Published in final edited form as:

Am J Ind Med. 2014 August ; 57(8): 915–927. doi:10.1002/ajim.22328.

Developing estimates of frequency and intensity of exposure to three types of metalworking fluids in a population-based case-control study of bladder cancer

Melissa C Friesen, Ph.D.¹, Dong-Uk Park, Ph.D., CIH^{1,2}, Joanne S Colt, M.P.H.¹, Dalsu Baris, MD; Ph.D.¹, Molly Schwenn, MD³, Margaret R. Karagas, Ph.D.⁴, Karla R. Armenti, Sc.D.⁵, Alison Johnson, M.B.A., C.T.R.⁶, Debra T Silverman, Sc.D.^{*,1}, and Patricia A Stewart, Ph.D.^{*,1,7}

¹Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, DHHS, 6120 Executive Blvd. EPS Room 8005, Rockville, Maryland, 20852, USA

²Department of Environmental Health, Korea National Open University, 169, Donsungdong, Jongroku, Seoul, Korea, 110-791

³Maine Cancer Registry, Augusta, Maine, USA

⁴ Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, USA

⁵New Hampshire Department of Health and Human Services, Division of Public Health Services, Bureau of Public Health Statistics and Informatics, Concord, New Hampshire, USA

⁶Vermont Department of Health, Burlington, Vermont, USA

⁷Stewart Exposure Assessments, LLC, 6045 N 27th. St, Arlington, VA 22207

Abstract

Background—A systematic, transparent, and data-driven approach was developed to estimate frequency and intensity of exposure to straight, soluble, and synthetic/semi-synthetic metalworking fluids (MWFs) within a case-control study of bladder cancer in New England.

Methods—We assessed frequency using individual-level information from job-specific questionnaires wherever possible, then derived and applied job group-level patterns to likely exposed jobs with less information. Intensity estimates were calculated using a statistical model developed from measurements and determinants extracted from the published literature.

Results—For jobs with probabilities of exposure 0.5, median frequencies were 8 to 10 hours per week, depending on MWF type. Median intensities for these jobs were 2.5, 2.1, and 1.0 mg/m³ for soluble, straight, and synthetic/semi-synthetic MWFs, respectively.

Corresponding author: Melissa C Friesen, Occupational and Environmental Epidemiology Branch, 9609 Medical Center Drive, Room 6E608, North Bethesda MD 20892. Tel: +1-240-276-7278. friesenmc@mail.nih.gov.

*denotes co-senior authors

Conflict of Interest: No conflicts of interest exist.

Conclusions—Compared to case-by-case assessment, these data-driven decision rules are transparent and reproducible and may result in less biased estimates. These rules can also aid future exposure assessments of MWFs in population-based studies.

Keywords

Metalworking fluids; bladder cancer; case-control study; retrospective exposure assessment; exposure prediction model

Introduction

In population-based case-control studies, exposure assessment efforts typically rely on participants' responses to occupational questionnaires such as lifetime occupational histories and more detailed job and industry-specific questionnaire modules. Exposure assessors use exposure-specific information captured in the reported job titles, industries, and work activities to assign exposure estimates to the agents of interest. This approach has been criticized as occurring in a 'black box' [Kromhout 2002, Teschke, et al. 2002], but the expert-based decision rules need not be opaque. For instance, an assessment of occupational diesel exhaust exposure in a case-control study used systematically asked occupational questions to algorithmically link questionnaire responses to exposure decisions [Pronk, et al. 2009] and software has been developed to facilitate the application of decision rules for newly initiated studies [Fritschi, et al. 2009].

In this paper, we describe a systematic, transparent approach for estimating the frequency and intensity of exposure to straight, soluble, and synthetic/semi-synthetic metalworking fluids (MWFs) for subjects within a population-based case-control study of bladder cancer in New England [Baris, et al. 2009, Colt, et al. 2011]. MWFs, oil mists, or metal machining activities have been previously associated with an increased risk of bladder cancer in this and other studies [Band, et al. 2005, Brownson, et al. 1987, Colt, et al. 2011, Cordier, et al. 1993, Friesen, et al. 2009, Howe, et al. 1980, Silverman, et al. 1983, Silverman, et al. 1989, Silverman, et al. 1989, Vineis and Di Prima 1983]. Studies varied in the level of detail in the exposure metrics, including using surrogates of exposure (e.g., metal machining activities) or treating MWFs as a single exposure agent; however, recent evidence from an autoworkers cohort suggest that bladder cancer risk may be isolated to straight MWFs [Friesen, et al. 2009, Friesen, et al. 2011].

The New England Bladder Cancer Study incorporated lifetime occupational history plus job-specific modules that asked study participants detailed questions on work activities, including metal machining and type and frequency of MWF use. Our approach was similar to previous decision rule-based exposure assessment efforts [Behrens, et al. 2012, Friesen, et al. 2012, Fritschi, et al. 2009, Macfarlane, et al. 2012, Pronk, et al. 2012]; however, we used a data driven assignment approach, rather than expert judgment wherever possible. We maximized the module information in a hierarchical, data-driven exposure assessment approach to estimate each subject's frequency and intensity of exposure to straight, soluble, and synthetic/semi-synthetic MWFs. Frequency of exposure was assessed using individual-level information wherever possible and, for likely exposed jobs with less information,

frequency estimates were based on job group-level patterns derived from the individual-level information. Intensity estimates were calculated using parameters from a statistical model of measurements extracted from the published literature that accounted for time-, industry-, operation-, and MWF-specific differences. The process for estimating the probability of MWF exposure for these same subjects is reported separately [Park, et al. submitted]. For the present study, we developed decision rules to estimate frequency and intensity of MWF exposure for the New England Bladder Cancer Study. These rules can be used to assess exposure to MWF in future population-based studies.

