of the classification method and the prevalence of exposure in the population being assessed.

The effect of misclassification of dichotomous exposure estimates has been described in a number of methodological papers (see Flegal and colleagues[76] and Dosemeci and Stewart[94]). Non-differential misclassification will usually attenuate relative risk estimates towards the null value. The resulting relative risk estimate will depend on the strength of the true relative risk and the extent of misclassification. If sensitivity and specificity are so low that their sum is less than 1, a relative risk estimate using the estimated exposure values will indicate an association opposite in direction to the true association.[76] When the prevalence of exposure is low, as in most population based case–control studies, it is important for the specificity to be as high as possible (that is, >0.9, and ideally very close to 1) to ensure that the small exposed group is not diluted by a large number of unexposed individuals.[94]

Common measures of *validity* when using *continuous* measures of exposure

- R^2—proportion of the variance in true exposure explained by the exposure estimation method being evaluated (values between 0 and 1).
- Pearson r—correlation coefficient (values between −1 and 1); sign the same as the slope of the relation between the true exposure and the estimated exposure, and magnitude related to degree of linear association between the two. The square of r is R^2.
- Spearman rank r—rank correlation coefficient (values between −1 and 1); same as Pearson r, except that it is based on the ranks of the true and estimated exposures, rather than the data itself.

The impact of misclassification of continuous exposure estimates is generally the same as for categorical data, and has been described in a number of papers (see Armstrong[121]). Non-differential misclassification will usually attenuate relative risk estimates towards the null value, with the degree of attenuation dependent on the true relative risk and the extent of misclassification. If the correlation coefficient is negative, a relative risk estimate using the estimated exposure values will indicate an association opposite in direction to the true association.

Common measures of *reliability*

- Percent agreement—percent of exposure estimates, estimated on two different occasions or by two different raters, which agree with each other (values between 0 and 100). This measure does not account for the proportion of agreement likely by chance alone.
- Kappa—proportion of agreement beyond that expected by chance alone (values between $-\infty$ and 1); for categorical measures of exposure.
- Intraclass correlation—proportion of the total variability as a result of differences in exposure between subjects (rather than differences between repeated estimates for individual subjects) (values between 0 and 1); for continuous estimates of exposure.

Reliability (precision) is a component of validity, with the effect of non-differential misclassification indicated above. Landis and Koch[129] gave the following verbal interpretations of the strength of the kappa statistic; these have also been used to describe intraclass correlations: poor = <0; 0–0.2 = slight agreement; 0.21–0.40 = fair agreement; 0.41–0.60 = moderate agreement; 0.61–0.80 = substantial agreement; 0.81–1 = almost perfect agreement.

Authors' affiliations

K Teschke, Department of Health Care and Epidemiology, University of British Columbia, Vancouver, BC, Canada
A F Olshan, C G Parks, M Schulz, Department of Epidemiology, University of North Carolina, Chapel Hill, NC, USA
J L Daniels, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA
A J De Roos, Occupational Eidemiology Branch, National Cancer Institute, Rockville, MD, USA
T L Vaughan, Department of Epidemiology, University of Washington, and the Fred Hutchinson Cancer Research Center, Seattle, WA, USA

REFERENCES

1 **Stewart PA**, Dosemeci M. A bibliography for occupational exposure assessment for epidemiologic studies. *Am Ind Hyg Assoc J* 1994;**55**:1178–87.
2 **McGuire V**, Nelson LM, Koepsell TD, *et al*. Assessment of occupational exposures in community-based case-control studies. *Annu Rev Public Health* 1998;**19**:35–53.
3 **Correa A**, Stewart WF, Yeh H, *et al*. Exposure measurement in case-control studies: reported methods and recommendations. *Epidemiol Rev* 1994;**16**:18–31.
4 **Samet JM**, Speizer FE, Gaensler AE. Questionnaire reliability and validity in asbestos exposed workers. *Bull Eur Physiopathol Respir* 1978;**14**:177–88.
5 **Baumgarten M**, Siemiatycki J, Gibbs GW. Validity of work histories obtained by interview for epidemiologic purposes. *Am J Epidemiol* 1983;**118**:583–91.
6 **Koskela R**, Kolari PJ, Jarvinen E, *et al*. Completeness of occupational history and occurrences of work-related diseases. *Scand J Work Environ Health* 1984;**10**:455–9.
7 **Rosenberg CR**, Mulvihill MN, Fischbein A, *et al*. An analysis of the validity of self reported occupational histories using a cohort of workers exposed to PCBs. *Br J Ind Med* 1987;**44**:702–10.

