UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Occupational Exposure to Metalworking Fluids and Incidence of Cancer in the United Autoworkers-General Motors Cohort: Assessing and Accounting for the Healthy Worker Survivor Effect

Permalink

https://escholarship.org/uc/item/7bf5w8q6

Author

Garcia, Erika

Publication Date

2017

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Occupational Exposure to Metalworking Fluids and Incidence of Cancer in the United Autoworkers-General Motors Cohort: Assessing and Accounting for the Healthy Worker Survivor Effect

By

Erika Garcia

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Environmental Health Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ellen A. Eisen, Chair
Professor John R. Balmes
Professor Patrick T. Bradshaw

Spring 2017

Occupational Exposure to Metalworking Fluids and Incidence of Cancer in the United Autoworkers-General Motors Cohort: Assessing and Accounting for the Healthy Worker Survivor Effect

Copyright 2017
by
Erika Garcia

Abstract

Occupational Exposure to Metalworking Fluids and Incidence of Cancer in the United Autoworkers-General Motors Cohort: Assessing and Accounting for the Healthy Worker Survivor Effect

By

Erika Garcia

Doctor of Philosophy in Environmental Health Sciences

University of California, Berkeley

Professor Ellen A. Eisen, Chair

Metalworking fluids (MWF) are coolants and lubricants used in industrial machining and grinding operations and have been linked with several cancers. Components found in MWFs, including polycyclic aromatic hydrocarbons found in oil-based straight fluids and nitrosamines found in water-based synthetic fluids, have been linked with mammary gland as well as respiratory organ tumors in laboratory animals. Epidemiologic evidence for these cancers, however, is less conclusive. Despite the provocative toxicologic evidence, there are few epidemiologic studies on MWF exposure and breast cancer. Results for lung cancer have been null for oil-based MWF and appeared inversely related to synthetic fluids, possible due to presence of endotoxin. The United Autoworkers-General Motors (UAW-GM) study includes an occupational cohort of 46,316 hourly workers in automotive manufacturing and extensive MWF exposure data, with annual average exposure estimates available for workers' full employment history, and over two decades of cancer incidence follow-up. To better understand the exposure-response relationship between MWF exposure and these cancer outcomes, this dissertation set out to improve upon prior studies by first evaluating the presence of a common source of bias in occupational epidemiology, the healthy worker survivor effect (HWSE) in the UAW-GM cohort, and second collecting additional data and using advanced methods to evaluate these relationships.

Chapter 1 is the assessment of the presence of the HWSE in cancer studies of the UAW-GM cohort. The HWSE can affect the validity of occupational studies when data are analyzed incorrectly. HWSE depends on three underlying conditions: (1) leaving work predicts future exposure, (2) leaving work is associated with disease outcome, and (3) prior exposure increases probability of leaving work. If all these conditions are satisfied employment status is a time-varying confounder affected by prior exposure and standard methods will produce bias. I evaluated the presence of these conditions for select cancer outcomes, including lung cancer, in the UAW-GM cohort and found evidence for all three conditions. This suggested that standard methods may underestimate the exposure-response for lung cancer and therefore a g-method should be applied to control for

employment status as a time-varying confounder affected by prior exposure. A secondary analysis examining breast cancer among female workers found insufficient evidence for condition (3), indicating that standard methods are appropriate for this outcome and will not produce bias due to HWSE.

In Chapter 2, I evaluated the exposure-response relation between cumulative MWF exposure and breast cancer incidence among female workers in the UAW-GM cohort. Additional data was obtained by extending follow-up four more years for female cohort members using data linkage with the Michigan Cancer Registry and the National Death Index. We identified 221 total incident breast cancer cases among 4,503 female workers. Risks associated with exposure to the three types of MWFs, straight, soluble, and synthetic, were evaluated using Cox proportional hazards models for all breast cancer cases as well as for pre-menopausal cases defined by age at diagnosis. Results suggest an increasing exposure-response curve for straight fluids and breast cancer. While the number of the pre-menopausal cancers is small, results are modestly suggestive of an increased risk associated with higher synthetic fluid exposure, suggesting a different mechanism for the younger cases.

In Chapter 3, I examined the relationship between lung cancer mortality and exposure to straight and synthetic MWF, as well as to the biocides that are added to water-based fluids to control microbial growth. Using the parametric g-formula, ratios were estimated comparing cumulative risk of lung cancer mortality under the hypothetical interventions always high exposure while at work and always unexposed. We also intervened on both synthetic MWF and biocides simultaneously to estimate independent effects. Results from this study suggest slightly elevated lung cancer mortality related to straight MWF exposure, albeit with wide confidence intervals. Our results do not support a negative association for synthetic fluids, as reported in earlier studies; instead, biocide in the fluid, a marker for the release of endotoxin, was associated with decreased lung cancer.

The hypotheses addressed in this dissertation are of public health importance in light of the extremely high incidence of both breast and lung cancer in the US. These studies provide information on the relationships between MWF exposure, breast and lung cancer incidence, adding to the scientific literature that informs regulatory measures and protects workers from the risks of MWF exposure.

To my mom, Rebeca, who always encouraged me to pursue education,
but never liked that I moved so very far away to Berkeley.
I'm from Los Angeles.

&

To my partner, Jack, who gave me a great deal of support throughout graduate school.

Table of Contents

List of Figures ..... **iv**

List of Tables ..... **v**

Acknowledgments ..... **vi**

Chapter 1 Introduction **1**

1.1 Healthy Worker Survivor Effect 1

1.2 Examination of the Healthy Worker Survivor Effect in the UAW-GM Cohort 2

1.3 Investigation of Metalworking Fluid Exposure and Breast and Lung Cancer in the UAW-GM cohort 2

1.4 Public Health Significance 3

1.5 References 3

Chapter 2 Assessment of the Healthy Worker Survivor Effect in Cancer Studies of the United Autoworkers-General Motors Cohort **6**

2.1 Abstract 6

2.2 Introduction 6

2.3 Methods 8

2.4 Results 10

2.5 Discussion 14

2.6 Conclusion 17

2.7 References 17

Chapter 3 Breast Cancer Incidence and Metalworking Fluid Exposure in a Cohort of Female Autoworkers **20**

3.1 Abstract 20

3.2 Introduction 20

3.3 Methods 21

3.4 Results 24

3.5 Discussion 28

3.6 Conclusion 30

3.7 References 30

Chapter 4 Metalworking Fluids Exposure and Lung Cancer Mortality; Controlling for the Healthy Worker Survivor Effect **35**

4.1 Abstract 35
4.2 Introduction 35
4.3 Methods 36
4.4 Results 39
4.5 Discussion 44
4.6 Conclusion 45
4.7 References 46
4.8 Appendix 49
Chapter 5 Conclusions 52
5.1 Summary of Findings 52
5.2 Conclusions and Future Directions 53
5.3 References 54
Dissertation Publications 55

List of Figures

Figure 1.1 Visual depiction of the Healthy Worker Survivor Effect……..1

Figure 2.1 Directed acyclic graph depicting the associations underlying healthy worker survivor effect……..6

Figure 2.2 Condition 2: Leaving working and cancer incidence……..12

Figure 3.1 Number of active workers and annual exposure prevalence to the three types of metalworking fluids by year……24

Figure 3.2 Adjusted hazard ratios for premenopausal breast cancer incidence as a smoothed function of 20-year lagged cumulative synthetic metalworking fluid……..27

Figure 4.1 Number of workers under follow-up and number of lung cancer deaths by age among 38,560 autoworkers……39

Figure 4.2 Observed and simulated cumulative lung cancer mortality in the UAW-GM cohort, under the natural course……41

Figure 4.3 Simulated cumulative lung cancer mortality in the UAW-GM cohort under two straight MWF interventions……42

Figure 4.4 Simulated cumulative lung cancer mortality in the UAW-GM cohort under six synthetic MWF interventions……..42

List of Tables

Table 2.1 Demographic and exposure characteristics of 31,485 autoworkers in the United Autoworker-General Motors incidence cohort……..10

Table 2.2 Demographic characteristics of incident cancer cases diagnosed 1985-1994 among autoworker……..11

Table 2.3 Condition 3: Cumulative exposure to metalworking fluids and leaving work among male and female autoworkers……..13

Table 3.1 Demographic and exposure characteristics of female breast cancer cases and cohort members……23

Table 3.2 Adjusted hazard ratios of breast cancer incidence (1985-2013) in relation to cumulative exposure to metalworking fluid……..25

Table 3.3 Adjusted hazard ratios of premenopausal breast cancer incidence (1985-2013) in relation to cumulative exposure to metalworking fluid……..26

Table 4.1 Exposure interventions simulated in the UAW-GM cohort study……..37

Table 4.2 Demographic characteristics of the UAW-GM cohort at baseline……..39

Table 4.3 Exposure characteristics over person-years in the observed data and simulated interventions a in the UAW-GM cohort……..40

Table 4.4 Cumulative lung cancer mortality by select ages under several simulated interventions a in the UAW-GM cohort……..43

Acknowledgments

I would first like to acknowledge the support of my wonderful advisor, Professor Ellen Eisen. I thank her for hiring me many years ago as a graduate student researcher during my first semester in the masters program and for later encouraging me to return to Berkeley for my PhD. Her mentorship, guidance, and wisdom these past years have been invaluable to my training as an occupational epidemiologist. I am very grateful for her time and dedication to my education.

I would also like to thank the other members of my dissertation committee, Professors John Balmes and Patrick Bradshaw, for their insightful feedback on my research, as well as the collaborators who shared their knowledge and worked with me to complete the studies presented here: Doctors Sally Picciotto, Sadie Costello, and Andreas Neophytou. I also thank the other members of the Eisen research group, particularly my fellow student Monika Izano for her constant comradery and support, Liza Lutzker for her fantastic and tireless work with the data linkage, and Dan Brown for providing his insights, often while heating lunch in the kitchen.