Methods

Study population

The study population has been previously described and were enumerated using a protocol approved by the National Cancer Institute Special Studies Institutional Review Board, as well as the human subjects review boards of each participating institution [Baris, et al. 2009]. Each participant provided signed informed consent. Briefly, 1171 urothelial carcinoma patients and 1418 control subjects were interviewed in Maine, New Hampshire, and Vermont from 2001 to 2004. Patients were identified through hospital pathology departments, hospital cancer registries, and the state cancer registries. Controls aged 30-64 were selected randomly from Department of Motor Vehicle records in each state and controls aged 65-79 were selected from beneficiary records of the Centers for Medicare and Medicaid Services. We excluded 13 cases and 16 controls from the following MWF exposure assessment because they were absentee business owners, they had never had paid work, or they were missing information on potential confounders. We further restricted the assessment to 895 male cases and 1031 male controls because MWF exposure was rare among women.

The occupational information collected by in-person interviews with study participants, as previously described [Colt, et al. 2011], provided whole life occupational histories for each job held for at least six months since the age of sixteen. For each job, the subject was asked open-ended questions for the job title, employer name, year job started and ended, type of business, products made or services provided, and the subject's main duties and activities, the tools and equipment used, and the chemicals and materials handled. If the job occurred in an occupation or industry of interest because of its potential for bladder cancer risk or if the subject's response to the open-ended occupational questions included *a priori* identified keywords, potentially appropriate job- or industry-specific modules appeared on the computer screen during the interview and the interviewer selected the most appropriate module.

The job modules asked questions about work tasks related to a variety of substances. In particular, MWF and machining-related questions were asked in the machinist, electrician, industrial machinery mechanic, plumber, sheet metal worker, tool and die worker and welder modules (hereafter, MWF modules) (question 1 to 5 in **Table I**). The type of MWF was identified using questions that included the specific MWF's color and texture, which differed by MWF, to aid recall (e.g., questions 2, 3 and 4 in **Table I**) [Park et al., submitted]. To reduce burden to the respondent, employment in jobs for less than 1000 hours

did not trigger a module. In addition, all MWF modules except the machinist module contained skip patterns that asked questions on specific MWFs only if the subjects machined for 5 hr/week.

Identifying potentially exposed jobs

We reviewed the occupational histories and MWF modules to identify questions that provided information on MWF use. The most useful questions are listed in **Table I**, which also lists the derived variables that are discussed below. All jobs were assigned a probability of exposure to MWF (scale: 0-1.00) based on information reported in the modules and occupational histories, as previously described [Park, et al. submitted]. Briefly, the probability of exposure was estimated for straight, soluble, and synthetic/semi-synthetic fluids. Synthetic and semi-synthetic MWF were considered together because the identification of these two fluid types was not expected to be easily distinguishable by study subjects. Possibly exposed jobs were assigned, after reviewing the information from the occupational history and (if available) the MWF module responses, to one of the following eleven job groups: machinist, electrician, farmer, industrial machinery mechanic, plumber, sheet metal worker, tool and die worker, welder, bystander (i.e., a supervisory or technical worker who was not expected to have machined but was likely to be in the area of machining), production (a blue-collar worker not expected to have machined but who was likely to be in the area of metal machining), and miscellaneous. A job may have been assigned a job group other than the module it received if, based on the occupational history and module information, the tasks suggested a more appropriate module (<10% of jobs with MWF-related modules). Each possibly exposed job was then assigned a probability rating, and a confidence rating for the probability estimate (scale of 1-4), based on its job group and responses to MWF-related questions based on the specificity of the information provided [Park et al., submitted].

Each job with a probability greater than 0 was assigned estimates of frequency (hours per week) and intensity (mg/m^3) of exposure to straight, soluble, and synthetic/semi-synthetic MWF using the process described below. Each job also received a confidence rating for the frequency and the intensity estimates ['frequency confidence' and 'intensity confidence,' ordinal scale: 1 (lowest) to 4 (highest)] based on the criteria described in sections 'Assignment of frequency estimates' and 'Assignment of intensity estimates'.

Assignment of frequency estimates

The rules for assigning frequency estimates to each job were hierarchical, using 1) subject-specific information where available and 2) assigning job group medians if subject-level information was unavailable (**Figure 1**). The subject-specific information came from jobs with a completed MWF module that had a valid response to the questions identified as 'FreqStraight', 'FreqSoluble' and 'FreqSynthetic' in **Table I**. These jobs were assigned the reported frequency from that question and given a frequency confidence rating of 4 (**Fig. 1, box A**). These responses were then used to calculate across all subjects with valid information in each job group: 1) the percentage of jobs machining (%JobsMachining, J_G); 2) the percentage of jobs using each MWF type (e.g., %JobsStraight, J_G); 3) the median frequency of machining overall (MedianFreqMachine, J_G); and 4) the median frequency of

use by MWF type (e.g., MedianFreqStraight_{JG}). For machinists, the median frequency for soluble, straight, and synthetic/semi-synthetic fluids was derived separately for each decade. For the other job groups, there were too few jobs to develop decade-specific estimates (n=0-24).

When either invalid or missing responses were provided (e.g., don't know), the information was missing because of the administration of an incorrect or a non-MWF (e.g., engineer) module, the job duration was less than 1000 hours, or the subject machined <5 hr/wk for specific modules, we moved to the job group level. For jobs with a valid 'FreqMachine' response, we multiplied that value by the median percentage of time each MWF type was used (compared to all MWF types) for each job group (e.g., 'Median%FreqStraight_{JG}') (**Fig. 1, box B**). The resulting value was assigned as the frequency estimate with a confidence of 3. If the job was in a job group that was missing the median percent time for a MWF, the median percent use for that MWF across all job groups (e.g., 'MedianFreqStraight_{MWFJobs}') was assigned with a confidence of 2. If, however, the job group was by-stander, plumber, production, or electrician, the confidence assigned was 1, because of the small number of responses in those groups (<5).

If both FreqMWF and FreqMachine were missing or invalid, the median of the job group's MWF frequency (e.g., 'MedianFreqStraight_{JG}') was assigned with a confidence of 2 (**Fig. 1, box C**). However, under the same conditions if the job group was by-stander, plumber, production, or electrician, the confidence assigned was 1, due to the small number of responses.

The farmer module did not include any MWF questions; however, it did query the amount of time spent repairing/maintaining equipment. Since many farmers machine as part of repairing/maintaining, an arbitrary 10% of the frequency reported for this task was assigned as the frequency estimate, with a confidence level of 2. All jobs with an assigned probability of 0 received a frequency of 0.