8 **Stewart WF**, Tonascia JA, Matanoski GM. The validity of questionnaire-reported work history in live respondents. *J Occup Med* 1987;**29**:795–800.
9 **Eskenazi B**, Pearson K. Validation of a self-administered questionnaire for assessing occupational and environmental exposures of pregnant women. *Am J Epidemiol* 1988;**128**:1117–29.
10 **Bond GG**, Bodner KM, Sobel W, *et al*. Validation of work histories obtained from interviews. *Am J Epidemiol* 1988;**128**:343–51.
11 **Bourbonnais R**, Meyer F, Theriault G. Validity of self reported work history. *Br J Ind Med* 1988;**45**:29–32.
12 **Rona RJ**, Mosbech J. Validity and repeatability of self-reported occupational and industrial history from patients in EEC countries. *Int J Epidemiol* 1989;**18**:674–9.
13 **Brisson C**, Vezina M, Bernard PM, *et al*. Validity of occupational histories obtained by interview with female workers. *Am J Ind Med* 1991;**19**:523–30.
14 **Warneryd B**, Thorslund M, Ostlin P. The quality of retrospective questions about occupational history—a comparison between survey and census data. *Scand J Soc Med* 1991;**19**:7–13.
15 **Booth-Jones AD**, Lemasters GK, Succop P, *et al*. Reliability of questionnaire information measuring muscolskeletal symptoms and work histories. *Am Ind Hyg Assoc J* 1998;**59**:20–4.
16 **Brower PS**, Attfield MD. Reliability of reported occupational history information for US coal miners, 1969 to 1977. *Am J Epidemiol* 1998;**148**:920–6.
17 **Rosenberg CR**. An analysis of the reliability of self reported work histories from a cohort of workers exposed to polychlorinated biphenyls. *Br J Ind Med* 1993;**50**:822–6.
18 **Birdsong WH**, Lash AA, Thayer S, *et al*. The validity of study group assignments based on occupational histories obtained from questionnaires. *J Occup Med* 1992;**34**:940–5.
19 **Kennedy SM**, Le Moual N, Choudat D, *et al*. Development of an asthma specific job exposure matrix and its application in the epidemiological study of genetics and environment in asthma. *Occup Environ Med* 2000;**57**:635–41.
20 **Hoar SK**, Morrison AS, Cole P, *et al*. An occupation and exposure linkage system for the study of occupational carcinogenesis. *J Occup Med* 1980;**22**:722–6.
21 **Pannett B**, Coggon D, Acheson ED. A job-exposure matrix for use in population based studies in England and Wales. *Br J Ind Med* 1985;**42**:777–83.
22 **Ferrario F**, Continenza D, Pisani P, *et al*. Description of a job-exposure matrix for sixteen agents which are or may be related to respiratory cancer. In: Hogstedt C, Reuterwall C, eds. *Progress in occupational epidemiology*. Amsterdam: Elsevier, 1988:379–82.
23 **Orlowski E**, Pohlabeln H, Berrino F, *et al*. Retrospective assessment of asbestos exposure—II. At the job level: complementarity of job-specific questionnaire and job exposure matrices. *Int J Epidemiol* 1993;**22**:S96–105.
24 **Plato N**, Steineck G. Methodology and utility of a job-exposure matrix. *Am J Ind Med* 1993;**23**:491–502.
25 **Sieber WK**, Sundin DS, Frazier TM, *et al*. Development, use, and availability of a job exposure matrix based on national occupational hazard survey data. *Am J Ind Med* 1991;**20**:163–74.
26 **Kauppinen T**, Toikkanen J, Pukkala E. From cross-tabulations to multipurpose exposure information systems: a new job-exposure matrix. *Am J Ind Med* 1998;**33**:409–17.
27 **Floderus B**, Persson T, Stenlund C. Magnetic field exposures in the workplace: reference distribution and exposures in occupational groups. *Int J Occup Environ Health* 1996;**2**:226–38.
28 **Cherry NM**, Labreche FP, McDonald JC. Organic brain damage and occupational solvent exposure. *Br J Ind Med* 1992;**49**:776–81.
29 **Hinds MW**, Kolonel LN, Lee J. Application of a job-exposure matrix to a case-control study of lung cancer. *JNCI* 1985;**75**:193–7.