Thank you to the many others who have contributed to my training. Professor Michael Bates and Doctor Amod Pokhrel, as well as Doctor Costello mentioned above, thank you for your contribution to and guidance on the short-lived Nepalese brick kiln and air pollution project. I thank the Center for Occupational and Environmental Health for their continual support and acknowledge that my support was possible through a National Institute of Occupational Safety and Health Targeted Research Training grant. I also thank Gayle Cepparo for her constant administrative support. And I thank the pillar of the EHS Division, Norma Firestone, for her many years of advice, encouragement, and help, and for taking the time to answer my questions even though she was quite busy.

I thank my friends in the EHS Division as well as other departments on campus. A good community can make a great difference during graduate school and I thank you for being mine. I especially want to thank my good friends Jessica An for keeping me accountable and meeting with me on a weekly basis for the last two years and Brooke Rhead for consistently working on Sundays with me and helping to keep me grounded and healthy.

Finally, I thank my partner, family, and friends for their love and support throughout the years. I wholeheartedly thank Jack Drale for his enduring support and constant belief in me throughout these many years. I thank my mom Rebeca and siblings Alvaro, Claudia, and Livier for their love. And I thank Sophia Horiuchi for helping me keep a good work-life balance throughout graduate school.

Chapter 1 Introduction

1.1 Healthy Worker Survivor Effect

Focusing on occupational populations for epidemiologic studies provides a number of advantages. These populations are usually well-defined, making it easier to enumerate a study population. Exposure levels in the workplace are usually higher than in the community, therefore studies will have greater statistical power to detect health risks associated with the exposure. The availability of employment and exposure monitoring–or job descriptions–records can be leveraged to more accurately estimate individual exposure levels. Since we are relying on observational data, however, we must also bear in mind potential sources of bias when using workers as study participants. A unique source of bias often of concern in occupational epidemiology is the Healthy Worker Survivor Effect (HWSE).[1-4] The HSWE is due to less healthy workers preferentially leaving the workforce, which results in a survivor worker population remaining at work. When examining the health risks of cumulative exposure, the HWSE can bias measures of association downward because subject who leave work accrue less exposure compared with subjects who "survive" and remain at work.[3]

Figure 1.1 is a visual demonstrating the HWSE. The rectangular color bands indicate the distribution of health status among hypothetical workers at the particular exposure level and time. Red indicates lower health, yellow moderate health, and green higher health. Because workers who stay in the workforce are more likely to be healthier (i.e., green in this example), over time the distribution of underlying health status among active workers shifts to be more healthy (i.e., more green). This is the HWSE. And it is expected to be stronger in jobs with higher exposure. As seen in the figure, while the color band shifts towards green for all exposure levels, it does so more dramatically for higher exposure levels. Because underlying health status is not typically available in occupational studies, employment status can be used as a proxy, since less healthy workers would be more likely to leave employment.

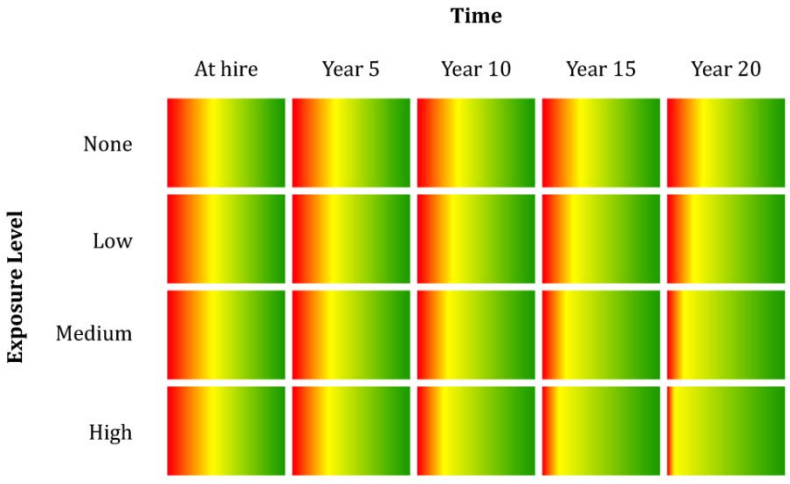


Figure 1.1 Visual depiction of the Healthy Worker Survivor Effect. The color band communicates the distribution of underlying health status of active workers: red indicates lower health, yellow moderate health, and green higher health.

1.2 Examination of the Healthy Worker Survivor Effect in the UAW-GM Cohort

In chapter 2, I used employment termination status as a proxy for underlying health status and evaluated the presence of the Healthy Worker Survivor Effect in the United Autoworkers-General Motors (UAW-GM) cohort. The UAW-GM cohort study was a joint labor-management funded study designed in the mid-1980s to examine cancer mortality and its relation to metalworking fluid (MWF) exposure, which are coolants and lubricants used in machining and grinding operations.[5,6] I used the results from chapter 2 assessment to guide my statistical approach for studies of breast (chapter 3) and lung cancer (chapter 4) in the UAW-GM cohort. In the presence of HWSE standard methods, such as covariate adjustment in Poisson or Cox proportional hazards regression, will likely produce biased results because these methods are unable to properly adjust for a time-varying confounder affected by prior exposure.[2,7] Because the HWSE was not evident for female workers in the UAW-GM for breast cancer, Cox proportional hazards models were used to evaluate the exposure-response relation for MWF and breast cancer incidence. I used a more advanced g-method, the parametric g-formula,[7,8] to assess the relationship between MWF exposure and lung cancer mortality, because I did detect the HWSE in the UAW-GM cohort for this outcome.

1.3 Investigation of Metalworking Fluid Exposure and Breast and Lung Cancer in the UAW-GM cohort

MWF are cooling and lubricating fluids used widely in industrial machining and grinding operations. These fluids are used to protect or treat the surfaces of the material being processed, carry away metal chips, and prolong the life of the machine tools.[9] Workers may be exposed via inhalation due to the aerosolization of MWF during use and splashing of fluids during application. MWF are classified into three categories based on composition: straight, soluble, and synthetic. Straight MWFs are complex mixtures of paraffinic, naphthenic, and aromatic compounds refined from mineral oil.[10] Mineral oil is classified as carcinogenic to humans[11] and its carcinogenic properties are thought to be primarily due to its polycyclic aromatic hydrocarbon (PAH) content.[12] Soluble MWFs are composed of oils emulsified in water. Synthetic fluids are water-soluble chemical lubricants without oil. MWF have been linked with several cancers, including skin,[6,13-16] laryngeal,[5,6,17-20] rectal,[5,6,17,21,22] pancreatic,[6,23,24] esophageal,[6,24,25] and bladder.[26-29]

Chapter 3 is an evaluation of the exposure-response relation between MWF exposure and breast cancer among 4,503 female autoworkers. Although breast cancer is the leading cancer diagnosed among women in the United States,[30] there are few known modifiable risk factors.[31] A 2012 Institute of Medicine review of environmental risk factors for breast cancer noted that the initial identifications of many known human carcinogens were based on studies of high exposures in occupational settings and recommended additional breast cancer studies of worker populations.[31] Following their recommendation I designed and carried out this study. Cancer incidence follow up was extended and I carried out the analyses based on 221 incident female breast cancer cases. Pooled breast cancer as well as premenopausal breast cancer, defined by age at diagnosed, were examined.

Chapter 4 employs the parametric g-formula to evaluate lung cancer mortality under multiple hypothetical interventions on straight and synthetic fluids as well as biocides, which are added to water-based fluids to control microbial growth. Based on observed data on 38,560 workers, including 873 cases of lung cancer deaths, I estimated ratios comparing cumulative risk of lung cancer mortality under the hypothetical interventions always high exposure while at work and always unexposed.

1.4 Public Health Significance

Occupational carcinogens are very relevant to public health and contribute to the global burden of disease. Occupational carcinogens account for almost half of all identified human carcinogens[32] and in 2015 there were an estimated 489 thousand deaths and 9.8 million disability-adjusted life years (DALYs) attributable to occupational carcinogens.[33] Based on a comparative risk assessment, 5% of all cancer DALYs and 25% of tracheal, bronchial, and lung cancer DALYs are estimated to be attributable to occupational carcinogens.[33] Taking advantage of the rich longitudinal data available in UAW-GM cohort, I investigated occupational exposure to MWF and incidence of breast and lung cancer. The hypotheses addressed are of public health importance in light of the extremely high incidence of both breast and lung cancer in the US. The studies described in this dissertation provide information on the relationship between MWF exposure and breast and lung cancer risk. Better specification of the exposure-response curves for MWF will improve the basis for setting protective health standards for workers.

1.5 References

1. Arrighi HM, Hertz-Picciotto I. The evolving concept of the healthy worker survivor effect. *Epidemiology* 1994;5(2):189-96.
2. Buckley JP, Keil AP, McGrath LJ, Edwards JK. Evolving methods for inference in the presence of healthy worker survivor bias. *Epidemiology* 2015;26(2):204-12.
3. Picciotto S, Hertz-Picciotto I. Commentary: healthy worker survivor bias: a still-evolving concept. *Epidemiology* 2015;26(2):213-5.
4. Eisen EA, Picciotto S, Robins JM. Healthy Worker Effect. *Encyclopedia of Environmetrics* John Wiley & Sons, Ltd, 2006.
5. Eisen EA, Tolbert PE, Monson RR, Smith TJ. Mortality studies of machining fluid exposure in the automobile industry I: A standardized mortality ratio analysis. *Am J Ind Med* 1992;22(6):809-24.
6. Eisen EA, Bardin J, Gore R, Woskie SR, Hallock MF, Monson RR. Exposure-response models based on extended follow-up of a cohort mortality study in the automobile industry. *Scand J Work Environ Health* 2001;27(4):240-9.
7. Robins J. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. *J Chronic Dis* 1987;40 Suppl 2:139S-161S.
8. Robins J. A New Approach to Causal Inference in Mortality Studies with a Sustained Exposure Period - Application to Control of the Healthy Worker Survivor Effect. *Mathematical Modelling* 1986;7(9-12):1393-1512.