Development of exposure intensity model

MWF aerosol measurements and exposure determinants had been previously extracted from the literature [Park, et al. 2009a,b]. Briefly, the arithmetic means of all personal and area MWF measurements identified in the literature with a sampling duration of one hour or more were compiled into a database. When the arithmetic mean was not available, it was estimated from the geometric mean and geometric standard deviation or from the exposure range using equations described by Hein et al. 2008 [Hein, et al. 2008]. For each arithmetic mean, we extracted the particle size fraction (total, thoracic, respirable), the sampling and analytical method characteristics (sample duration, sample type, filter type, analytic method), and several potential determinants of exposure.

We focused on exposure and sampling determinants that could be reported and recalled by study subjects and that could be extracted from the limited information described in the literature rather than on more technical determinants. Thus, the following determinants were evaluated: decade (<1970, 1970s, 1980s, 1990s, 2000s), type of industry (automobile, automobile parts, small job shops and other), type of operation (grinding or other

machining), MWF type (soluble, straight, synthetic, semi-synthetic), sample duration (1-2 hours, 2-7 hours, >7 hours), sample type (personal vs. area), sampling device (filter, impactor, unidentified), filter type (GF, PVC, PTFE, MCE, unidentified), and analytic method (gravimetric, extraction) [Park, et al. 2009b]. These factors had been previously found to be associated with MWF concentrations in single-factor analysis of variance models [Park, et al. 2009b], but had not been examined simultaneously in a multivariate model.

To develop a multivariate model to predict intensity of MWF exposure, we included only the 'total' aerosol MWF measurements because of the sparse measurements and coverage across the determinants of the respirable and inspirable fractions. These area and personal measurements were predominantly (82%) collected using 37mm closed-face cassettes [Park, et al. 2009a,b]. The measurements were also restricted to measurements with a minimum duration of 1 hour. Overall, 155 arithmetic means representing 9,379 'total' aerosol measurements were included in the database. The measurements were positively skewed and approximately log-normal (Park et al. 2008) so the natural logarithm of the concentration was used in all data analyses. Decade, industry, operation, MWF type, sample duration, sample type, sample device, filter type, and analytic method were offered as potential explanatory variables into a weighted regression model, where each log-transformed arithmetic mean was weighted by its number of measurements. Only main effects were included in the model, because there were too few observations to examine interactions. All variables were initially offered in the models; the variable with the highest p-value was removed and the model was re-run until only variables with p-values <0.05 remained. All analyses were conducted using Stata Version 9.9 (StataCorp, College Station, Texas, USA).

Assignment of intensity estimates

The final intensity model included five factors associated with MWF concentrations: sample type, decade, industry, operation, and MWF type.

Decade was assigned for each calendar year of each job (<1970, 1970s, 1980s, 1990s, 2000s).

Industry was assigned based on the type of business or service reported in the occupational history. If the industry produced metal products, the job was assigned to either the automobile, automotive parts, or small job shops industry category. Jobs in the industries of car, truck or other similar vehicle production were assigned the auto industry. Small job shops included operations that were characterized as making a variety of metal products on demand. The automotive parts industry was assigned to other metal products because the size of the parts was small compared to the auto industry and the size of the operation was often large and more routine compared to small job operations. If the industry produced only non-metal products (e.g., textiles) or was a service industry (schools, hospitals), the job was grouped with the small job shops category because the machining operations (e.g., in a maintenance shop) were likely to be small. The other industry category originally included in the model was not used because this category had too few measurements.

Type of operation was assigned based on the ‘TypeMachine’ question from the job modules. The number of grinders (#Grinders in **Table I**) and the number of other metal working machines (#OtherMachines), in response to “What kind of machining equipment did you use?” (question 2 in Table I), were counted for each reported job. Metal working machines reported in the occupational histories were also counted. Where the number of machines was missing or invalid, we assumed, based on the median number of machines reported to question 2 in Table I, 3 machines and 1 grinder if the job group was machinist and 2 machines and 1 grinder for all other job groups.

To derive estimates of intensity, the model parameters associated with the assigned decade and industry type were used to calculate the intensity for grinding operations (IntGrind) and for other machining operations (IntMachining) for each of the three MWF types, using the model parameter for personal sample type. For each MWF type, the final intensity assigned to the job was a weighted average based on the number of grinders and the number of other metal working machines: $(\#Grinders * IntGrind + \#OtherMachines * IntMachining) / (\#Grinders + \#OtherMachines)$. Because we combined synthetic and semi-synthetic in the exposure assessment, only the synthetic intensity estimates were used.

The intensity confidence rating was assigned based on the availability of information on the number of grinders and other metal working machines. The intensity confidence was a four if the number of machines was reported and a two where the number was missing or invalid.

Descriptive statistics of frequency and intensity metrics

We report the median and range of the frequency and intensity metrics summarized across all exposed calendar years for each job group, for all jobs with exposure probability rating >0 and for jobs with exposure probability rating 0.50. The latter is presented to coincide with the categories analyzed in the epidemiologic study.

For the epidemiologic analyses, the intensity estimate, frequency estimate, and job duration of each job with an assigned probability 0.5 were multiplied together and then summed across all jobs with probability 0.5 held by that subject to calculate cumulative exposure. The cumulative exposure estimates are described by Colt et al [submitted].

Results

For jobs with MWF modules, the number, the proportion of exposed jobs, and the median frequency of exposure for the three MWF types for each job group, and across all jobs combined, are listed in **Table II**. For jobs with MWF modules, 58% reported being exposed to soluble MWF, 63% to straight MWF, and 17% to synthetic MWF. The proportion of jobs with exposure to each type of MWF varied by job group, with no consistent patterns observed. Machinists had the highest frequency of exposure to all three MWF types (median = 20 hours per week for soluble and straight MWF; 15 hours per week for synthetic MWF). Other job groups had median frequencies ranging from 0.5 to 22, depending on the job group and MWF type. These values were used as job group estimates when frequency information was not obtained from subjects. Synthetic MWF had the fewest available data

and thus the median across all job groups of 10 hours per week was assigned for this MWF for most job groups.