30 **Linet MS**, Stewart WF, Van Natta ML, *et al*. Comparison of methods for determining occupational exposure in a case-control interview study of chronic lymphocytic leukemia. *J Occup Med* 1987;**29**:136–41.
31 **Cicioni C**, London SJ, Garabrant DH, *et al*. Occupational asbestos exposure and mesothelioma risk in Los Angeles County: application of an occupational hazard survey job-exposure matrix. *Am J Ind Med* 1991;**20**:371–9.
32 **Kauppinen TP**, Mutanen PO, Seitsamo JT. Magnitude of misclassification bias when using a job-exposure matrix. *Scand J Work Environ Health* 1992;**18**:105–12.
33 **Kromhout H**, Heederik D, Dalderup LM, *et al*. Performance of two general job-exposure matrices in a study of lung cancer morbidity in the Zutphen Cohort. *Am J Epidemiol* 1992;**136**:698–711.
34 **Ahrens W**, Jockel KH, Brochard P, *et al*. Retrospective assessment of asbestos exposure—I. Case-control analysis in a study of lung cancer efficiency of job-specific questionnaires and job exposure matrices. *Int J Epidemiol* 1993;**22**:S83–S95.
35 **Luce D**, Gerin M, Berrino F, *et al*. Sources of discrepancies between a job exposure matrix and a case by case expert assessment for occupational exposure to formaldehyde and wood-dust. *Int J Epidemiol* 1993;**22**:S113–20.
36 **Roeleveld N**, Zielhuis GA, Gabreels F. Mental retardation and parental occupation: a study on the applicability of job exposure matrices. *Br J Ind Med* 1993;**50**:945–54.
37 **Stengel B**, Pisani P, Limasset JC, *et al*. Retrospective evaluation of occupational exposure to organic solvents: questionnaire and job exposure matrix. *Int J Epidemiol* 1993;**22**:S72–82.
38 **Stucker I**, Bouyer J, Mandereau L, *et al*. Retrospective evaluation of the exposure to polycyclic aromatic hydrocarbons: comparative assessment with a job exposure matrix and by experts in industrial hygiene. *Int J Epidemiol* 1993;**22**:S106–12.
39 **Le Moual N**, Orlowski E, Schenker MB, *et al*. Occupational exposures estimated by means of job exposure matrices in relation to lung function in the PAARC survey. *Occup Environ Med* 1995;**52**:634–43.
40 **McNamee R**. Retrospective assessment of occupational exposure to hydrocarbons—job-exposure matrices versus expert evaluation of questionnaires. *Occup Hyg* 1996;**3**:137–43.
41 **Hawkes AP**, Wilkins JR. Assessing agreement between two job-exposure matrices. *Scand J Work Environ Health* 1997;**23**:140–8.
42 **Rybicki BA**, Johnson CC, Peterson EL, *et al*. Comparability of different methods of retrospective exposure assessment of metals in manufacturing industries. *Am J Ind Med* 1997;**31**:36–43.
43 **Tielemans E**, Heederik D, Burdorf A, *et al*. Assessment of occupational exposures in a general population: comparison of different methods. *Occup Environ Med* 1999;**56**:145–51.
44 **Louik C**, Frumkin H, Ellenbecker MJ, *et al*. Use of a job-exposure matrix to assess occupational exposures in relation to birth defects. *J Occup Environ Med* 2000;**42**:693–703.
45 **Benke G**, Sim M, Fritschi L, *et al*. Comparison of occupational exposure using three different methods: hygiene panel, job exposure matrix (JEM), and self-reports. *Appl Occup Environ Hyg* 2001;**16**:84–91.
46 **doPico GA**. Epidemiologic basis for dose-response criteria. *Ann Am Conf Gov Ind Hyg* 1982;**2**:189–95.
47 **Kromhout H**, Oostendorp Y, Heederik D, *et al*. Agreement between qualitative exposure estimates and quantitative exposure measurements. *Am J Ind Med* 1987;**12**:551–62.
48 **Järvolm B**, Sandén A. Estimating asbestos exposure: a comparison of methods. *J Occup Med* 1987;**29**:361–3.
49 **Hertzman C**, Teschke K, Dimich-Ward H, *et al*. Validity and reliability of a method for retrospective evaluation of chlorophenate exposure in the lumber industry. *Am J Ind Med* 1988;**14**:703–13.