9. NIOSH. Criteria for a Recommended Standard: Occupational Exposure to Metalworking Fluids. Vol. DHHS Publication No. 98-102. Cincinnati, OH: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health,, 1998.
10. Neale M. *Tribiology Handbook*. New York: John Wiley and Sons, 1973.
11. International Agency for Research on Cancer (IARC). Mineral Oils, Untreated or Mildly Treated. *Monographs, Chemical Agents and Related Occupations, vol 100F*. Vol. 100F. Lyon, France: IARC, 2012.
12. Catchpole WM, Macmillan E, Powell H. Specifications for cutting oils with special reference to carcinogenicity. *Ann Occup Hyg* 1971;14(2):171-9.
13. Jarvholm B, Lavenius B. Mortality and cancer morbidity in workers exposed to cutting fluids. *Arch Environ Health* 1987;42(6):361-6.
14. Jarvholm B, Fast K, Lavenius B, Tomsic P. Exposure to cutting oils and its relation to skin tumors and premalignant skin lesions on the hands and forearms. *Scand J Work Environ Health* 1985;11(5):365-9.
15. Roush GC, Kelly JA, Meigs JW, Flannery JT. Scrotal carcinoma in Connecticut metalworkers: sequel to a study of sinonasal cancer. *Am J Epidemiol* 1982;116(1):76-85.
16. Costello S, Friesen MC, Christiani DC, Eisen EA. Metalworking fluids and malignant melanoma in autoworkers. *Epidemiology* 2011;22(1):90-7.
17. Tolbert PE, Eisen EA, Pothier LJ, Monson RR, Hallock MF, Smith TJ. Mortality studies of machining-fluid exposure in the automobile industry. II. Risks associated with specific fluid types. *Scand J Work Environ Health* 1992;18(6):351-60.
18. Eisen EA, Tolbert PE, Hallock MF, Monson RR, Smith TJ, Woskie SR. Mortality studies of machining fluid exposure in the automobile industry. III: A case-control study of larynx cancer. *Am J Ind Med* 1994;26(2):185-202.
19. Zagraniski RT, Kelsey JL, Walter SD. Occupational risk factors for laryngeal carcinoma: Connecticut, 1975-1980. *Am J Epidemiol* 1986;124(1):67-76.
20. Wortley P, Vaughan TL, Davis S, Morgan MS, Thomas DB. A case-control study of occupational risk factors for laryngeal cancer. *Br J Ind Med* 1992;49(12):837-44.
21. Decoufle P. Further analysis of cancer mortality patterns among workers exposed to cutting oil mists. *J Natl Cancer Inst* 1978;61(4):1025-30.
22. Gerhardsson de Verdier M, Plato N, Steineck G, Peters JM. Occupational exposures and cancer of the colon and rectum. *Am J Ind Med* 1992;22(3):291-303.
23. Bardin JA, Eisen EA, Tolbert PE, Hallock MF, Hammond SK, Woskie SR, Smith TJ, Monson RR. Mortality studies of machining fluid exposure in the automobile industry. V: A case-control study of pancreatic cancer. *Am J Ind Med* 1997;32(3):240-7.
24. Silverstein M, Park R, Marmor M, Maizlish N, Mirer F. Mortality among bearing plant workers exposed to metalworking fluids and abrasives. *J Occup Med* 1988;30(9):706-14.
25. Sullivan PA, Eisen EA, Woskie SR, Kriebel D, Wegman DH, Hallock MF, Hammond SK, Tolbert PE, Smith TJ, Monson RR. Mortality studies of metalworking fluid exposure in the automobile industry: VI. A case-control study of esophageal cancer. *Am J Ind Med* 1998;34(1):36-48.

26. 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(12):1471-8.
27. Park RM, Mirer FE. A survey of mortality at two automotive engine manufacturing plants. *Am J Ind Med* 1996;30(6):664-73.
28. Schifflers E, Jamart J, Renard V. Tobacco and occupation as risk factors in bladder cancer: a case-control study in southern Belgium. *Int J Cancer* 1987;39(3):287-92.
29. 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(4):701-13.
30. U.S. Cancer Statistics Working Group. United States Cancer Statistics: 1999–2011 Incidence and Mortality Web-based Report. Atlanta, GA: Department of Health and Human Services, Center for Disease Control and Prevention, and National Cancer Institute, 2014.
31. Committee on Breast Cancer and the Environment. Breast Cancer and the Environment: A Life Course Approach. Insititute of Medicine, 2012;468.
32. Siemiatycki J, Richardson L, Boffetta P. Occupation. In: Schottenfeld D, Fraumeni JF, eds. *Cancer epidemiology and prevention*. 3rd ed. New York ; Oxford: Oxford University Press,, 2006;323-354.
33. Collaborators GBDRF. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;388(10053):1659-1724.

Chapter 2 Assessment of the Healthy Worker Survivor Effect in Cancer Studies of the United Autoworkers-General Motors Cohort

2.1 Abstract

The healthy worker survivor effect (HWSE) can affect the validity of occupational studies when data are analyzed incorrectly. HWSE depends on three underlying conditions: (1) leaving work predicts future exposure, (2) leaving work is associated with disease outcome, and (3) prior exposure increases probability of leaving work. If all these conditions are satisfied, then employment status is a time-varying confounder affected by prior exposure, and standard regression will produce bias. We assessed these conditions for cancer outcomes in a cohort of autoworkers exposed to metalworking fluids (MWF). The cohort includes 31,485 workers followed for cancer incidence from 1985-1994. Since occupational exposures to straight, soluble, and synthetic MWF are necessarily zero after leaving work, condition (1) is satisfied. Cox models for cancer incidence and for employment termination were used to assess conditions (2) and (3), respectively. Employment termination by select ages was examined to better gauge the presence of condition (2). The hazard ratio for leaving work as a predictor of all cancers combined and prostate cancer was null, but elevated for lung and colorectal cancers among men. Condition (2) was more clearly satisfied for all cancer outcomes when leaving work occurred younger. Higher exposures to all three MWF types were associated with increased rates of leaving work [condition (3)], with the exception of straight MWF among women. We found evidence for the structural conditions underlying HWSE in a cohort of autoworkers. G-methods should be applied to reduce HWSE bias in studies of all cancers presently examined.

2.2 Introduction

The healthy worker survivor effect (HWSE) can affect the validity of results from occupational health studies.[1,2] It has been extensively studied [1-6] and we describe it only briefly here. The HWSE arises due to less healthy workers preferentially leaving the workforce, resulting in workers who remain at work (i.e., survive) being healthier than those who have left. A directed acyclic graph (DAG) is a mathematical tool widely used in epidemiology to address problems of causal inference.[7] Introduced into epidemiology more

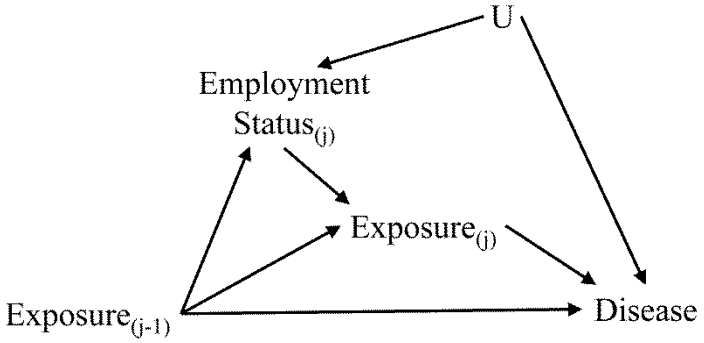


Figure 2.1 Directed acyclic graph depicting the associations underlying healthy worker survivor effect.

than a decade ago, these graphs explicitly present assumptions about the temporal relations between the exposures, covariates and the outcome, allowing us to identify variables that need to be controlled in order to estimate results unbiased by confounding.[8] In Figure 2.1 we show a DAG depicting the structure that gives rise to the HWSE. Unobserved underlying health status (U in Figure 2.1) may influence both leaving work (employment status at time *j*) and risk of the disease of interest. Leaving work may be affected by prior exposure (exposure at time *j-1*) and also directly affects subsequent exposure (exposure at time *j*) as occupational exposure will be zero after leaving work. In studies of cumulative exposure, employees who leave work accrue less exposure than their counterparts who remain on the job, causing bias induced by the HWSE.[3]

As initially described by Robins, the HWSE poses a challenge to conventional methods of analysis because employment termination status is a time-varying confounder affected by prior exposure.[6] Employment status is associated with the disease and predicts future exposure (making it a confounder of future exposure and disease outcome) and it is also predicted by past exposure (making it an intermediate between past exposure and disease outcome). This poses a dilemma whereby we both need to control for employment status as a confounder and need to *not* control for it since it is on the causal pathway between exposure and disease outcome. Furthermore, adjusting for employment status would induce collider-stratification bias by opening a confounding pathway as both past exposure and underlying health status are parents of employment status.[9]

The healthy worker survivor *effect* is sometimes referred to as healthy worker survivor *bias.*[2] If we define the effect as the data structure described above (depicted in Figure 2.1), however, then we can distinguish it from the bias. Bias from this source occurs only if inappropriate analytical methods are applied. When standard methods (e.g., covariate adjustment in Poisson or Cox proportional hazards regression) are applied where the HWSE is present, results will likely be biased.[2,6] More advanced g-methods, including the g-null test,[10] g-estimation of structural nested models,[11] the parametric g-formula,[6,10] and inverse-probability-weighted marginal structural models,[12] have been developed to control for this type of bias.

To guide our analytical approach in a cohort study of cancer incidence and metalworking fluids (MWF) exposure, we first assessed the potential for bias due to the HWSE. We regard evidence that employment status is a time-varying confounder affected by prior exposure as an indicator of the potential bias. Thus, the presence of the HWSE depends on satisfying three underlying conditions:

(1) leaving work predicts future exposure,
(2) leaving work is associated with disease outcome, and
(3) prior exposure increases probability of leaving work.

Since occupational exposure is necessarily zero after leaving work, (1) is inherently satisfied. In this study we assessed the presence of conditions (2) and (3), which would indicate that HWSE threatens the validity of traditional regression results in epidemiologic

studies of Michigan autoworkers exposed to MWF. This assessment of the presence of the HWSE will guide the choice of methodologic approach in future cancer incidence studies.

