

A Novel Approach to Data Collection in a Case-Control Study of Cancer and Occupational Exposures

PATRICIA A STEWART,* WALTER F STEWART,** ELLEN F HEINEMAN,* MUSTAFA DOSEMECI,* MARTHA LINET* AND PETER D INSKIP*†

Stewart P A (National Cancer Institute, Occupational Studies Section, 6130 Executive Blvd MSC 7364, Bethesda, MD 20892-7364, USA), Stewart W F, Heineman E F, Dosemeci M, Linet M and Inskip P D. A novel approach to data collection in a case-control study of cancer and occupational exposures. *International Journal of Epidemiology* 1996; 25: 744-752.

Background. In community and hospital-based case-control studies, the occupational data collected in interviews are usually limited to responses to general questions asked of all study subjects. A procedure is described in which more detailed information can be collected in an efficient, standardized and systematic way.

Methods. A generic work history is initially collected from all subjects using a computer-assisted interview. The work history includes job title, type of business, job activities, materials and chemicals, and tools and equipment used. After responses are entered into the computer by the interviewer, the computer searches a synonym file to identify possible job-specific modules relevant to the reported job. The modules are detailed questionnaires that address specific jobs administered after obtaining the generic work history. The modules are used to ask questions about the work environment; sources of exposure; factors affecting the movement of the agent from the source to the subject, such as local exhaust ventilation; and individual and job characteristics. After the interview is completed, the work history and responses to the modules are sent electronically to an industrial hygienist who reviews the information using a custom-designed software package. Where ambiguities or contradictions occur in information reported by the respondent, or for jobs for which no module had been developed, the industrial hygienist generates up to 10 additional questions per job. These questions are sent back to the interviewer for administration of a short, second interview.

Conclusions. These procedures, which are being successfully implemented in an on-going case-control study of brain tumours, should improve disease risk estimates over those derived from more traditional approaches to exposure assessment.

Keywords: case-control studies, exposure assessment, questionnaires, job histories, occupational risk

In community or hospital-based case-control studies, information obtained on work histories historically has often been generic and limited to job title, type of business and dates of employment. These data have been used to calculate odds of the disease for specific jobs compared to all other jobs. Associations of disease with exposures are inferred from knowledge of the jobs in which excesses are observed. Because evaluation of exposure associations is done after the analysis, inference regarding a particular exposure is speculative. Since the early 1980s, a different approach has been used where all jobs having a similar exposure have been

grouped before calculating the disease odds ratios. This approach often increases the confidence in the association with an exposure, but it ignores the variability of exposures among subjects reporting the same job.^{1,2} As an example, in this approach all mechanics would be assigned the same exposures, even though automobile mechanics may work solely on brakes (asbestos exposure), or mufflers (welding fumes), and others may perform a variety of tasks, including these two.

In some instances, the generic work history is supplemented with generic questions on work activities and materials used. Although this approach may decrease misclassification to some extent, important details about exposures are not collected in a systematic fashion. Instead, the respondent is usually asked an open-ended, general question (e.g. What materials did you use on this job?). The information reported is likely to be meaningful to the respondent but not necessarily relevant to

* National Cancer Institute, Epidemiology and Biostatistics Program, Rockville, MD 20892, USA.

** Department of Epidemiology, Johns Hopkins University, Baltimore, MD, USA.

† Currently at Department of Veterinary, Anatomy and Public Health, Texas A & M University, USA.

assessing exposures. Moreover, because the questions are unstructured, the interviewer and the respondent have considerable latitude. As such, variation in the details (work processes, tasks and equipment used, etc.) reported among subjects who hold the same occupation can be considerable even though they have the same exposures.^{1,3} This can cause differential misclassification between cases and controls if the two groups differ in the accuracy and completeness of the relevant details.