The parameter estimates for the ‘total’ aerosol intensity MWF model, which included the variables of sample type, decade, industry, operation, and fluid type, are listed in **Table III**. These variables explained 59% of the variance in the measurements. Exposure levels decreased with increasing decade, with concentrations in the 2000s only 11% of predicted concentrations for pre-1970s. Exposure levels in the automobile parts industries and small jobs operations were 2.08 and 1.13 times higher, respectively, than in the automobile industry. Exposure levels for non-grinding machining operations were 73% of the concentrations predicted for grinding operations. Exposure levels were highest for straight MWF. The other fluid types had predicted concentrations 60-71% of the straight MWF concentrations.

The medians and ranges of the assigned frequency and intensity estimates are presented for jobs assigned probabilities >0 and <0.5 to soluble, straight, and synthetic MWFs in **Table IV**. Farmers, machinists, mechanics, and by-standers were the most prevalent by number of jobs and exposed years for jobs with probabilities >0 . Machinists were the most prevalent for jobs with probabilities <0.5 . Median frequencies were much higher for jobs with probabilities <0.5 than >0 for all MWF types: 10 vs. 1.0 for soluble MWF; 8.0 vs. 1.2 for straight MWF, and 10 vs. 1.3 for synthetic MWFs. The frequency estimates were highly variable within job groups, indicating large within-job differences in the use of MWFs. Higher median frequencies (>4 hours per week) were observed for machinists, miscellaneous jobs, plumbers, and tool and die workers than for bystanders, electricians, farmers, mechanics, production workers, sheet metal workers and welders (<2 hours per week).

For jobs with probabilities >0 , the median intensities were 1.2, 1.8, and 1.0 mg/m^3 for soluble, straight, and synthetic/semi-synthetic MWFs, respectively (**Table IV**). The median intensities were somewhat higher for soluble MWF (2.5 mg/m^3) and straight MWF (2.1 mg/m^3) when restricted to jobs with probabilities <0.5 . The median and range of intensities varied by job group and MWF type; some differences were also observed by probability rating. By MWF type, the highest median intensities occurred for tool and die workers exposed to soluble MWF (4.7 mg/m^3), sheet metal workers exposed to straight MWF (7.3 mg/m^3), and sheet metal workers exposed to synthetic MWF (5.2 mg/m^3). The lowest median intensities occurred for farmers (soluble MWF, median=0.82 mg/m^3 ; straight MWF, median = 1.2 mg/m^3 ; synthetic MWF, median = 0.83 mg/m^3) and welders (soluble MWF, median=0.57-0.70 mg/m^3 ; straight MWF, median = 0.82-1.2 mg/m^3 ; synthetic MWF, median = 0.58-0.83 mg/m^3). Similarly low median concentrations were also observed for electricians (for probabilities >0) and miscellaneous jobs.

Discussion

This study is one of the few case-control studies to develop occupational exposure estimates using a systematic, objective, data-driven procedure. We directly linked responses to occupational histories and job-specific modules to expert-based exposure decision rules to

assess frequency and intensity of exposure to three types of MWFs. These data-driven decision rules are transparent and reproducible and may be more likely to result in less biased estimates than using expert judgment alone. In addition, the rules make it possible for sensitivity analyses to be conducted to test the robustness of exposure-disease associations to exposure decisions.

Frequency estimates were based on direct reports from study subjects in their responses to questions on the frequency of machining and the type and frequency of MWF use asked within several modules (e.g., machinist, tool and die, sheet metal worker) or, for subjects with less information, based on job group patterns derived from the more detailed responses. As a result, the assignments were data-driven, rather than based on expert judgment. The wide ranges in the reported frequency of machining activities and use of the various MWFs provide support for asking the more detailed module questions whenever possible to minimize exposure misclassification. While applying job group patterns may introduce some error, group-based assignment approaches have been previously found to follow a Berkson error structure and thus is unlikely to bias risk estimates [Armstrong 1998].

Intensity estimates were derived from a statistical model based on a large number of literature-based measurements collected over many years. Using a statistical model of the published data allowed us to implement a data-driven approach to account for decade-, industry-, operation-, and MWF-specific differences in exposure. The predictive factors and the magnitudes of their parameter estimates in the MWF model were consistent with the patterns observed in prior analyses of these data [Park, et al. 2009]. In the multi-variate model developed here, exposures decreased approximately 90% from pre-1970 to the 2000s, with the largest decline observed for the 1970s. Decade likely served as a proxy for changes in operations and technology that could not be extracted and modeled from the descriptive data in the published literature. For example, during the 1970s and 1980s, many work sites in the United States installed recirculating air cleaners, improved recirculating air filtration systems, partially enclosed machines, installed local exhaust ventilation, and upgraded work conditions (65, NIOSH, 1998). Predicted exposures were higher for grinding operations than for nongrinding machining operations. In many cases the subject reported multiple operations; thus, the resulting intensity estimates were weighted averages of grinding and other machining activities. Predicted exposures in the automobile part industry were twice as high as in the automobile industry and small jobs operations. Predicted concentrations were higher for straight MWF than for soluble, synthetic, and semi-synthetic MWFs. The medians and ranges of the resulting intensity estimates varied by job group, MWF type, decade and probability rating, demonstrating that the application of the model captured important contrasts in exposure intensities for the study subjects.

Our hierarchical, data-driven approach accounted for some within-job heterogeneity in frequency and intensity of exposure captured in the detailed module responses that could not have been captured through the use of alternative exposure assessment methods, such as job exposure matrices or analyses by job title, industry, or duration. Expert judgment was used transparently, based on the rules described in this paper, to assign subjects to job groups, to link the industries worked by the study subjects to the three industries for which there was sufficient monitoring data, and to decide how to use the available information, such as how

to weigh different operations for workers performing different machining operations and how to characterize the estimates' reliability (i.e., confidence ratings). Because expert judgment was used transparently, sensitivity analyses can be conducted to evaluate the robustness of the exposure decisions.

Our approach had several limitations that are inherent in any population-based study. The exposure data used to derive the exposure models were not specific to the participants' work sites and were based solely on determinants that could be easily recalled by study subjects, rather than previously identified determinants of exposure (e.g., level of engineering controls (Park et al. 2009b)). Similarly, the frequency estimates relied on subject recall of work activities, often decades after performing the job tasks.

