

Reliability of demographic, smoking and occupational data provided by mothers vs. fathers in a childhood cancer study

Roberta McKean-Cowdin^a, Susan Preston-Martin^a, Janice M. Pogoda^b, Beth A. Mueller^c and Elizabeth A. Holly^d

^aDepartment of Preventive Medicine, University of Southern California, Norris Comprehensive Cancer Center, Los Angeles, CA, ^bStatology, Truckee, CA, ^cFred Hutchinson Cancer Research Center, Seattle, WA, and ^dDepartment of Epidemiology, University of Washington, Seattle, WA, ^dDepartment of Epidemiology and Biostatistics, University of California, San Francisco, CA, USA

Summary

Correspondence:

Roberta McKean-Cowdin,
Department of Preventive
Medicine, University of
Southern California, Norris
Comprehensive Cancer
Center, 1441 Eastlake Ave., PO
Box 33800, MS 44, Los
Angeles, CA 90033-0800, USA.
E-mail: mckeanco@hsc.usc.edu

A large case-control study of children was used to test mothers' reporting of information on fathers' background, lifestyle and occupational factors. For a subset (104) of 1341 enrolled families, both parents were interviewed about fathers' characteristics. Reliability of reporting was determined for fathers' race, education, smoking status, non-recent job history and use of occupational agents. The ability of mothers to report fathers' race, education and smoking status was high ($\kappa > 0.70$). Mothers were generally able to report jobs held by the fathers in the 5 years preceding the birth of the child, but reliability was higher for jobs held for longer (κ typically above 0.70), rather than shorter periods (κ above 0.40). The finding that mothers' reporting on fathers' background, lifestyle and non-recent job history was reliable is encouraging, because many studies on childhood health rely exclusively on information from interviews with mothers. However, mothers were not reliably able to describe exposure to specific occupational agents.

Introduction

Exposure histories of families in retrospective studies are often obtained through interview. In studies of paediatric disease, exposure histories of both biological parents may be of importance, but are not necessarily available through records or personal interviews with each parent. In this situation, mothers' responses may be the only available source of information. Previous studies have shown that proxies (typically a spouse/partner or other relative) can reliably report demographic data¹⁻³ and smoking status,¹⁻⁶ but the reliability of occupational histories obtained through proxy has been less consistent.¹⁻¹⁰ While several reports indicate that respondents can describe the most recent or current job of a spouse/partner with good to excellent agreement,^{1,2,5} others show that they do not accurately report lifetime employment histories,^{3,4} nor do they accurately report occupational exposures to specific agents, such as

solvents, arsenic and asbestos.^{3,8-10} Whether a spouse/partner can reliably report non-recent job histories remains unclear. This study investigates the reliability of mothers' reporting of fathers' demographic details, smoking status, job history and exposure to occupational agents in a study of childhood cancer.

Using data from a case-control study of childhood brain tumours, we examined the ability of biological mothers to describe the race, education and smoking status of the child's biological father. We also examined the mother's ability to report jobs held by the father during the 5 years preceding the birth of the child enrolled in the study and her ability to report on his exposure to nine general types of occupational agents. We examined how reliability of proxy job history varied by characteristics of the mother-father pair and by level of detail about job history. We also looked for evidence of reporting bias among cases and controls.

Methods

Data for this analysis were collected as part of a multicentre, case-control study to examine the relationship between childhood brain tumours and potential exposures to N-nitroso compounds through diet, employment and lifestyle factors. Cases ($n = 540$) were enrolled in the study from 1984 to 1991 from three US West Coast tumour registries located in Seattle, Los Angeles and San Francisco. Controls ($n = 801$) were selected using a two-stage random digit dial procedure.¹¹ Detailed methods have been published previously.¹²

Personal interviews were conducted with all biological mothers and with 70% of biological fathers (on their own exposure histories), either by telephone or in person. For fathers, information was requested on race, religion, education level, medical, smoking and job history. Smoking history included whether the father had smoked at least once a day, for at least 3 months preceding the birth of the child enrolled in the study. Occupational histories included industry and job titles for any jobs held for at least 1 month during the 5 years preceding the birth of the child and were coded using standard classification systems.^{13,14} Additionally, respondents identified substances to which the father was exposed in the workplace from a study list of agents (13 broad categories and 88 subcategories). A total of 58 consecutive families (control mothers and fathers) were interviewed about fathers' exposure histories at the end of the study. Throughout the study, mothers were interviewed about the father's exposure history when an interview was not yet scheduled with the father at the time of the mother's interview. For 55 of these interviews, the fathers were willing to complete the interviews when subsequently contacted by the interviewer. Forty-six of the 55 couples (13 cases and 33 controls) provided complete demographic and job histories. Therefore, separate interviews on fathers' background, lifestyle and job history were completed with both parents for 104 (7.8%) of the 1341 families interviewed.