To facilitate evaluation of the variability of exposures within jobs and to reduce differential reporting, detailed information on jobs must be obtained systematically. Gerin and Siemietycki⁴ used structured questionnaires to obtain such detailed information on specific exposures in selected occupations, allowing these investigators to assess exposures specific to the individual. As a result, misclassification was diminished and statistical power enhanced compared to the approach where exposure groups were developed that were comprised of all subjects with the same job.³ Stewart and Stewart^{1,2} recommended several modifications to the methods of Gerin and Siemietycki⁴ to facilitate data collection, reduce cost and decrease the amount of interviewer training. Among the strategies proposed was the use of a computer-assisted interview equipped with job-specific modules (questionnaires that specifically address exposure details of particular jobs), and an automated search mechanism to identify possible matches of reported jobs to the modules. The use of the computer allows immediate electronic transfer of the work history and module responses to an industrial hygienist, who develops additional questions for a second, short interview to clarify ambiguous or contradictory module information and obtain detailed exposure information on jobs not covered by a module. In this paper, we describe the recent implementation of these procedures in a case-control study of brain tumours.

BACKGROUND

Study subjects are 800 patients newly diagnosed with primary glioma, meningioma or acoustic neuroma and 800 controls, admitted to one of three hospitals in Boston, MA, Phoenix, AZ and Pittsburgh, PA. Controls include patients with non-malignant neurosurgical, neurologic or general surgical conditions. In-person interviews are conducted in the hospital by trained interviewers within 8 weeks of first diagnosis for the cases or admission for the controls.

The study, initiated in response to public concern about a possible link between cellular telephone use

and risk of brain tumours, is investigating a variety of other possible risk factors, including occupational exposures, diet and vitamin supplements, home use of electrical appliances, personal and family history of selected diseases, medical exposure to non-occupational ionizing radiation and use of hair dyes. Occupational exposures being investigated include electromagnetic fields; organic solvents; heavy metals; polycyclic aromatic hydrocarbons; polychlorinated biphenyls; ionizing radiation and several other chemical and physical agents.

THE COMPUTER ASSISTED PERSONAL INTERVIEW [CAPI]

Information is obtained on possible occupational and non-occupational risk factors through the use of a computer-assisted personal interview [CAPI]. Participants also complete a self-administered questionnaire for some non-occupational factors. The occupational component of the CAPI consists of a generic work history and job-specific modules, and only this part is described here.

Generic Work History Data Collection and Linkage to the Job-Specific Modules

A work history is obtained for each job held at least 6 months after age 16. Job title, employer name, and start and stop dates are collected for each job. This process provides a brief overview of the subject's work history and a frame of reference for more detailed questioning. The outline is reviewed by the interviewer to identify possible redundancies, gaps, and inconsistencies in the work history (e.g. two jobs held during the same time). In addition, the interviewer combines into a single job two or more individual jobs reported by the respondent, such as a junior and senior programmer, if the employer and tasks are the same.

After completing this review, the interviewer returns to the first job. The computer searches for appropriate job-specific modules via a synonym file that links job titles and modules through any number of synonyms identified with the modules (Table 1). Words, not strings of letters or groups of words, are searched. For example, if 'repairs equipment' was entered as the job title, the following modules would be identified for selection by the interviewer: vehicle mechanic, aircraft mechanic, and industrial machinery mechanic, because of the word 'repair'. If a job is selected by the interviewer, the module questions from that module are immediately displayed and no additional generic information is obtained on that job. If no module match is identified, the interviewer asks additional generic

TABLE 1 Examples of synonyms associated with modules^a

Module	Synonyms
Bus driver	bus, children, drive, drove, school, streetcar
Taxicab/limo driver	cab, chauffeur, chaufferer, ^b drive, driver, drove, motor, pool, taxi, van, vehicle
Truck driver	deliver, distance, drive, driver, drove, haul, hostler, route, street, tractor, trailer, truck, van, vehicle
Vehicle mechanic	auto, body, bus, cab, car, fix, handyman, maintenance, motor, mechanic, railroad, repair, taxi, vehicle
Industrial machinery mechanic	fix, leaks, machine, maintenance, motors, pipes, pumps, repair
Aircraft mechanic	aircraft, airplane, electric, fibreglass, helicopter, install, jet, line, maintenance, plane, repair, service, troubleshoot

^a The computer software program was designed to search for entire words, not strings of letters.