50 **Pron GE**, Burch JD, Howe GR, *et al*. The reliability of passive smoking histories reported in a case-control study of lung cancer. *Am J Epidemiol* 1988;**127**:267–73.
51 **Teschke K**, Hertzman C, Dimich-Ward H, *et al*. A comparison of exposure estimates by worker raters and industrial hygienists. *Scand J Work Environ Health* 1989;**15**:424–9.
52 **Ahlborg GA**. Validity of exposure data obtained by questionnaire. Two examples from occupational reproductive studies. *Scand J Work Environ Health* 1990;**16**:284–8.
53 **Bachmann M**, Myers JE. Grain dust and respiratory health in South African milling workers. *Br J Ind Med* 1991;**48**:656–62.
54 **Holmes E**, Garshick E. The reproducibility of the self-report of occupational exposure to asbestos and dust. *J Occup Med* 1991;**33**:135–8.
55 **Joffe M**. Validity of exposure data derived from a structured questionnaire. *Am J Epidemiol* 1992;**135**:564–70.
56 **Walter SD**, Marrett LD, Shannon HS, *et al*. The association of cutaneous malignant melanoma and fluorescent light exposure. *Am J Epidemiol* 1992;**135**:749–62.
57 **Blair A**, Zahm SH. Patterns of pesticide use among farmers: implications for epidemiologic research. *Epidemiology* 1993;**4**:55–62.
58 **Fonn S**, Groeneveld HT, deBeer M, *et al*. Relationship of respiratory health status to grain dust in a Witwatersrand grain mill: comparison of workers' exposure assessments with industrial hygiene survey findings. *Am J Ind Med* 1993;**24**:401–11.
59 **van der Gulden JWJ**, Jansen IW, Verbeek ALM, *et al*. Repeatability of self-reported data on occupational exposure to particular compounds. *Int J Epidemiol* 1993;**22**:284–7.
60 **Halpin DMG**, Graneek BJ, Lacey J, *et al*. Respiratory symptoms, immunological responses, and aeroallergen concentrations at a sawmill. *Occup Environ Med* 1994;**51**:165–72.
61 **Savitz DA**, Baird N, Dole N. Agreement among textile industry exposures during pregnancy based on work description, job title, and self-report. *J Exp Anal Environ Epidemiol* 1994;**4**:513–24.
62 **Teschke K**, Kennedy SM, Olshan AF. Effect of different questionnaire formats on reporting of occupational exposures. *Am J Ind Med* 1994;**26**:327–37.
63 **Fritschi L**, Siemiatycki J, Richardson L. Self-assessed versus expert-assessed occupational exposures. *Am J Epidemiol* 1996;**144**:521–7.
64 **Rodvall Y**, Ahlbom A, Spannare B, *et al*. Glioma and occupational exposure in Sweden, a case-control study. *Occup Environ Med* 1996;**53**:526–32.
65 **Wiktorin C**, Hjelm E, Winkel J, Koster M, Stockholm MUSIC I Study Group. Reproducibility of a questionnaire for assessment of physical load during work and leisure time. *J Occup Environ Med* 1996;**38**:190–7.
66 **Calvert GM**, Mueller CA, O'Neill VL, *et al*. Agreement between company recorded and self-reported estimates of duration and frequency of occupational fumigant exposure. *Am J Ind Med* 1997;**32**:364–8.
67 **Ising H**, Babisch W, Kruppa B, *et al*. Subjective work noise: a major risk factor in myocardial infarction. *Sozial-Und Praventivmedizin* 1997;**42**:216–22.
68 **Nieuwenhuijsen MJ**, Noderer KS, Schenker MB. The relation between subjective dust exposure estimates and quantitative dust exposure measurements in California agriculture. *Am J Ind Med* 1997;**32**:355–63.
69 **Willemsen MC**, Brug J, Uges DRA, *et al*. Validity and reliability of self-reported exposure to environmental tobacco smoke in work offices. *J Occup Environ Med* 1997;**39**:1111–14.
70 **Nordstrom DL**, Vierkant RA, Layde PM, *et al*. Comparison of self-reported and expert-observed physical activities at work in a general population. *Am J Ind Med* 1998;**34**:29–35.