2.3 Methods

Study Population

The United Autoworkers-General Motors (UAW-GM) study is an occupational cohort study designed in the mid-1980s to examine cancer mortality and its relation to MWF exposure. The original cohort includes over 46,000 hourly workers employed for at least three years in one of three automobile manufacturing plants in Michigan, USA and hired before January 1, 1985.[13] The study population was further restricted to subjects alive and aged 75 years or younger on January 1, 1985, the year follow-up of cancer incidence began via the Michigan Cancer Registry. Since almost all subjects have left work by age 75 the age restriction was added both as an eligibility criterion and for administrative censoring to ensure adequate data coverage when examining employment status. Subjects missing more than half of their work history information were excluded from the analyses (3.1%). The final study population comprised 31,485 autoworkers. Information on date of birth, year of hire, race, sex, and plant were ascertained from workers' employment records.

Exposure Assessment

MWFs are coolants and lubricants used in industrial machining and grinding operations and are categorized into three classes based on composition: straight (oil-based), soluble (oils emulsified in water), and synthetic (water-soluble chemical lubricants without oil). Quantitative levels of exposure to each MWF class were assigned to subjects based on detailed employment records and a time-varying job-exposure-matrix (JEM).[14,15] The JEM was created by estimating MWF concentrations as an 8-hour time-weighted average (mg/m^3) based on 475 full-shift personal air samples collected in the mid-1980s in major exposure groups in the three manufacturing plants. A set of multipliers (scale factors) were developed to adjust MWF concentration for temporal trends based on 394 historical air sampling measurements, review of historical records, and interviews with plant personnel. Scale factors were updated in 1995 after revisiting the three plants. The JEM was combined with employment records to estimate time-weighted annual exposure to each of straight, soluble, and synthetic MWFs (mg/m^3). For each worker, cumulative exposures to straight, soluble, and synthetic MWFs (mg/m^3-years) were estimated by summing over annual exposures during employment.

Case ascertainment

The UAW-GM cancer incidence cohort was linked with the Michigan Cancer Registry to obtain incident cancer cases diagnosed between January 1, 1985 and December 31, 1994. Michigan cancer data are collected by the Michigan Department of Community Health as part of the Michigan Cancer Surveillance Program,[16] which participates in the National Program of Cancer Registries of the Centers for Disease Control and Prevention.[17] The outcomes of interest for this study are first diagnosis of any (International Classification of

Disease for Oncology Third Edition codes C00.0-C80.9), prostate (C61.0-C61.9), lung (C34.0-C34.9), and colorectal (C18.0-C18.9, C19.9, C20.9, C21.1-C21.2, C21.8) cancers. The site-specific cancers were selected based on having at least 100 identified incident cases during the follow-up period. Vital status was obtained through linkage with the National Death Index (National Center for Health Statistics, Hyattsville, Maryland).

Statistical Methods

We assessed the presence of two conditions necessary for the HWSE by estimating hazard ratios using Cox proportional hazards models to predict cancer incidence and employment termination, following an approach similar to Naimi et al.[18] For all models, we used age as the time metric. All models were additionally adjusted for race, plant, and calendar year (B-spline with 3 df and equally spaced knots). Separate models were fit for male and female workers due to temporal differences in hiring and exposure patterns.

For condition (2), we modeled cancer outcome as a function of employment termination (=1 if subject left work), which was treated as a time-varying indicator variable. Follow-up began January 1, 1985 and ended on either the date of first cancer diagnosis (any or site specific), date of death, 75th birthday, or December 31, 1994 (the last date for which we have employment data), whichever occurred first. Year of cohort entry (B-spline with 3 df and equally spaced knots) and duration of employment were additionally adjusted for in the models. In order to gauge whether condition (2) is present, we must block the backdoor path[8] leading from termination status through prior exposure to cancer incidence in Figure 2.1. To accomplish this we further adjusted the model for previous exposure using cumulative exposure accrued up through the prior year. Because of cancer latency time, we do not believe exposure in the current year would affect cancer diagnosis in that same year. As such, we did not include current exposure in the cancer incidence models. This differs from the methods used by Naimi and colleagues who were examining lung cancer mortality and did adjust for subsequent exposure.[18]

We also investigated whether an age-specific measure of leaving work better distinguishes employment termination for health-related reasons (at younger ages) from normal retirement (at older ages). Our hypothesis was that leaving work at a younger age would be a stronger predictor of future cancer risk than leaving at any age. To explore this hypothesis we examined condition (2) based on leaving work by age 50, 55, and 60. Instead of a single time-varying indicator of employment termination a set of two dummy variables were constructed to capture three levels of employment status: actively employed workers, subjects who left work by the age cut point, and subjects who left work after. This construction allowed the main model with no age cut point and these age specific models to have a common reference group (i.e., actively employed workers).

For assessing condition (3), cumulative MWF exposure was accrued up through the prior year and parameterized as a categorical variable with cut points defined based on the exposure distribution among subjects who had left work. Separate terms for each MWF class were included simultaneously in all the models. For these models we additionally adjusted for year of hire. Follow-up for leaving work began when subjects entered the

cohort (three years after date of hire) and ended on either the date of employment termination or December 31, 1994, whichever occurred first. We also examined cumulative exposure lagged by 15 years in our evaluation of condition (3), because we plan to lag exposure in the ultimate exposure-response models for cancer incidence. This approach is also consistent with an earlier proposal for reducing the HWSE by lagging exposures in order to ignore the more recent exposures most likely to affect leaving work.[5]

SAS software version 9.4 (SAS Institute, Cary, NC) was used for all analyses. All research protocols were approved by the Office for the Protection of Human Subjects at the University of California at Berkeley.

2.4 Results

Characteristics of the 31,485 autoworkers included in the cancer incidence cohort are presented in Table 2.1. These characteristics are shown according to the two different follow-up periods: for cancer incidence of any site in the assessment of condition (2) and

Table 2.1 Demographic and exposure characteristics of 31,485 autoworkers in the United Autoworker-General Motors incidence cohort who were alive in 1985, during follow-up for cancer incidence and follow-up for leaving work. (median (IQR), unless otherwise noted[a])

Characteristic	During cancer follow-up[b] (starting in 1985)	During employment follow-up (starting at date of hire[c])
N subjects	31,485	31,485
Person-years	286,023	504,836
Age	46 (38-58)	38 (31-46)
Sex; n (%)		
Male	247,220 (86.4)	447,576 (88.7)
Female	38,803 (13.6)	57,260 (11.3)
Race; n (%)		
White	231,186 (80.8)	399,730 (79.2)
African American	54,837 (19.2)	105,106 (20.8)
Year of Hire	1968 (1959-1976)	1966 (1953-1973)
Calendar Year	1989 (1987-1992)	1980 (1970-1986)
At work; n (%)	160,242 (56.0)	504,836 (100)
Employment terminations; n	–	20,586
Incident cancer cases; n	1,739	–
MWF exposure[d]		
Any straight; n (%)	24,367 (8.5)	125,287 (24.8)
Straight, mg/m^3	0.09 (0.04-0.25)	0.10 (0.04-0.39)
Cumulative straight, mg/m^3-years	0.64 (0.21-2.33)	0.72 (0.22-2.85)
Any soluble; n (%)	78,148 (27.3)	214,299 (42.4)
Soluble, mg/m^3	0.22 (0.13-0.34)	0.37 (0.18-0.87)
Cumulative soluble, mg/m^3-years	3.77 (1.54-9.54)	4.31 (1.54-11.51)
Any synthetic; n (%)	24,380 (8.5)	83,810 (16.6)
Synthetic, mg/m^3	0.05 (0.02-0.17)	0.04 (0.02-0.17)
Cumulative synthetic, mg/m^3-years	0.38 (0.15-1.33)	0.38 (0.13-1.39)

[a]All *n* statistics presented as number of person-years, expect for employment terminations and incident cancer cases which are presented as number of subjects
[b]Follow-up for first diagnosis of any cancer site
[c]Start of follow-up is date of hire + 3 years
[d]Annual and cumulative exposure levels reported among those exposed

Table 2.2 Demographic characteristics of incident cancer cases diagnosed 1985-1994 among autoworkers in the United Autoworker-General Motors incidence cohort who were alive in 1985. (median (IQR), unless otherwise noted)

	Male Workers				**Female Workers**
	All Cancers Combined	*Prostate Cancer*	*Lung Cancer*	*Colorectal Cancer*	*All Cancers Combined*
N	1,517	384	340	182	222
Year of birth	1926 (1920-1933)	1923 (1919-1929)	1925 (1920-1932)	1925 (1919-1931)	1928 (1920-1942)
Year entered cohort	1956 (1952-1969)	1956 (1951-1966)	1956 (1952-1967)	1956 (1951-1969)	1970 (1955-1979)
Race; n (%)					
White	1178 (77.7)	266 (69.3)	265 (77.9)	144 (79.1)	179 (80.6)
African American	339 (22.4)	118 (30.7)	75 (22.1)	38 (20.9)	43 (19.4)
Age at diagnosis	65.1 (57.3-70.4)	68.2 (63.1-72)	65.5 (57.5-71.5)	65.5 (59.5-69.8)	62.4 (50.1-70)
Deceased; n (%)	1276 (84.1)	286 (74.5)	336 (98.8)	155 (85.2)	161 (72.5)
Survival among deceased, years	1.4 (0.1-6.9)	6.5 (2.6-11.9)	0.2 (0-1.1)	2.7 (0.7-10.2)	1.5 (0.3-8.9)
Termination status; n (%)					
At Work	394 (26)	67 (17.5)	65 (19.1)	39 (21.4)	81 (36.5)
Left Work	1123 (74)	317 (82.6)	275 (80.9)	143 (78.6)	141 (63.5)
Age left work	54.9 (36.4-61.8)	57.4 (43.7-62.3)	52.4 (36.5-60.7)	57.1 (35.7-62)	49.4 (37.5-58.2)

for leaving work in the assessment of condition (3). The cohort is predominately white men, but includes 19% (N=5,972) African American and 13% (N=4,228) female workers. The person-years for both follow-up periods have similar distributions for most covariates. Workers were more likely to be exposed to soluble MWF than either of the two other fluid types. The exposure level among exposed workers was highest for soluble, measured both as annual and cumulative exposure. Male workers were hired earlier than female workers (median (IQR): 1966 (1953-1974) versus 1975 (1967-1977)), at younger ages (23 (19-28) versus 27 (22-34)), and were employed at the company for more years (19 (11-28) versus 17 (12-20)). Annual and cumulative exposures were higher for male than female workers for all fluid types, except for annual synthetic exposure, which was similar between the two groups. Because of the differences in work and exposure patterns among male and female workers, all analyses were stratified by sex.