^b Misspellings are included.

questions about the type of business: activities performed on the job; use of chemicals and materials (e.g. wood); and use of tools and equipment. Each of these questions triggers a new search for modules. If no module has been selected after the search on reported tools and equipment, data collection for that job is complete and the interviewer repeats the process for the next job in the work history.

The synonym file contains more than 4000 words identified from work histories reported in three previous case-control studies at the National Cancer Institute [NCI].⁵⁻⁷ To this list has been added other likely responses and misspellings. Thus, the same type of job can be reported in a variety of different ways and still be linked to the same module. The file is updated regularly as new words for describing jobs are reported by respondents.

Based on the first 127 interviews (121 subjects and six proxies), 683 jobs were reported for an average of 5.4 jobs per subject. This figure is similar to that found in other case-control studies.⁷ About 80% of the jobs received a module; 46% of these were linked to a module based on the job title. The generic question about the type of business was therefore asked on only 54% of the jobs that received a module. The remaining generic questions were asked even less frequently: type of activities was asked on 50% of the jobs receiving a module; chemicals and materials on 41%; and tools and equipment on 38%.

Rationale for Job-Specific Modules

Modules were developed because of the limitations of generic work histories. In response to open-ended generic questions (e.g. What chemicals did you use?) the respondent may provide vague or ambiguous information that results in assessments that do not reflect the actual variability of exposures across individuals reporting the same job.⁸ This type of question may also be more prone to differential reporting by cases and

controls than closed-ended questions (e.g. Were you exposed to benzene?). Additional probing by the interviewer may not necessarily clarify industrial hygiene concerns for several reasons. Interviewers in case-control studies usually lack experience or specialized training in the field of industrial hygiene. Thus, they may not be able to assess adequately the work environment to ask the appropriate questions. Probing may also vary by interviewer. Modules were also developed to collect the information needed by the industrial hygienist for exposure assessment, because the industrial hygienist does not generally review the work history data until after data collection has been completed. This late date in the study makes it difficult or impractical to resolve ambiguity or vagueness in the reported information, because study subjects may have died, become too ill to be interviewed or moved, becoming difficult to locate.¹

Modules were only developed for selected jobs for a number of reasons. First, the modules are optimal for jobs where the type or level of exposure varies among individuals having the same job. Modules were considered unnecessary for jobs that do not have exposures to the agents of interest in this study and not essential for jobs that were likely to be infrequently reported. To determine which modules should be developed, more than 50 000 jobs from the three NCI case-control studies⁵⁻⁷ were ranked by frequency of occurrence.

Modules were developed for the 38 most frequently reported jobs with potential exposure to the agents of interest in the brain tumour study, and these 38 jobs accounted for more than 40% of the jobs with exposures of interest in those studies. Modules were also developed for jobs that were expected to occur at a higher frequency than in the NCI studies because of the geographical location of the participating hospitals in the present study. For example, modules were developed for fishermen, shoemakers, and knitters/weavers (Boston); miners and the steel industry

TABLE 2 *Job specific module organization*

Industrial hygiene principles	Examples of questions	Knowledge Gained from Questions		
		Type of exposure	Exposure level	Rationale
Environment	Type of construction	X (Type of paint)	-	Type of paint varies by structure painted
	Number of customers	-	X	More customers may imply greater indirect exposure
Sources of exposure	Welding	X (Metal fumes)		Metal fumes from welding
	Soldering	X (Lead, nickel, silver)		Lead, nickel, silver from solder fumes
	Painting	X (Solvents, metals)		Solvents, metals from paint
Transport of exposure	Local ventilation exhaust	-	X	Exhaust likely to reduce level
	Indoor/outdoor operation	-	X	Level may be lower outdoors than indoors
Individual/job	Hours/work, weeks/year welded		X	Longer duration and greater frequency indicates greater exposure
	Distance from electrical equipment	-	X	Less distance implies greater exposure
	Skin wet with cutting oils	-	X	Dermal contact may increase exposure
	How pesticides were applied	-	X	Work practices can identify exposures and levels

(Pittsburgh); and farmers/ranchers and farm workers (Phoenix). Nineteen modules were developed specifically to address occupational exposure to electromagnetic fields.