In addition, our approach had several limitations that were specific to the data available in this study. The available MWF measurements were insufficient to detect possible differences in exposure related to sample duration, sampling method, and analytical method [Park et al. 2009b], to examine possible interactions between exposure determinants (e.g., industry- or MWF-specific time trends), or to predict MWF intensity by other particle size fractions that may have important health effects. In addition, no data were available for most of the specific metal product industries reported by the study subjects and thus the most similar broad industry group had to be assigned. The data used to derive job group patterns were usually too sparse, especially for synthetic MWF, to calculate time-specific patterns for most job groups. Finally, the wide range of frequency and intensity estimates observed within each job group suggests that there was likely random error in the frequency and intensity estimates for the jobs without individual-specific information. However, in epidemiologic analyses our overall objective is to identify broad differences in exposure, which were captured in the model intensity estimates and in the frequency estimates used here, that can be used to differentiate between low, medium, and highly-exposed subjects for epidemiologic analyses.

In summary, we describe a framework for estimating frequency and intensity of MWF exposure in a case-control study that systematically used the available MWF measurements and applied a tiered, data-driven approach that maximized individual-level module data where possible and applied job group-level patterns when less information was available. Our approach was transparent and accounted for some within-job heterogeneity that could not be captured through the use of job exposure matrices or analyses by job title, industry, or duration. The decision rules presented here may assist other researchers in assessing MWF exposure in epidemiologic studies with similar occupational information. In addition, the general framework used to develop those rules could be applied to other agents.

Acknowledgments

We thank Lonn Tremblay at IMS for data programming support. This project was funded by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health.

REFERENCES

- Armstrong BG. Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occup Environ Med.* 1998; 55:651–656. [PubMed: 9930084]
- Band PR, Le ND, MacArthur AC, Fang R, Gallagher RP. Identification of occupational cancer risks in British Columbia: a population-based case-control study of 1129 cases of bladder cancer. *J Occup Environ Med.* 2005; 47:854–858. [PubMed: 16093936]
- Baris D, Karagas MR, Verrill C, Johnson A, Andrew AS, Marsit CJ, Schwenn M, Colt JS, Cherala S, Samanic C, Waddell R, Cantor KP, Schned A, Rothman N, Lubin J, Fraumeni JF Jr, Hoover RN, Kelsey KT, Silverman DT. A case-control study of smoking and bladder cancer risk: emergent patterns over time. *J Natl Cancer Inst.* 2009; 101:1553–1561. [PubMed: 19917915]
- Behrens T, Mester B, Fritschi L. Sharing the knowledge gained from occupational cohort studies: a call for action. *Occup Environ Med.* 2012; 69:444–448. [PubMed: 22213839]
- Brownson RC, Chang JC, Davis JR. Occupation, smoking, and alcohol in the epidemiology of bladder cancer. *Am J Public Health.* 1987; 77:1298–1300. [PubMed: 3631363]
- Colt JS, Friesen MC, Stewart PA, Park DU, Johnson A, Schwenn M, Karagas MR, Armenti K, Waddell R, Verrill C, Ward MH, Beane Freeman LE, Moore LE, Baris D, Silverman DT. A case-control study of occupational exposure to metalworking fluids and bladder cancer risk among men. (Submitted).
- Colt JS, Karagas MR, Schwenn M, Baris D, Johnson A, Stewart P, Verrill C, Moore LE, Lubin J, Ward MH, Samanic C, Rothman N, Cantor KP, Beane Freeman LE, Schned A, Cherala S, Silverman DT. Occupation and bladder cancer in a population-based case-control study in Northern New England. *Occup Environ Med.* 2011; 68:239–249. [PubMed: 20864470]
- Cordier S, Clavel J, Limasset JC, Boccon-Gibod L, Le Moual N, Mandereau L, Hemon D. Occupational risks of bladder cancer in France: a multicentre case-control study. *Int J Epidemiol.* 1993; 22:403–411. [PubMed: 8359955]
- Friesen MC, Costello S, Eisen EA. Quantitative exposure to metalworking fluids and bladder cancer incidence in a cohort of autoworkers. *Am J Epidemiol.* 2009; 169:1471–1478. [PubMed: 19414495]
- Friesen MC, Costello S, Thurston SW, Eisen EA. Distinguishing the common components of oil- and water-based metalworking fluids for assessment of cancer incidence risk in autoworkers. *Am J Ind Med.* 2011; 54:450–460. [PubMed: 21328414]
- Friesen MC, Pronk A, Wheeler DC, Chen YC, Locke SJ, Zaebs DD, Schwenn M, Johnson A, Waddell R, Baris D, Colt JS, Silverman DT, Stewart PA, Katki HA. Comparison of Algorithm-based Estimates of Occupational Diesel Exhaust Exposure to Those of Multiple Independent Raters in a Population-based Case-Control Study. *Ann Occup Hyg.* 2012
- Fritschi L, Friesen MC, Glass D, Benke G, Girschik J, Sadkowsky T. OccIDEAS: Retrospective Occupational Exposure Assessment in Community-Based Studies Made Easier. *J Environ Public Health.* 2009; 2009:957023. [PubMed: 20041014]
- Hein MJ, Waters MA, van Wijngaarden E, Deddens JA, Stewart PA. Issues when modeling benzene, toluene, and xylene exposures using a literature database. *J Occup Environ Hyg.* 2008; 5:36–47. [PubMed: 18041643]
- Howe GR, Burch JD, Miller AB, Cook GM, Esteve J, Morrison B, Gordon P, Chambers LW, Fodor G, Winsor GM. Tobacco use, occupation, coffee, various nutrients, and bladder cancer. *J Natl Cancer Inst.* 1980; 64:701–713. [PubMed: 6928984]
- Kromhout H. Commentary. *Occupational and Environmental Medicine.* 2002; 59:594–594.
- Macfarlane E, Benke G, Sim MR, Fritschi L. OccIDEAS: An Innovative Tool to Assess Past Asbestos Exposure in the Australian Mesothelioma Registry. *Safety and health at work.* 2012; 3:71–76. [PubMed: 22953234]
- Park D, Stewart PA, Coble JB. A comprehensive review of the literature on exposure to metalworking fluids. *J Occup Environ Hyg.* 2009a; 6:530–541. [PubMed: 19544177]
- Park D, Stewart PA, Coble JB. Determinants of exposure to metalworking fluid aerosols: a literature review and analysis of reported measurements. *Ann Occup Hyg.* 2009b; 53:271–288. [PubMed: 19329796]