The reliability of proxy exposure histories was tested using kappa scores. Kappa is a measure of reliability for categorical data representing expected agreement beyond chance,^{15,16} and was interpreted as 'good' for kappa > 0.60 and 'poor' for kappa < 0.40 (based on Fleiss'¹⁶ description of kappa scores of 0.4–0.75 as fair to good and Landis and Koch's¹⁷ description of kappa from 0.6 to 0.8 as substantial). Smoking status was examined as a binary variable

(ever/never), whereas race (e.g. white, Latino) and education (e.g. less than high school, high school graduate) were examined as multiple categorical variables. Estimates for employment were based on coded data broken into major (first digit, e.g. manufacturing) and minor (first and second digit, e.g. manufacturing of paper products and printing) industrial and occupational categories. The proportion of specific agreement (P_s , the conditional probability that the father will assess himself as exposed, given the mother already has) is also reported for exposure to specific occupational agents, because kappa may underestimate reliability for substances of very low prevalence.¹⁶ Family characteristics that may influence parents' ability to recall jobs occurring near the birth of the child were analysed separately (using stratified analysis) and in a multiple logistic regression. These characteristics were (1) time from birth of the child to interview; (2) age of the mother at the time of the interview; (3) number of biological children of the mother; (4) total number of reported jobs; and (5) education level of parents.

Results

The sample of 104 mother-father pairs consisted mostly of white, non-Latino individuals who were college educated. The sample was similar to the study populations (Seattle, San Francisco, Los Angeles) from which it was drawn with respect to racial/ethnic group and age of the child at interview (i.e. time to interview), but included more college graduates. Mothers reported slightly more jobs for the father on average than fathers reported for themselves (Table 1). Kappa scores for proxy-reported employment histories were good whether one or two digits of the code were used, and highest for employment of longest duration (Table 2).

Kappa scores for industry or job tasks of longest duration were consistently high regardless of the characteristics of the child's family; selected findings are shown in Table 3. In contrast, kappa for industry and job tasks of shortest duration declined as the number of children in the family increased, or the number of jobs held by the father during the 5-year study interval increased. There was no clear pattern in kappa by time since event (i.e. age of the child at interview), using 10-year (Table 3) or 5-year intervals, or by mother's age at interview (data not shown). Mothers' reporting of fathers' exposure to specific

Table 1. Demographics and job characteristics for 104 mother-father pairs, West Coast Childhood Brain Tumor Study 1984-91

Demographic factors	No.	%
Mothers' race		
White (non-Latina)	88	85
Latino	8	8
African-American	5	5
Other	3	3
Fathers' education		
Less than high school	2	2
High school graduate	3	3
Partial college/technical	18	17
College graduate	81	78
Mothers' age at time of interview (years)		
24-35	42	40
36-47	51	49
48-60	11	11
Number of biological children of mother		
1-2	59	57
3-5	41	39
≥ 6	4	4
Age of child at interview (years)		
0-5	14	14
6-10	39	38
11-15	28	27
16 +	23	22
Job characteristics		
	Mean	Standard error
Duration of longest job in years:		
Mother reported for father	5.7	0.51
Father reported for himself	4.7	0.39
Duration of shortest job in years:		
Mother reported for father	3.6	0.50
Father reported for himself	3.5	0.39
Mean number of jobs:		
Mother reported for father	2.3	0.13
Father reported for himself	1.7	0.09

occupational agents was poor (Table 4). The prevalence of reported exposure was low, and the mother was less likely to report exposure for the father than was the father himself.

Reliability of reporting for demographic data and smoking status was high. Kappa scores for race ($K=0.79$; 95% confidence interval [CI] 0.68, 0.91), education ($K=0.71$; 95% CI 0.61, 0.80) and smoking status ($K=0.84$; 95% CI 0.65, 1.0) were good.