As seen in Table 2.2, among the 1,739 incident cases of a first cancer diagnosed at any site, female cancer cases were younger at diagnosis and more likely to be actively employed than male cancer cases. For women, no site-specific cancer had more than 100 cases and so none were assessed. For men, we focused on the three most common cancers: prostate, lung, and colorectal. Prostate cancer cases were diagnosed at older ages than the other site-specific cancers, and survival was worst for lung cancer with regard to both case fatality

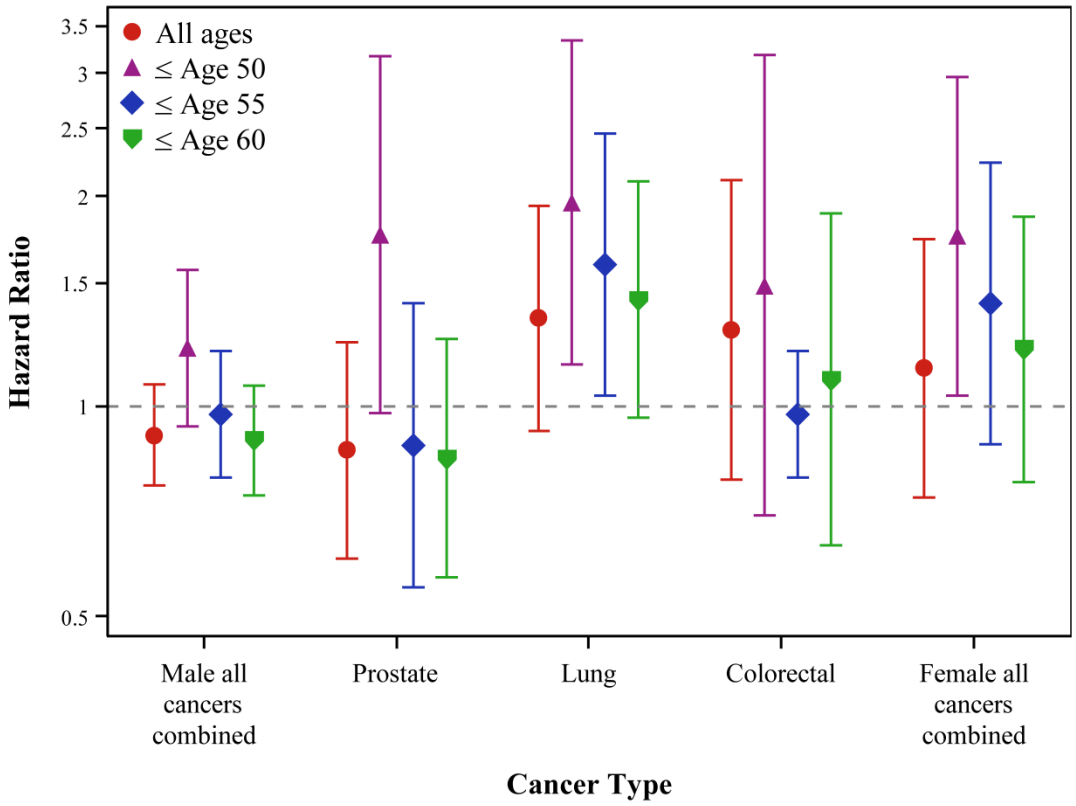


Figure 2.2 Condition 2: Leaving working and cancer incidence. Site-specific cancers examined only among male workers. Hazard ratios with 95% confidence intervals. All modeled as Cox regression with age as time scale and adjusted for race, plant, year of cohort entry, calendar year, duration of employment, and prior exposure

Table 2.3 Condition 3: Cumulative exposure to metalworking fluids (MWF) and leaving work among male and female autoworkers[a]

Male Workers (N= 27,257)				Female Workers (N= 4,228)			
MWF Exposure[b] (mg/m^3-years)	N Left Work	Person-years	Adjusted HR (95% CI)	MWF Exposure[b] (mg/m^3-years)	N Left Work	Person-years	Adjusted HR (95% CI)
Straight				Straight			
0	7979	206,053.5	Ref	0	1126	30,432.5	Ref
0 to 0.38482	3374	85,952.0	1.13 (1.08-1.18)	0 to 0.180	452	9,333.3	1.02 (0.90-1.15)
0.38482 to 2.058	3375	81,198.7	1.05 (1.00-1.10)	0.180 to 0.963	453	10,108.7	0.95 (0.84-1.09)
>2.058	3374	74,371.5	1.03 (0.98-1.07)	>0.963	453	7,385.3	1.01 (0.89-1.15)
Soluble				Soluble			
0 to 1.2742	4625	140,658.5	Ref	0 to 0.11	621	18,314.0	Ref
1.2742 to 4.658	4526	115,882.0	1.16 (1.11-1.21)	0.11 to 1.522	621	16,836.4	1.06 (0.94-1.19)
4.658 to 14.573	4526	111,218.9	1.01 (0.97-1.06)	1.522 to 4.07	621	12,636.9	1.13 (0.99-1.28)
>14.573	4525	79,816.3	1.01 (0.96-1.06)	>4.07	621	9,472.4	0.98 (0.86-1.13)
Synthetic				Synthetic			
0	12652	314,853.2	Ref	0	1604	39,879.6	Ref
0 to 0.2555	1816	54,235.8	0.91 (0.86-0.97)	0 to 0.167	293	6,945.7	1.15 (0.99-1.33)
0.2555 to 1.464	1817	46,134.1	0.95 (0.89-1.01)	0.167 to 0.81	294	6,076.7	1.12 (0.96-1.31)
>1.464	1817	32,352.7	1.08 (1.02-1.15)	>0.81	293	4,357.8	1.11 (0.96-1.29)

[a]Modeled stratified by sex using Cox models with age as time scale and adjusted for race, plant, calendar year, and year of hire. All MWF types were included in the same model.

[b]Upper bound included in range

and survival time. Colorectal cancer cases were more likely to be diagnosed while actively employed, whereas prostate cancer cases were least likely to be actively employed at the time of diagnosis.

The results of the assessment for condition (2) were not entirely consistent across the five cancer outcomes examined (Figure 2.2). Based on the main model adjusted for covariates and past exposure, we found the hazard ratio (HR) associated with having left work was null for all cancers combined and for prostate cancer, slightly elevated for colorectal cancers, and elevated with borderline significance for lung cancer. However, when we considered the age at having left work, the hazard ratios increased for all outcomes.

Based on magnitude and statistical significance, at least one category of MWF exposure was associated with leaving work for all fluid types among males (Table 2.3). Compared to lowest exposed males, those in the second quartile of cumulative exposure to oil-based straight and soluble MWFs were at a higher risk for leaving work. By contrast, only those in the highest exposure category of synthetic MWFs were at a higher risk for leaving work. Among female workers MWF exposure was associated with leaving work only for the water-based soluble and synthetic fluid types. Examining condition (3) using exposure lagged by 15-years did not change the overall conclusion that MWF exposure is associated with leaving work in both male and female workers (data not shown).

2.5 Discussion

The aim of this study was to assess the potential for bias due to the HWSE in a longitudinal study of cancer incidence and metalworking fluid (MWF) exposure in order to guide our future analytical approach. The first condition required for the HWSE is present by definition: (1) leaving work predicts future MWF exposure, because in this, as in all occupational cohort studies, workers cannot be occupationally exposed after employment termination. One caveat is that after terminating work at one of the three study plants, subjects may have been hired at another automobile manufacturing plant where MWF were also in use. Although leaving work is not necessarily an indicator of zero future exposure, we have assumed that it is. This is a limitation of the study, as described below.

Evidence of a higher risk of all cancer outcomes was observed for those who left work by age 50 [condition (2)]. Across all outcomes leaving work by age 50 was a stronger predictor of cancer than leaving work without the age restriction. This pattern is consistent with our hypothesis that workers who terminate employment prematurely are more likely to have a higher risk of the disease outcome. Evidence for condition (2) was stronger for lung and prostate among men, and all cancers combined among women, than for colorectal and all cancers combined among men. Although we cannot state for certain whether this association is due to a causal link between leaving work and these cancers or by an unmeasured confounder, we believe it is most likely the latter. An unmeasured factor, such as health status, could lead to both increased risk of employment termination (e.g., preferential work termination by less healthy workers) and cancer. We did observe younger age at diagnosis among those who left work by age 50 compared with those who

left work after age of 50 for prostate, lung, colorectal, and all cancers combined among men (range 3.8-4.8 years earlier). For all cancers combined among women this difference was only 1.4 years.
We do not have any information on why subjects in this cohort left work. Ideally, we would separate employee-initiated termination (more likely to be due to workers' underlying health status) from lay-offs or other company-initiated terminations. In the absence of any way to distinguish these, we assume all employment termination was voluntary and serves as a proxy for underlying health status. Together with condition (1), these results suggest that employment status is a time-varying confounder and researchers should control for it in their analyses.