Organization of the Job-Specific Modules

Each module has a similar organization and format that is consistent with cognitive function and industrial hygiene principles, and is designed to facilitate exposure assessment (Table 2). First, questions are asked about the type of business and the general work environment. For example, painters are asked about the type of building or construction work they painted (e.g. homes, commercial buildings, industrial buildings, roads, tunnels or bridges) or type of manufacturing operation. Hairdressers are asked about the number of customers the beauty salon served per week and laundrers are asked about the specific type of work setting (e.g. laundry only, dry cleaning facility, hotel, etc.).

Next, questions are asked about specific tasks possibly performed in the job to identify the types and sources of exposure and to qualitatively and quantitatively evaluate exposures. From the source, the quantity of a substance being used or produced, production variables (e.g. temperature and pressure) and the mechanism by which exposures are released into the environment (e.g. shaping, grinding or mechanical) can be inferred for evaluating the level of exposure. For example, an aircraft mechanic is asked about the hours per week and weeks per year spent welding, soldering,

painting or varnishing, lubricating machinery, cleaning metal parts and machining metal parts. Responses to these questions identify specific types of exposures (e.g. metal fumes, solvents, etc.) and information on the level of exposure (from the type of task and its duration and frequency).

More detailed questions about the tasks are then asked to add further information on the level of exposure, including factors that affect the transport or movement of the agent from the source to the individual, such as the presence of exhaust ventilation, and factors on the job and individual level. The latter two factors include questions on duration and frequency of exposure; the distance from the source; use of chemicals and materials; occurrence of dermal exposure; work practices; and use of personal protective equipment. Because these more detailed questions may be more difficult for the respondent to recall, they are generally asked only if a specified amount of time spent performing a task is exceeded. If the respondent is a proxy, fewer and more general questions are asked, because proxies are less likely to be able to report detailed workplace information.⁹⁻¹³

Administration of the occupational component of the questionnaire has taken an average of 46 minutes per subject. In the first 127 interviews, modules were administered for 77% of the jobs reported (Table 3), an average of four modules per subject. The modules administered most frequently have been for office professional, manager, salesperson, teacher, nurse,

TABLE 3 Administration of job specific modules and supplemental questions

Job modules ^a	Job modules			Supplemental questions ^b	
	Number of respondents administered at least one module	Number administered	Mean number of time modules asked	Number of times asked	Mean number of questions
Office professional	41	81	2.0	57	1.2
Manager	42	77	1.8	17	2.3
Salesperson ^a	32	54	1.7	12	2.0
Teacher	17	33	1.9	4	2.7
Nurse	12	28	2.3	7	1.1
Computer user ^a	19	27	1.4	6	2.0
Waitress	14	19	1.4	4	1.3
Kitchen worker	14	17	1.2	0	-
Extremely low frequency job ^{a,c}	14	16	1.1	14	4.3
Labourer	11	14	1.3	13	4.0
Electrical machinery & VDT operator ^a	9	13	1.4	11	3.4
Janitor	11	11	1.0	4	1.0
Military ^a	11	11	1.0	7	2.0
Painter	9	9	1.0	4	3.0
Machinist	7	8	1.1	7	3.4
Gas station attendant	7	7	1.0	1	2.0
Butcher/meat cutter	4	6	1.5	0	-
Electrician/repairer of electrical equipment	4	6	1.5	3	4.3
Police, detective	6	6	1.0	2	2.0
Traffic, shipping & receiving clerk	6	6	1.0	3	1.3
Fork lift operator ^a	4	5	1.2	3	3.0
Physician/health professional ^a	3	5	1.7	2	2.0
Production inspector, checker or examiner	5	5	1.0	3	-
Shoemaker/repairer	2	5	2.5	1	1.0
Truck driver	5	5	1.0	1	4.0
Vehicle mechanic	5	5	1.0	4	3.3
Welder	4	5	1.2	4	4.7
Computer operator ^a	4	4	1.0	1	2.0
Farm worker	3	3	1.0	1	6.0
Packaging or filling operator	3	3	1.0	1	2.0
Radio operator ^a	3	3	1.0	1	3.5
Radio, VDT operator ^a	3	3	1.0	0	-
Aircraft mechanic	2	2	1.0	2	2.0
Barber/hairdresser	2	2	1.0	0	-
Bartender	2	2	1.0	0	-
Concrete & terrazzo worker	2	2	2.0	2	5.0
Dry cleaner	2	2	1.0	1	4.0
Electronic/electrical engineer technician ^a	2	2	1.0	2	6.0
Gardener, groundskeeper	2	2	1.0	1	1.0
Mail carrier	2	2	1.0	1	1.0
Steel industry	2	2	1.0	1	10.0
Cabinet maker/bench carpenter	1	1	1.0	1	4.0
Farmer	1	1	1.0	0	-
Fire fighter	1	1	1.0	1	2.0