71 **Neale AV**, Demers RY, Severson RK. Consistency of occupational exposure history from pattern and model makers. *J Occup Environ Med* 2000;**42**:76–82.
72 **Palmer KT**, Haward B, Griffin MJ, *et al.* Validity of self reported occupational exposures to hand transmitted and whole body vibration. *Occup Environ Med* 2000;**57**:237–41.
73 **Zahm SH**, Blair A. Cancer among migrant and seasonal farmworkers: an epidemiologic review and research agenda. *Am J Ind Med* 1993;**24**:753–66.
74 **Blair A**, Stewart PA, Kross B, *et al.* Comparison of two techniques to obtain information on pesticide use from Iowa farmers by interview. *J Agric Safety Health* 1997;**3**:1–8.
75 **Teschke K**, Smith JC, Olshan AF. Evidence of recall bias in volunteered vs. prompted responses about occupational exposures. *Am J Ind Med* 2000;**38**:385–8.
76 **Flegal KM**, Brownie C, Haas JD. The effects of exposure misclassification on estimates of relative risk. *Am J Epidemiol* 1986;**123**:736–51.
77 **Woitowitz HJ**, Schacke G, Woitowitz R. Ranking estimation of the dust exposure and industrial-medical epidemiology. *Staub-Reinhalt Luft* 1970;**30**:15–20.
78 **Goldberg MS**, Siemiatycki J, Gerin M. Inter-rater agreement in assessing occupational exposure in a case-control study. *Br J Ind Med* 1986;**43**:667–76.
79 **Ciccone G**, Vineis P. Inter-rater agreement in the assessment of occupational exposure to herbicides. *Med Lav* 1988;**79**:363–7.
80 **Hawkins NC**, Evans JS. Subjective estimation of toluene exposures: a calibration study of industrial hygienists. *Appl Ind Hyg* 1989;**4**:61–8.
81 **Post W**, Kromhout H, Heederik D, *et al.* Semiquantitative estimates of exposure to methylene chloride and styrene: the influence of quantitative exposure data. *Appl Occup Environ Hyg* 1991;**6**:197–204.
82 **Dovan T**, Kaune WT, Savitz DA. Repeatability of measurements of residential magnetic fields and wire codes. *Bioelectromagnetics* 1993;**14**:145–59.
83 **Macaluso M**, Delzell E, Rose V, *et al.* Inter-rater agreement in the assessment of solvent exposure at a car assembly plant. *Am Ind Hyg Assoc J* 1993;**54**:351–9.
84 **Takahashi K**, Case BW, Dufresne A, *et al.* Relation between lung asbestos fibre burden and exposure indices based on job history. *Occup Environ Med* 1994;**51**:461–9.
85 **Armstrong TW**, Pearlman ED, Schnatter AR, *et al.* Retrospective benzene and total hydrocarbon exposure assessment for a petroleum marketing and distribution worker epidemiology study. *Am Ind Hyg Assoc J* 1996;**57**:333–43.
86 **de Cock J**, Kromhout J, Heederik D, *et al.* Experts' subjective assessment of pesticide exposure in fruit growing. *Scand J Work Environ Health* 1996;**22**:425–32.
87 **Segnan N**, Ponti A, Ronco GF, *et al.* Comparison of methods for assessing the probability of exposure in metal plating, shoe and leather goods manufacture and vine growing. *Occup Hyg* 1996;**3**:199–208.
88 **Benke G**, Sim M, Forbes A, *et al.* Retrospective assessment of occupational exposure to chemicals in community-based studies: validity and repeatability of industrial hygiene panel ratings. *Int J Epidemiol* 1997;**26**:635–42.
89 **McGuire V**, Longstreth WT, Nelson LM, *et al.* Occupational exposures and amytrophic lateral sclerosis. A population-based case-control study. *Am J Epidemiol* 1997;**145**:1076–88.
90 **Siemiatycki J**, Fritschi L, Nadon L, *et al.* Reliability of an expert rating procedure for retrospective assessment of occupational exposures in community-based case-control studies. *Am J Ind Med* 1997;**31**:280–6.