We found evidence for the third condition required for the HSWE among all workers: (3) prior MWF exposure increases probability of leaving work. In our analyses, exposure to each of the three types of MWF (straight, soluble, and synthetic) was associated with leaving work among males, and exposure to water-based soluble and synthetic MWF were each associated with leaving work among females. Lagging exposure did not alter these associations, and therefore would not be an effective strategy to avoid bias due to the HWSE. The presence of conditions (1) and (3) imply that employment status should not be adjusted for in our cancer risk models because it is an intermediate between exposure and disease outcome.[18]

The evidence for all three conditions indicates that employment status is a time-varying confounder [conditions (1) and (2)] affected by prior exposure [condition (3)]. We need both to control for employment status as a confounder and to not control for it since it is an intermediate. The presence of the HWSE does not in itself dictate that results will be biased; it is the analytical method applied in the presence of the HWSE that determines whether results will be biased. The observed results for the three conditions suggest that g-methods should be used to address the HWSE and avoid bias in exposure-response analyses for all cancers outcomes examined.[18] G-methods were originally developed to address the HWSE.[10] This is achieved by unpacking cumulative exposure into a series of time varying exposures and building separate models for outcome, exposure, and sometimes other covariates. A large literature on this class of analytic methods explains how these additional models are used to adjust properly for time-varying confounding (by leaving work) and estimate marginal differences in survival within a counterfactual framework.[1,2,6,10,12,19]

These results are generally consistent with findings from a 2012 study of this same cohort comparing standard Cox models with g-estimation models for straight MWF and several cancer mortality outcomes.[11] Chevrier et al reported statistically significant results from g-estimation, but null results using standard methods for lung and prostate cancer mortality, suggesting the presence of the HWSE. For rectal cancer mortality standard methods produced null results, while elevated, though non-significant, risks were found using g-estimation. There was no marked difference between standard and g-estimation results for colon cancer mortality. Costello et al. reported less evidence for the HWSE among women compared to men in an ischaemic heart disease analysis of this same cohort.[20] Differences in the evidence for the HWSE between men and women in the present study are difficult to

distinguish since small numbers prevented the examination of site-specific cancers among female workers.

We presented one DAG for HWSE. Slight variations of this DAG have also been described.[18] Rather than a direct arrow from prior exposure to employment status, there could be, instead or in addition, an unmeasured factor leading to both of these nodes, thereby inducing an association. Similarly, where we have a U leading to both employment status and disease there could be, instead or in addition, a direct arrow leading from employment status to the outcome. Each of these DAGs contains a time-varying confounder affected by prior exposure and its presence can be assessed using the methods presented here. See Naimi 2013 supplemental appendix A1 for details.[18]

Underlying health status may affect exposure in several ways: less healthy employees may take more intermittent time off work, transfer to lower exposed jobs, use personal protective equipment, or terminate employment earlier. In our assessment of the conditions for the HWSE, we only explored the association through the latter pathway. An analysis that takes into account these additional potential pathways between exposure and underlying health status may provide a more accurate assessment of the potential for bias due to HWSE. However, we expect terminating employment to be the strongest source of healthy worker survivor bias, assuming those who left work were not subsequently exposed elsewhere. This is both because the other actions would lead to lower but not necessarily zero exposure, and because a worker who terminates employment is likely to be in worse health than one who manages to transfer to lower exposed job at the plant. A last notable limitation for our study is the small amount of covariate information. Due to this lack of data we were unable to account for other potential confounders of the relationships examined here, such as smoking status or other lifestyle factors.

It has been claimed that the HWSE is unlikely to have an effect on cancer outcomes, particularly cancers with late-stage symptoms and poor survival, such as lung cancer.[21] In the current study of MWF exposed workers, however, we found evidence of the HWSE in relation to incidence of lung cancer. By comparing results from g-methods and standard methods, the presence of the HWSE has also been demonstrated in other occupational studies of lung cancer in relation to arsenic,[10,22] diesel,[23] radon,[24,25] and asbestos,[26,27] as well as MWFs.[11,28] Coupled with the theoretical underpinnings of the HWSE, these studies provide evidence that the HWSE can affect the study of cancer outcomes in occupational settings and should be taken into account when considering different methodologic approaches.

There are several g-methods available for an exposure-responses analysis of MWF and cancer incidence, though not all would be appropriate for this study. Follow-up continues after subjects leave work when exposure is zero. This violates the positivity assumption and precludes the use of inverse probability of exposure weighting.[19] Accelerated failure-time models used with g-estimation assume the outcome is inevitable and assess whether exposure accelerates time to the event. When the outcome is rare, as is cancer, this assumption may not hold.[29] The parametric g-formula[6,10] is more appropriate for rare outcomes and can incorporate quantitative exposure and dynamic treatment regimes,

though it requires many parametric assumptions. When applying any of these methods in this study we will need to adjust for leaving work by a particular age (e.g., 50 years), rather than leaving work regardless of age. Based on our results, this more refined variable definition for leaving work better captures the essence of the healthy worker survivor effect.

2.6 Conclusion

We found evidence for all three conditions necessary for the HWSE to exist in a cohort study of Michigan autoworkers exposed to MWFs. Evidence was strongest for leaving work by age 50, consistent with the HWSE. Lagging exposure by 15 years, in keeping with the latency period for the cancer outcomes, did not alter these results. This suggests that standard methods may underestimate the exposure-response for these outcomes and therefore a g-method should be applied to control for employment status as a time-varying confounder affected by prior exposure.[6,10] Examination of the three conditions for the HWSE should precede exposure-response analyses in other occupational cohort studies to assess the evidence that it is present.

2.7 References

1. Arrighi HM, Hertz-Picciotto I. The evolving concept of the healthy worker survivor effect. *Epidemiology* 1994;5(2):189-96.
2. Buckley JP, Keil AP, McGrath LJ, Edwards JK. Evolving methods for inference in the presence of healthy worker survivor bias. *Epidemiology* 2015;26(2):204-12.
3. Picciotto S, Hertz-Picciotto I. Commentary: healthy worker survivor bias: a still-evolving concept. *Epidemiology* 2015;26(2):213-5.
4. Fox AJ, Collier PF. Low mortality rates in industrial cohort studies due to selection for work and survival in the industry. *Br J Prev Soc Med* 1976;30(4):225-30.
5. Gilbert ES. Some confounding factors in the study of mortality and occupational exposures. *Am J Epidemiol* 1982;116(1):177-88.
6. Robins J. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. *J Chronic Dis* 1987;40 Suppl 2:139S-161S.
7. Pearl J, Glymour M, Jewell NP. *Causal Inference in Statistics: A Primer*. West Sussex, UK: Wiley, 2016.
8. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10(1):37-48.
9. Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology* 2003;14(3):300-6.
10. Robins J. A New Approach to Causal Inference in Mortality Studies with a Sustained Exposure Period - Application to Control of the Healthy Worker Survivor Effect. *Mathematical Modelling* 1986;7(9-12):1393-1512.
11. Chevrier J, Picciotto S, Eisen EA. A comparison of standard methods with g-estimation of accelerated failure-time models to address the healthy-worker survivor effect: application in a cohort of autoworkers exposed to metalworking fluids. *Epidemiology* 2012;23(2):212-9.

12. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11(5):550-60.
13. Eisen EA, Tolbert PE, Monson RR, Smith TJ. Mortality studies of machining fluid exposure in the automobile industry I: A standardized mortality ratio analysis. *Am J Ind Med* 1992;22(6):809-24.
14. Hallock MF, Smith TJ, Woskie SR, Hammond SK. Estimation of historical exposures to machining fluids in the automotive industry. *Am J Ind Med* 1994;26(5):621-34.
15. Woskie SR, Smith TJ, Hallock MF, Hammond SK, Rosenthal F, Eisen EA, Kriebel D, Greaves IA. Size-selective pulmonary dose indices for metal-working fluid aerosols in machining and grinding operations in the automobile manufacturing industry. *Am Ind Hyg Assoc J* 1994;55(1):20-9.
16. Michigan Cancer Surveillance Program. Cancer Statistics. http://www.michigan.gov/mdch/0,4612,7-132-2945_5221-16586--,00.html Accessed May 5, 2015.
17. Division of Cancer Prevention and Control National Center for Chronic Disease Prevention and Health Promotion. National Program of Cancer Registries. http://www.cdc.gov/cancer/npcr/ Accessed November 5, 2014.
18. Naimi AI, Cole SR, Hudgens MG, Brookhart MA, Richardson DB. Assessing the component associations of the healthy worker survivor bias: occupational asbestos exposure and lung cancer mortality. *Ann Epidemiol* 2013;23(6):334-41.
19. Hernan MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health* 2006;60(7):578-86.
20. Costello S, Picciotto S, Rehkopf DH, Eisen EA. Social disparities in heart disease risk and survivor bias among autoworkers: an examination based on survival models and g-estimation. *Occup Environ Med* 2015;72(2):138-44.
21. Choi BC. Definition, sources, magnitude, effect modifiers, and strategies of reduction of the healthy worker effect. *J Occup Med* 1992;34(10):979-88.
22. Arrighi HM, Hertz-Picciotto I. Controlling the healthy worker survivor effect: an example of arsenic exposure and respiratory cancer. *Occup Environ Med* 1996;53(7):455-62.
23. Neophytou AM, Picciotto S, Costello S, Eisen EA. Occupational Diesel Exposure, Duration of Employment, and Lung Cancer: An Application of the Parametric G-Formula. *Epidemiology* 2016;27(1):21-8.
24. Edwards JK, McGrath LJ, Buckley JP, Schubauer-Berigan MK, Cole SR, Richardson DB. Occupational radon exposure and lung cancer mortality: estimating intervention effects using the parametric g-formula. *Epidemiology* 2014;25(6):829-34.
25. Keil AP, Richardson DB, Troester MA. Healthy worker survivor bias in the Colorado Plateau uranium miners cohort. *Am J Epidemiol* 2015;181(10):762-70.
26. Cole SR, Richardson DB, Chu H, Naimi AI. Analysis of occupational asbestos exposure and lung cancer mortality using the g formula. *Am J Epidemiol* 2013;177(9):989-96.
27. Naimi AI, Cole SR, Hudgens MG, Richardson DB. Estimating the effect of cumulative occupational asbestos exposure on time to lung cancer mortality: using structural nested failure-time models to account for healthy-worker survivor bias. *Epidemiology* 2014;25(2):246-54.