TABLE 3 (continued)

Job modules ^a	Job modules			Supplemental questions ^b	
	Number of respondents administered at least one module	Number administered	Mean number of time modules asked	Number of times asked	Mean number of questions
Miner ^a	1	1	1.0	1	2.0
Plumber/pipefitter/steamfitter	1	1	1.0	0	-
Radar operator ^a	1	1	1.0	0	-
Sheet metal worker	1	1	1.0	1	6.0
Telephone line installer ^a	1	1	1.0	1	6.0
Railroad operator ^a	1	1	1.0	1	3.0
Taxi/limo driver	1	1	1.0	0	-
Antenna repair ^a	0	0	0	0	-
Bus driver	0	0	0	0	-
Brick, block & stone mason	0	0	0	0	-
Carpenter	0	0	0	0	-
Dentist/dental hygienist ^a	0	0	0	0	-
Electrical power line installer ^a	0	0	0	0	-
Fisherman	0	0	0	0	-
Heating equipment operator ^a	0	0	0	0	-
Industrial machinery mechanic	0	0	0	0	-
Insulator	0	0	0	0	-
Knitter/weaver	0	0	0	0	-
Roofer	0	0	0	0	-
Tool & die worker	0	0	6	0	-
Total	361	528			
No module		94		75	3.5

^a Indicates that these modules only include questions on electromagnetic fields.

^b Excludes questions asked due to computer or operator error.

^c Examples include assemblers, crane and hoist operators and typesetters.

computer user, waitress and kitchen worker, constituting almost 50% of the modules administered (primarily addressing EMF exposures). Often, the same module is administered several times to the same individual for different employers. For example, nurses in this study have tended to change employers frequently but continued to work as nurses. They therefore have received the nurse module several times. The five most frequently administered modules have, on average, been administered twice per subject.

INDUSTRIAL HYGIENE REVIEW AND SUPPLEMENTAL INTERVIEW

In most occupational case-control studies the industrial hygienist typically reviews the exposure information

long after data collection has been completed. As a result, the industrial hygienist has no means of resolving questions about exposures if the reported information is vague or incomplete.¹ To minimize this problem, a procedure called SCORE [Self-Corrected Occupational Report] was developed in the present study to allow review of the work histories and modules by an industrial hygienist within 2 weeks of the interview.¹⁴

In the SCORE procedure, the work histories and responses to the modules from the CAPI interview are sent electronically to the study industrial hygienist for review. The industrial hygienist is blinded to the case/control status of the study subject. A software package was developed for this study that reformats and displays the work history information and responses to module questions to facilitate review.

TABLE 4 Examples of jobs for which no module was available

Activity director	Closer funder (wrote up mortgages)
Take care of children	Housekeeper
Meter reader	Wire bonder ^a
Pianist	Played music
Selector (glass) ^a	Jockey
Rode horses	Proof reader
Reconciler	Massage therapist
Clean house	Made coke machines ^a
Made ads	Go to meetings
Building sewer system ^a	Set up equipment
Read scripts, organized trips	Production assistant for TV
On the air performer	Reporter on the air
Press secretary	Making chlorine gas ^a
Gift wrapper	Shopper
Page in library	Lifeguard
Singing telegrams	Carpet installer ^a
Librarian assistant	Silk winder ^a
Crossing guard	Stock boy (grocery store)
Parking attendant	

^a Jobs for which questions were asked in the supplemental interview.