91 **Rybicki BA**, Peterson EL, Johnson CC, *et al.* Intra-and inter-rater agreement in the assessment of occupational exposure to metals. *Int J Epidemiol* 1998;**27**:269–73.
92 **Cherrie JW**, Schneider T. Validation of a new method for structured subjective assessment of past concentrations. *Ann Occup Hyg* 1999;**43**:235–45.
93 **Stewart PA**, Carel R, Schairer C, *et al.* Comparison of industrial hygienists' exposure evaluations for an epidemiological study. *Scand J Work Environ Health* 2000;**26**:44–51.
94 **Dosemeci M**, Stewart PA. Recommendations for reducing the effects of misclassification on relative risk estimates. *Occup Hyg* 1996;**3**:169–76.
95 **Gérin M**, Siemiatycki J, Kemper H, *et al.* Obtaining occupational exposure histories in epidemiologic case-control studies. *J Occup Med* 1985;**27**:420–6.
96 **Stewart PA**, Stewart WF, Siemiatycki J, *et al.* Questionnaires for collecting detailed occupational information for community-based case-control studies. *Am Ind Hyg Assoc J* 1998;**58**:39–44.
97 **Stewart PA**, Stewart WF. Occupational case-control studies: II. Recommendations for exposure assessment. *Am J Ind Med* 1994;**26**:313–26.
98 **Ashmore JP**, Krewski D, Zielinski JM, *et al.* First analysis of mortality and occupational radiation exposure based on the National Dose Registry of Canada. *Am J Epidemiol* 1998;**148**:564–74.
99 **Watts WF**, Parker DR. Mine inspection data analysis system. *Appl Occup Environ Hyg* 1995;**10**:323–30.
100 **Stamm R**. MEGA-database: one million data since 1972. *Appl Occup Environ Hyg* 2001;**16**:159–63.
101 **Vinzents P**, Carton B, Fjeldstad P, *et al.* Comparison of exposure measurements stored in European databases on occupational air pollutants and definition of core information. *Appl Occup Environ Hyg* 1995;**10**:351–4.
102 **Stewart PA**, Rice C. A source of exposure data for occupational epidemiology studies. *Appl Occup Environ Hyg* 1990;**5**:359–63.
103 **Nelson DI**, Nelson RY, Hart KJ, *et al.* Hydrocarbon exposure assessment methodology for an epidemiologic study of renal disease. *Appl Occup Environ Hyg* 1995;**10**:299–310.
104 **Fjeldstad PE**, T Woldbaek. A national exposure database. In: Brown RH, Curtis M, Saunders KM, *et al*, eds. *Clean air at work: new trends in assessment and measurement for the 1990s.* Melksham, Wiltshire: Redwood Press, 1992:303–10.
105 **Burns KD**, Beaumont PL. The HSE national exposure database—(NEDB). *Ann Occup Hyg* 1989;**33**:1–14.
106 **Vincent R**, Jeandel B. COLCHIC—Occupational exposure to chemical agents database: current content and development perspectives. *Appl Occup Environ Hyg* 2001;**16**:115–21.
107 **Leighton TM**, Nielsen AP. The United States Environmental Protection Agency, Health Canada, and National Agricultural Chemicals Association pesticide handlers exposure database. *Appl Occup Environ Hyg* 1995;**10**:270–3.
108 **Lippmann M**, Gomez MR, Rawls GM, *et al.* Data elements for occupational exposure databases: guidelines and recommendations for airborne hazards and noise. *Appl Occup Environ Hyg* 1996;**11**:1294–311.
109 **Rajan B**, Alesbury R, Carton B, *et al.* European proposal for core information for the storage and exchange of workplace exposure measurements on chemical agents. *Appl Occup Environ Hyg* 1997;**12**:31–9.