- Park D, Colt JS, Baris D, Schwenn M, Karagas MR, Armenti KR, Johnson A, Silverman DT, Stewart PA. Estimation of the probability of exposure to metalworking fluids in a population-based case-control study. Submitted.
- Pronk A, Coble J, Stewart PA. Occupational exposure to diesel engine exhaust: a literature review. *J Expo Sci Environ Epidemiol*. 2009; 19:443–457. [PubMed: 19277070]
- Pronk A, Stewart PA, Coble JB, Katki HA, Wheeler DC, Colt JS, Baris D, Schwenn M, Karagas MR, Johnson A, Waddell R, Verrill C, Cherala S, Silverman DT, Friesen MC. Comparison of two expert-based assessments of diesel exhaust exposure in a case-control study: programmable decision rules versus expert review of individual jobs. *Occup Environ Med*. 2012; 69:752–758. [PubMed: 22843440]
- Silverman DT, Hoover RN, Albert S, Graff KM. Occupation and cancer of the lower urinary tract in Detroit. *J Natl Cancer Inst*. 1983; 70:237–245. [PubMed: 6571931]
- Silverman DT, Levin LI, Hoover RN. Occupational risks of bladder cancer in the United States: II Nonwhite men. *J Natl Cancer Inst*. 1989; 81:1480–1483. [PubMed: 2778835]
- Silverman DT, Levin LI, Hoover RN, Hartge P. Occupational risks of bladder cancer in the United States: I. White men. *J Natl Cancer Inst*. 1989; 81:1472–1480.
- Teschke K, Olshan AF, Daniels JL, De Roos AJ, Parks CG, Schulz M, Vaughan TL. Occupational exposure assessment in case-control studies: opportunities for improvement. *Occup Environ Med*. 2002; 59:575–593. discussion 594. [PubMed: 12205230]
- Vineis P, Di Prima S. Cutting oils and bladder cancer. *Scand J Work Environ Health*. 1983; 9:449–450. [PubMed: 6673102]

Figure 1.

Hierarchical rules for assigning frequency of MWF exposure for jobs with probability >0 , where MWF denotes soluble, straight, and synthetic MWF, *SUBJ* denotes subject-level information, *JG* denotes job group-level information, and *MWFJobs* denotes information from all jobs receiving a MWF module. Definition and source of variables listed in Table I.

Table I

Exposure information extracted and derived from modules and occupational histories.

Question	Variable Name	Units or Categories ^a
<i>Variables from modules</i>		
1. On average, how often did you do any kind of machining or making of metal parts? (answers converted to hours per week)	FreqMachine	Units: Hours per week
2. What kind of machining equipment did you use?	TypeMachine	Categories: Specific types of metal working machines
3. On average, how often did you use straight cutting oils or oils that look and feel like motor oil? (answers converted to hours per week)	FreqStraight	Units: Hours per week
4. On average, how often did you use soluble cutting oils or oils that are milky white? (answers converted to hours per week)	FreqSoluble	Units: Hours per week
5. On average, how often did you use synthetic cutting oils or oils that feel like water and may be brightly colored, such as green or blue? (answers converted to hours per week)	FreqSynthetic	Units: Hours per week
<i>Derived Variables from occupational history</i>		
Job group obtained from job title.	Job Group	Machinist, electrician, industrial machinery mechanic, plumber, sheet metal worker, tool and die, welder, bystander, production, farmer, miscellaneous.
Industry type obtained from industry business or service.	TypeIndustry	Metal, non-metal, automobile, automobile parts, small job shops.
Job begin year/end year	Decade	<1970, 1970s, 1980s, 1990, 2000s
<i>Derived variables at subject-level, extracted from TypeMachine</i>		
Number of grinders used	#Grinders	Continuous, discrete
Number of other machines used	#OtherMachines	Continuous, discrete
Number of machines	#Machines = #Grinders + #OtherMachines	Continuous, discrete
<i>Derived variables for each job group (denoted J_G)</i>		
Median FreqMachine of all jobs in job group if FreqMachine>0.	MedianFreqMachine _{JG}	Units: Hours per week
Number of jobs reporting FreqMachine>0 divided by the number of jobs reporting FreqMachine = 0, by job group. ^b	%JobsMachining _{JG}	Continuous, discrete
Number of jobs reporting FreqStraight>0 divided by the number of jobs reporting FreqStraight = 0, by job group. ^b	%JobsStraight _{JG}	Continuous, discrete
Median FreqStraight if FreqStraight>0, by job group. ^b	MedianFreqStraight _{JG}	Units: Hours per week
Median of FreqStraight divided by the sum of FreqStraight + FreqSoluble + FreqSynthetic for all jobs, by job group. ^b	Median%FreqStraight _{JG}	Continuous
Median FreqStraight of all jobs receiving MWF module if FreqStraight>0. ^c	MedianFreqStraight _{MWFJobs}	Units: Hours per week
Number of jobs reporting FreqSoluble>0 divided by the number of jobs reporting FreqSoluble = 0, by job group. ^b	%JobsSoluble _{JG}	Continuous, discrete
Median FreqSoluble if FreqSoluble>0, by job group. ^b	MedianFreqSoluble _{JG}	Units: Hours per week
Median of FreqSoluble divided by the sum of FreqStraight + FreqSoluble + FreqSynthetic for all jobs, by job group. ^b	Median%FreqSoluble _{JG}	Continuous

Question	Variable Name	Units or Categories ^a
Median FreqSoluble of all jobs receiving MWF module if FreqSoluble>0. ^c	MedianFreqSoluble _{MWFJobs}	Units: Hours per week
Number of jobs reporting FreqSynthetic>0 divided by the number of jobs reporting FreqSynthetic=0, by job group. ^b	%JobsSynthetic _{JG}	Continuous, discrete
Median FreqSynthetic ^b if FreqSynthetic>0, by job group.	MedianFreqSynthetic _{JG}	Units: Hours per week
Median of FreqSynthetic divided by the sum of FreqStraight + FreqSoluble + FreqSynthetic for all jobs, by job group. ^b	Median%FreqSynthetic _{JG}	Continuous
Median FreqSynthetic of all jobs receiving MWF module if FreqSynthetic>0. ^c	MedianFreqSynthetic _{MWFJobs}	Units: Hours per week

#, number; Freq., frequency; MWF, metalworking fluids

^aFinal categories, after grouping original responses.