Although the data set included a limited number of case parents for comparison, there was no evidence of differential reporting by disease status for race (kappa for cases = 0.79 [95% CI 0.48, 1.0] vs. kappa for controls = 0.79 [95% CI 0.65, 0.94]), smoking (kappa for cases = 0.84 [95% CI 0.35, 1.0] vs. kappa for controls = 0.86 [95% CI 0.66, 1.0]), industry of longest duration (kappa for cases = 0.89 [95% CI 0.57, 1.0] vs. kappa for controls = 0.85 [95% CI 0.75, 0.95]) or job title of longest duration (kappa for cases = 0.74 [95% CI 0.53, 0.94] vs. kappa for controls = 0.68 [95% CI 0.60, 0.77]). However, case mothers reported fathers' education levels (kappa for cases = 0.59 [95% CI 0.33, 0.83] vs. kappa for controls = 0.72 [95% CI 0.62, 0.82]) and employment of shortest duration (kappa for case industry = 0.39 [95% CI 0.12, 0.65] vs. kappa for control industry = 0.69 [0.59, 0.78]) with less reliability than control mothers (regardless of whether the comparison group consisted of all control pairs or only control pairs interviewed during the same time period as the cases). Confidence limits for these kappa estimates were wide. Nevertheless, disease status remained a discriminating factor for agreement on employment of shortest duration in a multiple logistic regression model including characteristics of the child's family (listed in Methods).

Table 2. Kappa for fathers' industry and job title self-report vs. mothers' response by degree of detail for 104 mother-father pairs, West Coast Childhood Brain Tumor Study 1984-91

	Industry		Job title		No. of pairs
	Kappa	[95% CI]	Kappa	[95% CI]	
Longest duration^a					
One digit of code	0.85	[0.76, 0.95]	0.70	[0.62, 0.77]	104
Two digits of code	0.84	[0.79, 0.89]	0.66	[0.62, 0.69]	104
Shortest duration^b					
One digit of code	0.65	[0.56, 0.74]	0.57	[0.49, 0.65]	104
Two digits of code	0.65	[0.59, 0.71]	0.50	[0.47, 0.54]	104

^aJob held by father for the longest duration during the 5 years preceding the birth of the study-enrolled child.

^bJob held by father for the shortest duration during the 5 years preceding the birth of the study-enrolled child.

Table 3. Select kappa findings for fathers' industry and job title self-report vs. mothers' response by family characteristics for 104 mother-father pairs, West Coast Childhood Brain Tumor Study 1984-91

	Industry		Job title		No. of pairs
	Kappa	[95% CI]	Kappa	[95% CI]	
Longest duration ^a					
All mother-father pairs	0.85	[0.76, 0.95]	0.70	[0.62, 0.77]	104
1-2 children ^b	0.83	[0.71, 0.95]	0.70	[0.61, 0.81]	59
3-11 children	0.89	[0.75, 1.0]	0.71	[0.59, 0.83]	45
1-2 jobs ^c	0.87	[0.76, 0.98]	0.71	[0.61, 0.80]	69
3-6 jobs	0.81	[0.63, 0.99]	0.68	[0.56, 0.80]	35
Age of child at interview					
0-10 years	0.78	[0.64, 0.91]	0.73	[0.62, 0.84]	53
11-15 years	0.85	[0.72, 0.98]	0.64	[0.53, 0.74]	51
Shortest duration ^d					
All mother-father pairs	0.65	[0.56, 0.74]	0.57	[0.49, 0.65]	104
1-2 children ^b	0.71	[0.59, 0.82]	0.68	[0.58, 0.79]	59
3-11 children	0.58	[0.44, 0.72]	0.40	[0.28, 0.53]	45
1-2 jobs	0.74	[0.63, 0.85]	0.62	[0.52, 0.72]	69
3-6 jobs	0.46	[0.31, 0.62]	0.47	[0.34, 0.59]	35
Age of child at interview					
0-10 years	0.71	[0.58, 0.85]	0.55	[0.44, 0.66]	53
11-15 years	0.59	[0.46, 0.72]	0.58	[0.47, 0.69]	51

^aJob held by father for the longest duration during the 5 years preceding the birth of the study-enrolled child.

^bNumber of children born to the biological mother of the study-enrolled child.

^cNumber of jobs held by the father during the 5-year interval as reported by the mother.

^dJob held by father for the shortest duration during the 5 years preceding the birth of the study-enrolled child.