110 **Morgan DA**. Occupational exposure databases and their application for the next millennium: symposium framework and workshop introduction. *Appl Occup Environ Hyg* 2001;**16**:111–14.
111 **Burstyn I**, Kromhout H, Kauppinen T, *et al.* Statistical modelling of the determinants of historical exposure to bitumen and polycyclic aromatic hydrocarbons among paving workers. *Ann Occup Hyg* 2000;**44**:43–56.
112 **Burstyn I**, Teschke K. Studying the determinants of exposure: a review of methods. *Am Ind Hyg Assoc J* 1999;**60**:57–72.
113 **Preller L**, Kromhout H, Heederik D, *et al.* Modeling long-term average exposure in occupational exposure-response analysis. *Scand J Work Environ Health* 1995;**21**:504–12.
114 **Baris D**, Kwak LW, Rothman N, *et al.* Blood levels of organochlorines before and after chemotherapy among non-Hodgkin's lymphoma patients. *Cancer Epidemiol Biomarkers Prev* 2000;**9**:193–7.
115 **Floderus B**, Persson T, Stenlund C, *et al.* Occupational exposure to electromagnetic fields in relation to leukemia and brain tumors: a case-control study in Sweden. *Cancer Causes Control* 1993;**4**:465–76.
116 **Veulemans H**, Steeno O, Masschelein R, *et al.* Exposure to ethylene glycol ethers and spermatogenic disorders in man: a case-control study. *Br J Ind Med* 1993;**50**:71–8.
117 **Caldwell GG**, Cannon SB, Pratt CB, *et al.* Serum pesticide levels in patients with childhood colorectal carcinoma. *Cancer* 1981;**48**:774–8.
118 **Scheele J**, Teufel M, Niessen KH. Chlorinated hydrocarbons in the bone marrow of children: studies on their association with leukemia. *Eur J Pediatr* 1992;**151**:802–5.
119 **Scheele J**, Teufel M, Niessen KH Chlorinated hydrocarbons in bone marrow of healthy individuals and leukemia patients. *Arch Environ Health* 1996;**51**:220–5.
120 **Kromhout H**, Loomis DP, Mihlan GJ, *et al.* Assessment and grouping of occupational magnetic field exposure in five electric utility companies. *Scand J Work Environ Health* 1995;**21**:43–50.
121 **Armstrong BG**. The effects of measurement errors on relative risk regressions. *Am J Epidemiol* 1990;**132**:1176–84.
122 **Kromhout H**, Heederik D. Occupational epidemiology in the rubber industry: implications of exposure variability. *Am J Ind Med* 1995;**27**:171–85.
123 **Seixas NP**, Sheppard L. Maximizing accuracy and precision using individual and group exposure assessments. *Scand J Work Environ Health* 1996;**22**:94–101.
124 **Tielemans E**, Kupper L, Kromhout H, *et al.* Individual-based and group-based exposure assessment: some equations to evaluate different strategies. *Ann Occup Hyg* 1998;**42**:115–19.
125 **Le Moual N**, Bakke P, Orlowski E, *et al.* Performance of population specific job exposure matrices (JEMs): European collaborative analyses on occupational risk factors for chronic obstructive pulmonary disease with job exposure matrices (ECOJEM). *Occup Environ Med* 2000;**57**:126–32.
126 **Hoppin JA**, Tolbert PE, Flagg EW, *et al.* Use of a life events calendar approach to elicit occupational history from farmers. *Am J Ind Med* 1998;**34**:470–6.
127 **Rappaport SM**. Smoothing of exposure variability at the receptor: implications for health standards. *Ann Occup Hyg* 1985;**29**:201–14.
128 **Werner MA**, Attfield MD. Effect of different grouping strategies in developing estimates of personal exposures: specificity versus precision. *Appl Occup Environ Hyg* 2000;**15**:21–5.
129 **Landis JR**, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74.