28. Picciotto S, Brown DM, Chevrier J, Eisen EA. Healthy worker survivor bias: implications of truncating follow-up at employment termination. *Occup Environ Med* 2013;70(10):736-42.
29. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In: Halloran ME, Berry DA, eds. *Statistical models in epidemiology, the environment and clinical trials*. The IMA volumes in mathematics and its applications. New York: Springer, 2000;95-133.

Chapter 3 Breast Cancer Incidence and Metalworking Fluid Exposure in a Cohort of Female Autoworkers

3.1 Abstract

Breast cancer is the leading cancer diagnosed among women and environmental studies have produced few leads on modifiable risk factors. Following an Institute of Medicine recommendation for occupational studies of highly exposed women, we took advantage of an existing cohort of 4,503 female hourly autoworkers exposed to metalworking fluid (MWF), complex mixtures of oils and chemicals widely used in metal manufacturing worldwide. Cox proportional hazards models were fit to estimate hazard ratios (HR) for incident breast cancer and cumulative exposure (20-year lag) to straight mineral oils (a known human carcinogen), and water-based soluble and synthetic MWF. Because the state cancer registry began in 1985, decades after the cohort was defined, we restricted analyses to sub-cohorts hired closer to the start of cancer follow-up. Among those hired after 1969, the HR associated with an increase of one interquartile range in straight MWF exposure was 1.13 (95% confidence interval: 1.03, 1.23). In separate analyses of premenopausal breast cancer, as defined by age at diagnosis, the HR was elevated for exposure to synthetic MWF, chemical lubricants with no oil content, suggesting a different mechanism for the younger cases. This study adds to the limited literature regarding quantitative chemical exposures and breast cancer risk.

3.2 Introduction

Breast cancer is the leading cancer diagnosed among women in the United States.[1] It is estimated that in 2016 there will be 246,660 new cases and 40,450 deaths attributable to breast cancer among women in the United States.[2] Despite the large impact of this disease, well-established risk factors are estimated to account for only 41% of breast cancer cases[3] and studies of environmental exposures and breast cancer risk have produced few additional leads on modifiable risk factors over the past 20 years.[4] A 2012 Institute of Medicine review of environmental risk factors for breast cancer noted that the initial identifications of many known human carcinogens were based on studies of high exposures in occupational settings and recommended additional breast cancer studies of worker populations.[4]

Despite provocative evidence that metalworking fluid (MWF) contaminants cause mammary gland tumors in laboratory animals[5] and that occupational exposure levels are appreciable, few studies have been conducted on breast cancer risk. MWF are coolants and lubricants widely used in industrial machining and grinding operations, and are categorized into three classes based on composition: straight, soluble, and synthetic. Straight MWF are complex mixtures of paraffinic, naphthenic, and aromatic compounds refined from mineral oil.[6] The carcinogenic properties of mineral oil, classified as carcinogenic to humans,[7] are thought to be primarily due to their polycyclic aromatic hydrocarbons (PAH) content.[8] Soluble MWF are oils emulsified in water. Synthetic fluids are water-soluble chemical lubricants without oil. Ethanolamines and nitrites, added to

synthetic MWF to inhibit corrosion and adjust pH, interact to form nitrosamines. Although no toxicologic studies have specifically examined the effects of MWF and breast carcinogenic effects, MWF components including PAH and nitrosamines have been implicated as mammary gland carcinogens.[5] Additionally, studies on ambient air PAH exposure are suggestive of increased breast cancer risk[9,10] and nitrosamines are suspected endocrine disruptors.[11] Women comprise a growing proportion of the 4.4 million potentially exposed U.S. workers,[12] as well as of the growing workforce employed in metal manufacturing worldwide. The global manufacturing workforce is estimated to be about 30% female.[13] The global market volume of MWF in 2013 was 2.3 million tons; Asia had the largest share, 41.5%, compared to North America, 28.0%.[14]

Few epidemiologic studies on breast cancer risk and MWF exposure exist.[15-22] An early study observed reduced breast cancer mortality among women in automotive manufacturing based on standardized mortality ratios.[15,16] More recently, case-control studies reported elevated breast cancer risk associated with occupation in automotive manufacturing and metal products/metal work.[19-21] These studies, however, did not examine a specific compound, but instead used occupation as a proxy for exposure. Two breast cancer studies have been conducted in the United Autoworker-General Motors (UAW-GM) cohort; both were limited in power and relied on either a combination of incident and mortal cases or only mortal cases. We took advantage of this large existing cohort of female hourly autoworkers exposed to MWF to examine the relationship between quantitative MWF exposure and breast cancer incidence. With an additional nine years of follow-up, we have enough power to restrict our analysis to registry-identified incident cases. The aim of this study was to examine the exposure-response relationship between cumulative MWF exposure estimates and breast cancer risk in a cohort of occupationally-exposed female autoworkers.

3.3 Methods

Study population

The UAW-GM cohort was a joint labor-management funded study designed in 1984 to examine cancer mortality and its relation to MWF exposure. This cohort has been described in detail elsewhere.[23,24] Briefly, the original study included 46,316 hourly workers from one of three automobile manufacturing plants in Michigan exposed to MWF primarily via inhalation and possibly dermal contact.[23] All hourly employees who had worked at least 3 years prior to January 1, 1985 were included in the cohort. Subjects alive on January 1, 1985 (N=33,915) when the Michigan Cancer Registry began were included in the incidence cohort. Analyses for the present study were restricted to female workers in the incidence cohort (N = 4,572). We excluded women who were missing more than 50% of their employment history (N = 59) or were hired prior to 1938 (N = 5). The final study population comprised 4,503 female workers.

Outcome assessment

The UAW-GM incidence cohort was linked with the Michigan Cancer Registry to identify incident cancer cases diagnosed between January 1, 1985 and December 31, 2013. Michigan cancer data are collected by the Michigan Department of Community Health as part of the Michigan Cancer Surveillance Program,[25] which participates in the National Program of Cancer Registries of the Centers for Disease Control and Prevention.[26] For this study, the outcome of interest is first diagnosis of breast cancer (International Classification of Disease for Oncology Third Edition codes C50.0-C50.9). We identified 221 incident cases of breast cancer in the study population, including 72 cases diagnosed at age 55 years or younger. Data on vital status were obtained from the National Death Index (National Center for Health Statistics, Hyattsville, Maryland).[27]

Exposure assessment

Cumulative exposure estimates for each MWF type were calculated for each subject in the UAW-GM cohort based on detailed employment records and a time-varying job-exposure-matrix. An extensive retrospective exposure assessment was conducted to develop this job-exposure-matrix.[28,29] Size fractionated MWF concentrations were estimated as an 8-hour time-weighted average (mg/m^3) based on several hundred personal and area airborne exposure measurements collected by study industrial hygienists in the three manufacturing plants during the mid-1980s. A set of multipliers to adjust MWF concentration for temporal trends were developed based on nearly 400 historical air sampling measurements (collected 1958 through 1987), review of historical records, and interviews with plant personnel. One of the original study industrial hygienists revisited the three plants in 1995 to update the scale factors. This study relies on the respirable size fraction (<3.5 μm) of the MWF exposure estimates, which mostly deposit in the alveolar region. The job-exposure-matrix was combined with employment records to estimate cumulative time-weighted exposure to straight, soluble, and synthetic MWF (mg/m^3-years). Missing employment information was interpolated by averaging exposures from previous and subsequent jobs. To account for breast cancer latency cumulative exposures for each fluid type were lagged by 20 years.

Data analyses

Cox proportional hazards models were fit to estimate hazard ratios and 95% confidence intervals associated with exposure to straight, soluble, and synthetic MWF (20-year lag) on breast cancer incidence. Exposure was defined as cumulative exposure to each fluid type for the respirable size fraction. MWF exposures were modeled both as continuous and categorical. For continuous exposure, HR were computed for an increase of one interquartile range among exposed subjects in the full study population; for the three fluid types, straight, soluble, and synthetic, these are 0.318, 0.979, and 0.270 mg/m^3-years, respectively. For categorical exposure, the referent groups were those subjects with no fluid type exposure. For each fluid type, the median level among exposed cases was used as the cut points for the exposure categories. Age was used as the time metric for all models. All models also included baseline covariates for race (white/black), an established breast cancer risk factor, manufacturing plant to adjust for plant-specific characteristic not captured elsewhere, as well as social-economic status and regional difference, and year of

hire (B-spline with 3 degrees of freedom and equally spaced knots[30]) to account for secular trends in exposure including PAH content, personal protective equipment, and plant ventilation. A time-varying covariate for calendar year (B-spline with 3 degrees of freedom and equally spaced knots) was additionally included to account for secular trends in breast cancer diagnosis. To assess the linearity of exposure-response relationships we evaluated additional Cox models with penalized splines (2 degrees of freedom) for MWF exposure.

By necessity only UAW-GM cohort members alive on January 1, 1985 were included in the incidence cohort, thereby creating a left-truncated cohort. Left truncation occurs when not all otherwise eligible subjects are enrolled in the cohort. Downward bias arises from left truncation because the proportion of subjects susceptible to the effect of exposure decreases over time.[31] To reduce this potential bias, we restricted analyses to a series of sub-cohorts defined by year of hire. Narrowing the time interval between hire and start of the cancer registry reduces the opportunity for susceptible subjects to die prior to start of follow-up in 1985. We defined two sub-cohorts; hired after 1959 and hired after 1969 (20 and 10 years, respectively, prior to the last year of hire among cases). The greater the restriction, the less bias due to left truncation we expect.

We also examined premenopausal breast cancer as an outcome using age at diagnosis as an indicator for menopausal status. The statistical methods employed were the same as those described above for the main analysis. For these analyses, however, a series of four age cut

Table 3.1 Demographic and exposure characteristics of female breast cancer cases and cohort members in the UAW-GM incidence cohort who were alive in 1985 (mean (range) unless otherwise noted).