^bDerived separately for each job group. For machinists, derived for each decade. Derivation of these variables described in Park et al. (in preparation).

^cIncludes all job groups except farmers, who did not receive a MWF module.

Table II

Proportion of jobs exposed and the median exposed hours per week, by MWF type and job group, for subjects reporting >0 hours working with specified MWF in the New England Bladder Cancer Case-Control Study.

Job Group	Soluble MWF			Straight MWF			Synthetic MWF		
	N Responses to FreqSoluble ^a	% exposed ^b	Median hours per week exposed ^c	N Responses to FreqStraight ^a	% exposed ^b	Median hours per week exposed ^c	N Responses to FreqSynthetic ^a	% exposed ^b	Median hours per week exposed ^c
By-stander	4	50	1.6	4	25	1	4	25	1
Electrician	1	<i>d</i>	<i>d</i>	1	<i>d</i>	<i>d</i>	1	<i>d</i>	<i>d</i>
Machinist	94	66	20	92	65	20	92	22	15
Mechanic	22	50	0.5	21	52	1.2	16	<i>d</i>	<i>d</i>
Miscellaneous	6	67	4.2	6	67	0.7	3	<i>d</i>	<i>d</i>
Plumber	4	<i>d</i>	<i>d</i>	4	50	10.5	4	<i>d</i>	<i>d</i>
Production	2	<i>d</i>	<i>d</i>	2	50	1	2	<i>d</i>	<i>d</i>
Sheet metal worker	8	38	1	8	63	1	2	<i>d</i>	<i>d</i>
Tool and die	14	64	8	14	71	22	2	<i>d</i>	<i>d</i>
Welder	9	44	1	9	78	0.5	1	<i>d</i>	<i>d</i>
All jobs receiving a MWF module ^e	164	58	12	161	63	10	127	17	10

^aIncludes responses 0 hours per week to question related to specified variable (see Table 1). Does not include 'don't know', 'refused', or 'not asked'.

^bVariables '%JobsSolubleJG', '%JobsStraightJG', and '%JobsSyntheticJG' in Table 1.

^cVariables 'MedianFreqSolubleJG', 'MedianFreqStraightJG', and 'MedianFreqSyntheticJG' in Table 1.

^dThe overall median for all jobs receiving a MWF module was used for the MWF frequency when reports for a job group/MWF were sparse.

^eExcludes farmers, who did not receive a MWF module and thus were not asked the FreqSoluble, FreqStraight, or FreqSynthetic questions.

Table III

Parameter estimates from a regression model developed to estimate exposure intensity to MWF (units, natural-log transformation of mg/m³).

Independent variable (number of measurements in each category)	Coefficient, β	SE	P-value	$\exp(\beta)^a$
Intercept	1.46	0.40	<0.001	4.31
Sample type				
Personal (n=6652)	Reference			1.00
Area (n=1339)	0.34	0.19	0.085	1.40
Both (n=1388)	1.16	0.20	<0.001	3.19
Decade				
<1970s (n=311)	Reference			1.00
1970s (n=874)	-1.22	0.37	0.001	0.30
1980s (n=1085)	-1.62	0.36	<0.001	0.20
1990s (n=6002)	-1.58	0.30	<0.001	0.21
2000s (n=1107)	-2.22	0.36	<0.001	0.11
Industry				
Auto (n=1775)	Reference			1.00
Auto part (n=1126)	0.73	0.29	0.012	2.08
Small-jobs (n=4751)	0.12	0.26	0.628	1.13
Other/Unidentified (n=1727) ^b	0.00	0.25	0.999	1.00
Operation				
Grinding (n=1005)	Reference			1.00
Non-grinding (n=3583)	-0.32	0.16	0.049	0.73
Other/Unidentified (n=4791) ^b	-0.49	0.15	0.002	0.61
MWF type				
Straight (n=1406)	Reference			1.00
Soluble (n=2233)	-0.36	0.16	0.022	0.70
Synthetic (n=321)	-0.34	0.27	0.199	0.71
Semi-synthetic (n=551) ^b	-0.51	0.22	0.023	0.60
Other/Unidentified (n=4868) ^b	-0.51	0.15	0.001	0.60

^aPredicted geometric mean (GM) of the AMs in the literature = $\exp(\beta_{\text{intercept}}) * \exp(\beta_{\text{decade}}) * \exp(\beta_{\text{industry}}) * \exp(\beta_{\text{operation}}) * \exp(\beta_{\text{MWF Type}})$. Used to calculate intensity for grinding operations (IntGrind) and for other machining operations (IntMachining) for each of the three MWF types.

^bNot used in the development of the intensity estimates.

Descriptive statistics of the frequency (in hours per week) and intensity (in mg/m³) metrics by MWF type and job group for MWF-exposed subjects in the New England Bladder Cancer Case-Control Study.