Table 4. Kappa values and proportion of specific agreement (P_s) for fathers' self-reported occupational exposures vs. mothers' responses for 104 mother-father pairs, West Coast Childhood Brain Tumor Study 1984-91

Occupational exposure ^a	Kappa	P_s	Father ever exposed as reported by:	
			Father (No.)	Mother (No.)
Dust	0.51	0.66	43	24
Oil/coal products	0.42	0.42	30	16
Insulation	0.29	0.37	20	9
Paints	0.28	0.36	25	9
Plastics	0.22	0.22	9	1
Radiation	0.18	0.18	4	8
Animal/animal products	0.15	0.15	12	1
Metals/fumes	0.13	0.22	13	5
Fertiliser/pesticides	0.08	0.13	13	2

^aNine of 13 possible broad categories were reported by the 104 mother-father pairs. Remaining non-reported categories include beauty products, solvents, high electrical currents and other substances.

Discussion

Although previous studies have shown that proxy respondents do not provide reliable lifetime employment histories,^{3,4} our data indicate that mothers can reliably report select periods of fathers' past job histories. Kappa levels for our study are consistent with scores reported previously for recent or current jobs (kappa levels typically above 65%^{1,2,5}). The level of reliability for jobs of shortest duration varied by the number of children in the family and by the number of jobs reported by the mother. Because the child's birth date served as the reference point for jobs of interest to the study, it was not surprising to detect more error in recall as the reference point becomes less unique (i.e. multiple births in family). A decrease in accuracy with increasing number of reported jobs has been described previously^{18,19} and may be the result of error in reporting of short duration jobs.

In contrast to previous studies, which found that wives reported fewer jobs from their spouse's lifetime employment history than the husbands themselves,^{4,20} we found that mothers reported more paternal jobs on average than fathers (2.3 vs. 1.7). The mean number of

jobs (1.7) reported by fathers who completed telephone interviews (84%) was lower than the mean number of jobs (2.1) reported by fathers completing in-person interviews, suggesting that interview type may have influenced response completeness.

The reliability of mothers' reporting of fathers' exposures to specific agents was poor. The small kappa values may be explained by the low prevalence of agents reported by fathers (as prevalence approaches zero, the value of kappa also approaches zero²¹), but is also likely to be attributed to the mothers' lack of familiarity with details of the fathers' work experiences. For each category of occupational hazard (with the exception of radiation), twice as many fathers were classified as exposed when personal rather than proxy interview data were used. Such missing information has been shown to result in differential bias and underestimation of risk.²² Low kappa scores were reported in previous studies for occupational exposure, such as to fertilisers (kappa = 0.48)⁸ and solvents (kappa = 0.18).⁹

In this study, mothers were able to provide reliable data on fathers' race, education and smoking status. Previous studies have shown that proxies can report a spouse's education¹⁻³ and smoking status, but not quantity of tobacco use.¹⁻⁶

These data show that mothers were able to serve as proxy respondents for fathers' demographic data, prior smoking status and non-recent jobs in this study of childhood disease. However, they were not reliably able to report fathers' exposure to occupational agents. The focus of questions around a major life event (birth of child) shared by both parents probably contributed to the high levels of reliability for job history, in contrast to the relatively low levels reported previously for lifetime employment. Although parents' proportion of time spent together during the entire 5-year period in this study is not known, it has been shown that agreement is high when exposure occurs during time periods held in common by proxy and direct respondents.¹ These results suggest that the ability to report non-recent jobs may be influenced more by how familiar both respondents are with the events in a specific time period, rather than the duration of time since the event occurred.

Acknowledgements

This work was supported by NIH grant CA47082 from the National Cancer Institute, NRSA-NIH Cancer Epidemiology Training Grant CA09142-20 and a grant

from the Cancer Research Foundation of America. In Los Angeles, cancer incidence data were collected under contracts CA17054 from the National Institutes of Health and 050 (C-G)-8709 from the State of California Department of Health Services. In Seattle, this research was also supported by the Cancer Surveillance System of the Fred Hutchinson Cancer Research Center, which is funded by contract N01-CN-05230 from the Surveillance, Epidemiology and End Results Program of the National Cancer Institute with additional support from the Fred Hutchinson Cancer Research Center. This work was part of an international collaborative study of childhood brain tumours co-ordinated by the SEARCH Program of the International Agency for Research on Cancer. The authors thank the other collaborators who gave valuable input, particularly during the design of the study and the development of the study questions. Additionally, the authors thank Ann Hadley, Tina Paoff, Maria Paul, Isabel Gaeta, Dr Katherine Newton, Barbara Nist, Jennifer Kristiansen, Paige Bracci, Dr Jess Kraus and Susan Parrisher for their contributions to this work.