SCORE is used by the industrial hygienist to ask up to 10 additional questions per job. The number of questions asked is limited to minimize the duration of the follow-up interview. Questions can be retrieved from an existing electronic library or developed *de novo*. Questions in the library are indexed and can be accessed by job, type of business, exposure agent, or by industrial hygiene assessment characteristics (general environment, source of exposure, transport of the agent to the individual, and individual/job characteristics).² For example, the question 'How many times a year did you work on brakes?' is indexed as a question for truck drivers (among others), the trucking industries, asbestos and source and frequency of exposure. Once retrieved, the question can either be used as is or modified. If there is no appropriate question in the library, a new question can be developed and added to the library under appropriate indexing variables assigned by the industrial hygienist.

SCORE is used to clarify information reported in the modules. The modules for which SCORE questions have been most frequently asked thus far include jobs identified as having extremely low EMF exposures (average = 4.3 additional questions), labourer (4.0), vehicle mechanics (3.3), machinists (3.4) and workers using electrical machinery and video display units (3.4) (Table 3). SCORE is particularly useful, however, for jobs for which there is no module, such as a 'glass selector', a 'wire bonder', 'building sewer system', and 'making chlorine gas' (Table 4). Jobs such as these

receive questions that probe for specific exposure information. Many other jobs for which no module had been developed (e.g. jockey, housekeeper, pianist) (Table 4) did not require supplemental questions ($n = 61$). A total of 326 supplemental questions have been asked on 94 jobs that had no module.

Coding of Occupation and Industry

When a job-specific module is used, the computer automatically assigns the appropriate Standard Occupational Classification¹⁵ [SOC] code to the reported job from the precoded modules. For jobs that did not have a module, the industrial hygienist can assign the SOC and, for all jobs, the industry Standard Industrial Classification¹⁶ [SIC] codes, or reassign codes, while reviewing the work history. Two different procedures can be used to code the industry. First, business information (employers' name, address, type of business [i.e. SIC code] and number of employees) was incorporated into the software from an existing economic data base.¹⁷ The industrial hygienist can search this data base on employer name, or city and state, to assign the appropriate SIC code. Alternatively, the industrial hygienist can assign job and industry codes using a library of SIC and SOC codes incorporated in the SCORE software package.

DISCUSSION

We have developed and successfully implemented a new approach to facilitate the collection of detailed occupational exposure information in a hospital-based case-control study. The approach uses a CAPI, job-specific modules, timely review of the work history and module data by an industrial hygienist, and a supplemental interview developed by the industrial hygienist to resolve ambiguities and to obtain additional information.

The CAPI enables the interviewer to obtain detailed industrial hygiene information about reported jobs in an efficient, standardized and systematic fashion. First, automating the module selection procedure with the synonym file reduces the time required by the interviewer to identify the correct module. This is done by using a search procedure that checks a synonym file as soon as the data are entered and by having possible matching modules identified on the screen. Second, the computerized identification of modules minimizes the variability of the module selection process and the use of the modules minimizes the variability associated with probing that might occur due to differences in the interviewers' experience or knowledge of jobs or the subject's disease status. Third, many of the modules are

highly detailed, covering numerous potential exposures and tasks. For example, the carpenter module has 41 questions (What per cent of your time did you ...) that contain 105 sub-questions (... saw wood; paint or varnish; etc.). The CAPI allows sophisticated branching and skip patterns (22 in the carpenter module) so that the interviewer asks only the most important questions for each particular study subject. Using a CAPI also minimizes inadvertent failure to ask all questions that could occur with a paper questionnaire. Fourth, the CAPI eliminates the need for coding of responses and data entry at a later date. This allows for timely review by the industrial hygienist and enhances opportunity for corrective action through the follow-up interview. Finally, it ensures that detailed exposure information is collected from cases and controls across interviewers at different field centres in a standardized fashion.

The modules were formatted to promote respondent recall. The initial overview of the work history primes the respondent's recollection of the job for the questions that follow. Questions are first asked on the work setting to create a frame of reference for the respondent. This framework facilitates the respondent's recall of tasks, which are the next level of detail and which are asked to prime the recall of the more detailed questions on the transport of the agent and the job and individual characteristics.