Table IV

Job Group by probability level	Soluble					Straight					Synthetic				
	Frequency, Median (Range)	Intensity, Median (Range)	N jobs	N subjects	N exposure years	Frequency, Median (Range)	Intensity, Median (Range)	N jobs	N subjects	N exposure years	Frequency, Median (Range)	Intensity, Median (Range)	N jobs	N subjects	N exposure years
Bystander															
Prob ≥0	1.6 (0.007-3.0)	1.5 (0.30-6.2)	79	51	617	1.0 (0.002-1.0)	2.2 (0.43-7.3)	80	52	627	1.0 (0.002-1.0)	1.5 (0.31-5.2)	80	52	624
Prob 0.5	0.20 (0.007-3.0)	1.8 (1.0-6.2)	3	3	50	1.0 (1.0-1.0)	1.4 (1.4-1.4)	1	1	6	2.0 (1.0-1.0)	1.0 (1.0-1.0)	1	1	6
Electrician															
Prob ≥0	0.50 (0.28-12)	0.82 (0.30-2.8)	22	14	145	0.64 (0.30-10)	1.2 (0.43-7.3)	27	16	169	0.26 (0.13-10)	0.83 (0.31-5.2)	22	14	133
Prob 0.5	1.5 (0.25-2.5)	5.1 (2.8-5.1)	5	4	24	--					0.13 (0.13-0.26)	0.58 (0.31-0.58)	3	3	30
Farmer															
Prob ≥0	0.5 (0.003-4.2)	0.82 (0.30-2.8)	133	110	1908	0.50 (0.003-4.2)	1.2 (0.43-4.0)	13	110	1908	0.00 (0.0-0.0)	0.83 (0.31-2.8)	12	101	1776
Prob 0.5	--					--									
Machinist															
Prob ≥0	20 (0.30-84)	1.5 (0.27-6.2)	167	143	1140	20 (0.02-84)	2.1 (0.38-8.9)	85	68	413	15 (2.0-84)	1.4 (0.27-6.4)	80	68	371
Prob 0.5	20 (0.30-84)	2.5 (0.28-6.2)	125	109	886	20 (0.02-84)	2.1 (0.40-8.9)	78	61	740	20 (2.0-84)	0.95 (0.27-5.1)	33	19	334
Mechanic															
Prob ≥0	0.50 (0.01-32)	1.1 (0.27-5.4)	149	116	1295	1.2 (0.01-32)	1.5 (0.38-7.7)	15	119	1352	10 (0.003-16)	1.0 (0.27-5.5)	13	105	1254
Prob 0.5	1.2 (0.04-32)	0.82 (0.30-5.1)	14	13	209	1.2 (0.04-32)	1.2 (0.70-6.5)	11	10	156	1.3 (0.003-6.4)	0.58 (0.31-0.58)	6	6	54
Miscellaneous															
Prob ≥0	4.2 (0.30-12)	0.82 (0.30-5.1)	8	8	112	0.70 (0.10-4.0)	1.5 (0.43-7.3)	9	9	139	10 (0.26-10)	0.84 (0.31-5.2)	11	11	144
Prob 0.5	8.0 (0.30-12)	0.93 (0.49-4.5)	4	4	86	1.0 (0.10-4.0)	3.8 (2.2-7.3)	4	4	51	--				
Plumber															
Prob ≥0	12 (0.09-15)	0.90 (0.30-5.1)	26	19	203	11 (0.15-24)	1.2 (0.43-7.3)	34	21	254	7.7 (0.05-10)	0.83 (0.31-5.2)	31	20	224
Prob 0.5	8.4 (0.09-12)	2.8 (2.5-5.1)	9	6	51	10 (10-20)	1.2 (0.79-7.3)	3	3	23	7.7 (1.6-7.7)	0.95 (0.50-1.3)	4	4	26
Production															
Prob ≥0	1.2 (0.00-40)	1.5 (0.27-5.4)	3	3	67	1.0 (0.005-2.0)	1.5 (0.38-7.3)	6	6	95	0.64 (0.0-10)	1.0 (0.27-5.2)	5	5	86
Prob 0.5	1.2 (0.00-40)	2.8 (1.3-5.4)	7	7	42	1.0 (1.0-2.0)	1.9 (1.5-4.0)	2	2	4	0.0 (0.0-10)	0.58 (0.27-0.58)	2	2	9
Sheet metal worker															
Prob ≥0	1.0 (1.0-1.0)	2.8 (0.73-5.1)	4	4	14	1.0 (0.30-5.0)	7.3 (1.0-8.9)	4	4	14	3.9 (1.9-10)	5.2 (0.74-6.4)	8	8	34

Job Group by probability level	Soluble						Straight						Synthetic					
	Frequency, Median (Range)	Intensity, Median (Range)	N jobs	N subjects	N exposure years	Frequency, Median (Range)	Intensity, Median (Range)	N jobs	N subjects	N exposure years	Frequency, Median (Range)	Intensity, Median (Range)	N jobs	N subjects	N exposure years	Frequency, Median (Range)	Intensity, Median (Range)	N jobs
Prob 0.5	1.0 (1.0-1.0)	2.8 (0.73-5.1)	3	3	12	1.0 (0.30-5.0)	7.3 (1.0-8.9)	5	5	30	--	--	--	--	--	--	--	--
Tool & die																		
Prob >0	8.0 (0.20-48)	4.7 (0.49-5.1)	13	10	71	22 (0.64-40)	6.7 (0.70-7.3)	14	11	73	10 (0.26-22)	2.8 (0.50-5.2)	25	16	150			
Prob 0.5	7.0 (0.20-48)	4.7 (0.75-5.0)	10	7	55	23 (12-40)	6.7 (1.1-7.1)	10	6	60	--	--						
Welder																		
Prob >0	1.0 (1.0-5.0)	0.70 (0.30-5.1)	24	17	65	1.0 (0.005-4, 0)	1.2 (0.43-7.3)	25	17	68	2.6 (0.64-14)	0.83 (0.31-5.2)	28	19	123			
Prob 0.5	1.0 (1.0-5.0)	0.57 (0.30-2.8)	6	4	85	1.2 (0.005-4.0)	0.82 (0.43-4.0)	8	6	101	2.6 (1.9-2.6)	0.58 (0.31-0.58)	2	2	26			
All jobs ^a																		
Prob >0	1.0 (0.00-84)	1.2 (0.27-6.2)	727	462	6683	1.2 (0.001-84)	1.8 (0.38-8.9)	73	469	6748	1.29 (0.0-84)	1.0 (0.27-6.4)	63	409	5780			
Prob 0.5	10 (0.00-84)	2.5 (0.28-6.2)	197	155	1667	8.0 (0.005-84)	2.1 (0.40-8.9)	12	101	1340	10 (0.0-84)	0.91 (0.27-5.1)	56	42	593			

^aThe sum the number of jobs, subjects, and person-years from the 11 job groups does not necessarily equal the numbers listed for all jobs because jobs that belonged to multiple job groups (i.e., both electrician and mechanic) are included only in the 'all job' category and not the individual job groups.