References

- 1 Pickle LW, Brown LM, Blot WJ. Information available from surrogate respondents in case-control interview studies. *American Journal of Epidemiology* 1983; **118**:99-108.
- 2 Rocca WA, Fratiglioni L, Bracco L, Pedone D, Groppi C, Schoenberg BS. The use of surrogate respondents to obtain questionnaire data in case-control studies of neurologic diseases. *Journal of Chronic Diseases* 1986; **39**:902-912.
- 3 Boyle CA, Brann EA. Proxy respondents and the validity of occupational and other exposure data. *American Journal of Epidemiology* 1992; **136**:712-721.
- 4 Lerchen ML, Samet JM. An assessment of the validity of questionnaire responses provided by a surviving spouse. *American Journal of Epidemiology* 1986; **123**:481-489.
- 5 Hatch MC, Misra D, Kabat GC, Kartzmer S. Proxy respondents in reproductive research: a comparison of self- to partner-reported data. *American Journal of Epidemiology* 1991; **133**:826-831.
- 6 Kolonel LN, Hirohata T, Nomura AMY. Adequacy of survey data collected from substitute respondents. *American Journal of Epidemiology* 1977; **106**:476-484.
- 7 Greenberg ER, Rosner B, Hennekens C, Rinsky R, Colton T. An investigation of bias in a study of nuclear shipyard workers. *American Journal of Epidemiology* 1985; **121**:301-308.
- 8 Wang FL, Semchuk KM, Love EJ. Reliability of environmental and occupational exposure data provided by surrogate respondents in a case-control study of Parkinson's disease. *Journal of Clinical Epidemiology* 1994; **47**:797-807.
- 9 Shalat SL, Christiani DC, Baker EL. Accuracy of work history obtained from a spouse. *Scandinavian Journal of Work Environment and Health* 1987; **13**:67-79.

- 10 Roeleveld N, Kiemeny L, Schattenberg G, Peer P. Information bias in a case-referent study on mental retardation and parental occupation: colleagues as dual respondents. *Epidemiology* 1990; **1**:292–297.
- 11 Hartge P, Brinton LA, Rosenthal JF, Cahill JI, Hoover RN, Waksberg J. Random digit dialing in selecting a population-based control group. *American Journal of Epidemiology* 1984; **120**:836–833.
- 12 Preston Martin S, Pogoda JM, Mueller BA, Holly EA, Lijinsky W, Davis RL. Maternal consumption of cured meats and vitamins in relation to pediatric brain tumors. *Cancer Epidemiology Biomarkers and Prevention* 1996; **5**:599–605.
- 13 Statistical Office of the United Nations Department of Economic and Social Affairs. *Indexes to the International Standard Industrial Classification of All Economic Activities (ISIC)*. Statistical Papers, Series M, no. 4, Rev. 2, Add. 1. New York: United Nations, 1971.
- 14 *International Standard Classification of Occupations*, revised edn. Geneva, Switzerland: International Labour Office, 1968.
- 15 Siegel DG, Podgor MJ, Remaley NA. Acceptable values of Kappa for comparison of two groups. *American Journal of Epidemiology* 1992; **135**:571–578.
- 16 Fleiss JL. *Statistical Methods for Rates and Proportions*. New York. John Wiley & Sons, 1981.
- 17 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**:159–174.
- 18 Bond GG, Kenneth M, Bodner M, Sobel W, Shellenberger RJ, Flores GH. Validation of work histories obtained from interviews. *American Journal of Epidemiology* 1988; **128**:343–351.
- 19 Baumgarten M, Siemiatyck J, Gibbs GW. Validity of work histories obtained by interview for epidemiologic purposes. *American Journal of Epidemiology* 1983; **118**:583–591.
- 20 Blot WJ, Harrington JM, Toledo A, Hoover R, Heath CW, Fraumeni JF. Lung cancer after employment in shipyards during World War II. *New England Journal of Medicine* 1978; **299**:620–624.
- 21 Thompson WD, Walter SD. A reappraisal of the kappa coefficient. *Journal of Clinical Epidemiology* 1988; **41**:949–958.
- 22 Ahlborg GA. Validity of exposure data obtained by questionnaire. Two examples from occupational reproductive studies. *Scandinavian Journal of Work Environment and Health* 1990; **16**:284–288.