The modules were developed specifically for this study and consequently emphasize the exposures of special interest to brain tumour aetiology and to the geographical location of the three collaborating hospitals. Most of the modules, however, cover exposures of general interest (e.g. solvents, metals, polycyclic aromatic hydrocarbons). Because of the particular interest in EMF exposure, 17 modules were developed for jobs that have EMF exposures but may also have other exposures as well (e.g. dentists to ionizing radiation and mercury and fork lift truck operators to polycyclic aromatic hydrocarbons). Due to the limited amount of time available for questionnaire development before the study was initiated, only EMF exposures were explored in detail in these modules. Information on other exposures is collected in the supplemental interview and could be added to the modules in the future.

We were surprised to find that almost 80% of the jobs reported to date had corresponding modules, compared to the expected 40%. Initially we had developed a secretary module primarily to obtain information on video display terminals [VDT], but the use of VDT is so widespread among current office and other jobs that the module was renamed as office professional to incorporate other frequently held jobs, e.g. accountants, in which computers are used. Similarly, many managers,

sales people and other computer users now have VDT exposures where previously they would have been considered as non exposed. These four jobs contributed 45% of the modules asked and therefore probably resulted in this increased coverage of jobs.

SCORE enables the industrial hygienist to evaluate the quality of information collected in the interview during the review of the work histories and modules and to make suggestions on how to improve module selection. Of the 155 jobs that were not evaluated using a module, 53 (8% of all the jobs) should have been linked by synonyms to a corresponding module but were not. Most of these occurred in the early stages of the study and were due to computer or interviewer error. Where the latter occurred, the industrial hygienist provided feedback to the interviewer of the more appropriate module. The other reason for lack of a module was due to the lack of an appropriate synonym. When this occurred, the generic work history questions were asked and reviewed by the industrial hygienist to determine whether additional questions were necessary. Additional synonyms were added to the synonym file as these failures were identified. Forty-one other jobs had no module but were likely to have exposures of concern (e.g. wire bonder, building sewer system, making chlorine gas (from Table 4)); supplemental questions were asked about these jobs using the SCORE procedure. Sixty-one jobs did not require any additional questions of the respondent.

The proportion of subjects being asked supplemental questions, as well as the number of questions asked, should decrease as more experience is gained. For example, a question repeatedly asked of office professionals in the early stages of the study in the supplemental interview was later added to that module. The jobs for which most of the supplemental questions are asked were anticipated because the modules for these jobs did not provide sufficient information without a second interview. For example, labourers and machinists can vary considerably in their job activities and exposures. The modules were helpful, however, because they allowed the questions in the second interview to be more focused. Similarly, two modules developed for EMF exposures only (jobs with extremely low EMF exposures, and electrical machinery and VDT operators) had a large number of supplemental questions to address other exposures.

Coding of the jobs and industries was done using the SIC and SOC systems. Although these systems are frequently used in case-control studies to group jobs, they suffer from the limitation that they were developed for purposes other than exposure assessment. Thus, often there are many codes for jobs or industries that

are fairly uniform in exposures (there are 15 types of registered nurse) but few codes for jobs or industries that can have extremely varied exposures (one code for manufacturers of inorganic chemicals). This nonspecificity can reduce observed associations.¹⁸ Codes were assigned, however, because they will be used as search variables in databases in the exposure assessment phase of the study.² We strongly urge the development of a coding system for jobs and industries that is based on industrial hygiene principles.

Assignment of exposure levels for the reported job will not begin until data collection has been completed. It will follow the same industrial hygiene principles as described in the data collection phase of this study and in another report.² Standardized procedures for data collection and exposure assessment should decrease nondifferential misclassification due to inconsistency among interviewers or the exposure assessor and due to assessor fatigue. Differential misclassification of exposure between cases and controls could occur if respondents identify exposures differentially, but this type of bias is probably less likely for the other types of questions (e.g. tasks) that are asked. Implementation of the procedures described in this paper should, therefore, improve both the accuracy and the reliability of the assessment, since there should be less misclassification than when only generic information is obtained. This, in turn, will provide better estimates of exposure-disease associations.