COMMENTARY

In their very comprehensive review on methods for assessment of occupational exposure in case–control studies, Teschke *et al* state that "among the exposure estimation methods in common use today, expert assessment is usually the best approach". They do so, despite the fact that it is well known that subjective assessments by experts is of a relative nature[1] and that in order to have a more quantitative assessment the experts have to be calibrated.[2 3] The main reason for choosing experts can be traced back to the alternative methods of self reported exposures and generic job–exposure matrices (JEM) which, as they claim, suffer from severe limitations. Recently, the limitations and possibilities of exposure assessment on the basis of JEM were extensively discussed.[4] From a somewhat broader perspective, expert assessment and JEM are not as different as often is being suggested. A study in which an expert judges the job history of every case and control, is actually applying a very detailed (job) exposure matrix where the input axis is made up by exposure determinants which the expert think of as being important. The problem with the case by case expert assessment is that the process of assigning exposure to an individual on the basis of determinants of exposure generally takes place in the black box made up by the mind and heart of an occupational hygienist or exposure assessor (in the best case). Teschke *et al* show that recently results of determinants of exposure studies (pointing at determinants of exposure such as physical properties of the agent, work environment, tasks, and use of control measures, including personal protective equipment) have increasingly become available to the expert and the field at large. With this in mind, I would like to propose that we use the result of such studies together with the hidden treasures in the mind and hearts of experts to elaborate deterministic exposure models. These models can subsequently be used to assign exposure to individual subjects on the basis of information collected on a priori identified determinants of exposure in standardised interviews (of next of kin) or questionnaires.[5] In other words, experts should be used collectively to devise these deterministic–exposure models (DEM). The models will combine the specificity of experts and the structured approach of the JEM. Exposure assessment for case–control studies in this way will become more reproducible and reliable and less prone to biases and the resulting harsh critiques it is often (justifiably) exposed to.[6]

With occupational risk assessment becoming more quantitative, it is conceivable that case–control studies (in the general population) will become less popular. The main reason for this is that the retrospective nature and resulting limitations of the exposure assessment will at best produce semiquantitative estimates of past exposures. However, case–control studies on short term health effects, such as reproductive effects,[7 8] as discussed by Teschke *et al*, point into a new direction. Banking of biological material in large community based studies (for instance, the European Community Respiratory Health Survey)[9] together with adequate collection of deterministic information will enable the future exposure assessor to produce more quantitative estimates of (internal) exposure. In addition, much needed expert calibration studies have been shown to be possible with the introduction of simple sampling methods based on passive monitoring.[7] Self assessment of occupational exposure[10] and a more rigorous use of experts as described above are needed in order to have a future for community based occupational case–control studies. Nevertheless, everyone considering such a study should not go along that way without consulting the insightful review of exposure assessment methods by Teschke and her colleagues.

H Kromhout

Environmental and Occupational Health Division, Institute for Risk Assessment Sciences, University of Utrecht, Yalelaan 2, Utrecht, Netherlands; H.Kromhout@iras.uu.nl

REFERENCES

1 **Kromhout H**, Oostendorp Y, Heederik D, *et al*. Agreement between qualitative exposure estimates and quantitative exposure measurements. *Am J Ind Med* 1987;**12**:551–62.
2 **Hawkins NC**, Evans JS. Subjective estimation of toluene exposures: a calibration study of industrial hygienists. *Appl Ind Hyg* 1989;**4**:61–8.
3 **Post WK**, Heederik D, Kromhout H, *et al*. Occupational exposures estimated by a population specific job exposure matrix and 25 year incidence rate of chronic nonspecific lung disease (CNSLD): the Zutphen Study. *Eur Respir J* 1994;**7**:1048–55.
4 **Kromhout H**, Vermeulen R. Application of job-exposure matrices in studies of the general population: some clues to their performance. *Eur Respir Rev* 2001;**11**:80–90.
5 **Stewart PA**, Stewart WF, Siemiatycki J, *et al*. Questionnaires for collecting detailed occupational information for community-based case control studies. *Am Ind Hyg Assoc J* 1998;**59**:39–44.
6 **Cherrie JW**, Kromhout H, Semple S. The importance of reliable exposure estimates in deciding whether trichloroethylene can cause kidney cancer. *J Cancer Res Clin Oncol* 2001;**127**:400–4.
7 **Tielemans E**, Heederik D, Burdorf A, *et al*. Assessment of occupational exposures in a general population: comparison of different methods. *Occup Environ Med* 1999;**56**:145–51.
8 **Veulemans H**, Steeno O, Masschelein R, *et al*. Exposure to ethylene glycol ethers and spermatogenic disorders in man: a case-control study. *Br J Ind Med* 1993;**50**:71–8.
9 **Burney PG**, Luczynska C, Chinn S, *et al*. The European Community Respiratory Health Survey. *Eur Respir J* 1994;**7**:954–60.
10 **Liljelind IE**, Stromback AE, Jarvholm BG, *et al*. Self-assessment of exposure—a pilot study of assessment of exposure to benzene in tank truck drivers. *Appl Occup Environ Hyg* 2000;**15**:195–202.