ACKNOWLEDGEMENTS

A special thank you must be given to a number of people who have contributed to this project. Jack Siemiatycki, of the Institut Armand-Frappier (SRA contract #B112-01) provided support in the development of the questionnaires. Robert Saal, Sandra Daughton, Kathy Chimes, and Diane Fuchs of Westat, Inc. provided assistance in formatting the questionnaires. Douglas Watson and Timothy Wilcosky of RTI, have served as study managers for the brain tumour study. Janet Whelan, Renee Karlsen and Patricia Yost of RTI designed the computerized software program. Helen Triolo of SRA provided computer programming support for the SCORE program.

REFERENCES

- ¹ Stewart W F, Stewart P A. Occupational case-control studies: I. Collecting information on work histories and work-related exposures. *Am J Ind Med* 1994; **26**: 297-312.
- ² Stewart P A, Stewart W F. Occupational case-control studies: II. Recommendations for exposure assessment. *Am J Ind Med* 1994; **26**: 313-26.
- ³ Dewar R, Siemiatycki J, Gerin M. Loss of statistical power associated with the use of a job-exposure matrix in occupational case-control studies. *Appl Occup Environ Hyg* 1991; **6**: 508-15.
- ⁴ Gerin M, Siemiatycki J. Obtaining occupational exposure histories in epidemiologic case-control studies. *J Occup Med* 1985; **27**: 420-46.
- ⁵ Thomas T L, Stewart P A, Steinhagen A *et al*. Risk of astrocytic brain tumors associated with occupational chemical exposures: A case-reference study. *Scand J Work Environ Health* 1987; **13**: 417-23.
- ⁶ Vineis P, Thomas T, Hayes R B *et al*. Proportion of lung cancers due to occupation in different areas in the U.S. *Int J Cancer* 1988; **48**: 851-56.
- ⁷ Blair A, Linos A, Stewart P A *et al*. Evaluation of risks for non-Hodgkin's lymphoma by occupation and industry exposure from a case-control study. *Am J Ind Med* 1993; **2**: 301-12.
- ⁸ Teshke K, Kennedy S M, Olshan A F. Effect of different questionnaire formats on reporting of occupational exposures. *Am J Ind Med* 1994; **26**: 327-37.
- ⁹ Coggan D, Pippard E C, Acheson E D. Accuracy of occupational histories obtained from wives. *Br J Ind Med* 1985; **42**: 563-64.
- ¹⁰ Correa-Villasenor A, Batista C, Rothman N *et al*. Job histories obtained by interview from semi-conductor manufacturing workers: A reliability study. *Am J Epidemiol* 1991; **134**: 737-38.
- ¹¹ Lerchen M L, Samet J M. An assessment of the validity of questionnaire responses provided by a surviving spouse. *Am J Epidemiol* 1986; **23**: 481-89.
- ¹² Boyle C A, Brann E A. The Selected Cancers Cooperative Study Group: Proxy respondents and the validity of occupational and other exposure data. *Am J Epidemiol* 1992; **136**: 712-21.
- ¹³ Chong J P, Turpie I, Haines T *et al*. Concordance of occupational and environmental exposure information elicited from patients with Alzheimer's Disease and surrogate respondents. *Am J Ind Med* 1989; **15**: 73-89.
- ¹⁴ Stewart W F, Stewart P A, Heineman E *et al*. A method for collecting work history information in occupational community-based case-control studies. *Occ Hyg* 1996; **3**: 145-52.
- ¹⁵ Office of Management and Budget. *Standard Occupational Classification Manual*. Washington, DC: US Government Printing Office, 1977.
- ¹⁶ Office of Management and Budget. *Standard Industrial Classification Manual*. Washington, DC: US Government Printing Office, 1972.
- ¹⁷ Dun & Bradstreet. *Dun's Electronic Yellow Pages*. Dialog Information Services Inc., Palo Alto, CA, 1988.
- ¹⁸ Dosemeci M, Cocco P, Gomez M *et al*. Effects of three features of a job-exposure matrix on risk estimates. *Epidemiology* 1994; **5**: 124-27.

(Revised version received January 1996)