	Cases	**Cohort**
No. subjects	221	4,503
No. person-years	3,519	108,595
Year of birth	1938 (1907, 1959)	1940 (1893, 1961)
Year of hire	1969 (1942, 1979)	1969 (1940, 1981)
Age at hire (years)	30.2 (18.0, 55.1)	29.3 (16.7, 58.0)
Duration of employment (years)	18.1 (3.2, 47.8)	16.2 (2.9, 47.8)
Race; *n* (%)		
White	155 (70.1)	3,256 (72.3)
African American	66 (29.9)	1,247 (27.7)
Plant; *n* (%)		
Plant 1	25 (11.3)	509 (11.3)
Plant 2	122 (55.2)	2,612 (58.0)
Plant 3	74 (33.5)	1,382 (30.7)
Vital status as of 2013; *n* (%)		
Alive	132 (59.7)	2,947 (65.5)
Deceased	89 (40.3)	1,556 (34.6)
Year of diagnosis	1999 (1985, 2013)	
Age at diagnosis (years)	62.1 (34.8, 94.1)	
Exposure characteristics (mg/m^3-years); mean (SD)[a]		
Straight	0.29 (1.08)	0.20 (1.75)
Soluble	0.68 (1.28)	0.51 (1.42)
Synthetic	0.09 (0.43)	0.08 (0.44)

Abbreviations: UAW-GM, United Autoworkers-General Motors.

[a] Cumulative exposure lagged 20 years

points, 55, 54, 53, and 52 years of age, were used to define premenopausal breast cancer cases and follow-up ended upon reaching that age. A separate model was made for each age cut point. The younger cut points improved specificity but reduced sensitivity in identifying premenopausal breast cancer cases. For these additional analyses, we did not examine sub-cohorts defined by year of hire since these cases were hired later in time. Among cases diagnosed at age 55 or younger, the earliest year of hire was 1966. Thus, left truncation bias is less of a concern here. No apparent violation of the underlying assumption of proportional hazards was detected based on correlations between the Schoenfeld residuals for each MWF of interest and the ranked failure times. SAS software version 9.4 (SAS Institute, Cary, NC) was used for all analyses, except for Cox models with penalized splines which were conducted in R (version 3.2.3, R Core Team, Vienna, Austria). Use of human subjects data in this study was reviewed and approved by the Office for the Protection of Human Subjects at the University of California, Berkeley.

3.4 Results

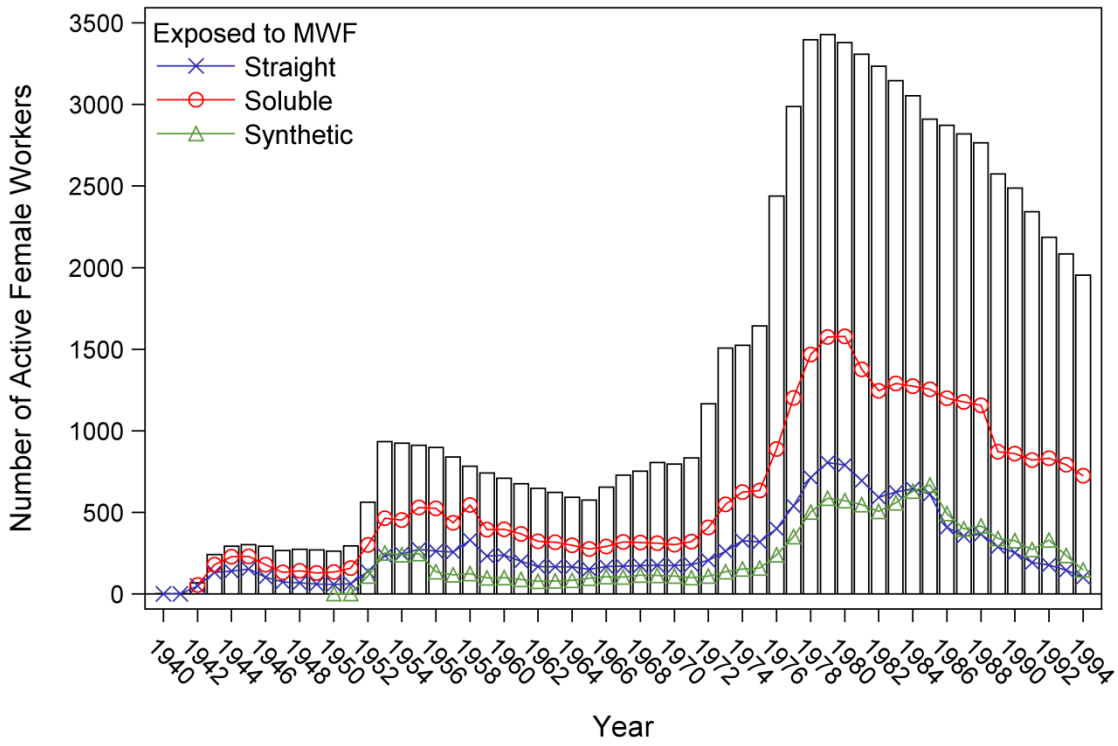


Figure 3.1 Number of active workers and annual exposure prevalence to the three types of metalworking fluids (MWF) by year among female members of the United Autoworker-General Motor incidence cohort who were alive in 1985 (n = 4,503), Michigan, 1940-1994.

Table 3.1 provides demographic and exposure characteristics of the 4,503 female autoworkers comprising the study population, including the 221 incident breast cancer cases. The cohort is predominantly white, but does have more than a quarter African American women. Total number of active female workers in the three plants and their MWF exposure by year is presented in Figure 3.1. Just over half of the women, 50.4%, were hired in 1974 or later. Soluble MWF was the predominant fluid exposure type, followed by

Table 3.2 Adjusted hazard ratios[a] of breast cancer incidence (1985-2013) in relation to cumulative exposure to metalworking fluid size <3.5µm (20-year lag) in female autoworkers based on the full UAW-GM incidence cohort and in subcohorts defined by year of hire.

				Restricted by Year of Hire					
	All *N* = 4,503 (221 Cases)			Hired ≥ 1959 *N* = 3,431 (172 Cases)			Hired ≥ 1969 *N* = 3,182 (157 Cases)		
Cumulative Exposure (mg/m³-years)	**No. Cases (Person-years)**	**HR**	**95% CI**	**No. Cases (Person-years)**	**HR**	**95% CI**	**No. Cases (Person-years)**	**HR**	**95% CI**
Categorical									
Straight									
0	124 (74,802)	1.00	Referent	106 (69,019)	1.00	Referent	95 (65,022)	1.00	Referent
>0-0.112	48 (17,440)	1.39	0.90, 2.14	36 (12,760)	1.64	0.99, 2.73	33 (11,888)	1.67	0.99, 2.83
>0.112	49 (16,353)	1.32	0.86, 2.01	30 (9,251)	1.64	1.01, 2.67	29 (8,406)	1.74	1.05, 2.86
Soluble									
0	80 (52,727)	1.00	Referent	74 (50,819)	1.00	Referent	65 (48,070)	1.00	Referent
>0-0.5	70 (29,660)	0.90	0.61, 1.34	61 (25,716)	0.90	0.58, 1.39	59 (24,371)	1.00	0.63, 1.59
>0.5	71 (26,208)	0.88	0.56, 1.40	37 (14,496)	0.73	0.42, 1.26	33 (12,875)	0.77	0.43, 1.39
Synthetic									
0	168 (86,772)	1.00	Referent	133 (75,619)	1.00	Referent	121 (71,094)	1.00	Referent
>0-0.07	26 (10,213)	0.89	0.53, 1.49	22 (9,099)	0.72	0.40, 1.31	19 (8,555)	0.62	0.33, 1.17
>0.07	27 (11,610)	0.77	0.47, 1.26	17 (6,312)	0.80	0.43, 1.48	17 (5,667)	0.87	0.46, 1.64
Continuous[b]									
Straight		1.00	0.98, 1.02		1.11	1.01, 1.22		1.13	1.03, 1.23
Soluble		1.00	0.91, 1.09		0.97	0.74, 1.28		0.98	0.73, 1.30
Synthetic		0.98	0.90, 1.07		0.82	0.53, 1.29		0.91	0.60, 1.38

Abbreviations: UAW-GM, United Autoworkers-General Motors; CI, confidence interval; HR, hazard ratio.

[a] Each Cox regression models included cumulative exposure to the three fluid types, used age as time scale, and adjusted for year of hire, calendar year, race, and manufacturing plant.

[b] HR per increase of one interquartile range among exposed subjects in the full study population; for the three fluid types these are 0.318, 0.979, and 0.270 mg/m³-years, respectively.

Table 3.3 Adjusted hazard ratios[a] of premenopausal breast cancer incidence (1985-2013) in relation to cumulative exposure to metalworking fluid size <3.5µm (20-year lag) in female autoworkers in the UAW-GM cohort using multiple age cut points for premenopausal definition.

	Age at Diagnosis											
	Age 55 or Younger *N = 3,263 (72 Cases)*			**Age 54 or Younger** *N = 3,211 (65 Cases)*			**Age 53 or Younger** *N = 3,148 (60 Cases)*			**Age 52 or Younger** *N = 3,092 (49 Cases)*		
Cumulative Exposure (mg/m³-years)	**No. Cases (Person-years)**	**HR**	**95% CI**	**No. Cases (Person-years)**	**HR**	**95% CI**	**No. Cases (Person-years)**	**HR**	**95% CI**	**No. Cases (Person-years)**	**HR**	**95% CI**
Categorical												
Straight												
0	54 (48,105)	1.00	Referent	48 (46,151)	1.00	Referent	45 (44,131)	1.00	Referent	39 (42,082)	1.00	Referent
>0-0.08	9 (5,470)	0.95	0.36, 2.52	8 (4,968)	1.08	0.37, 3.10	7 (4,488)	0.82	0.26, 2.52	5 (4,022)	0.88	0.23, 3.31
>0.08	9 (4,208)	1.08	0.44, 2.66	9 (3,692)	1.47	0.56, 3.82	8 (3,196)	1.23	0.45, 3.37	5 (2,728)	1.15	0.34, 3.91
Soluble												
0	42 (39,347)	1.00	Referent	39 (38,184)	1.00	Referent	36 (36,963)	1.00	Referent	33 (35,675)	1.00	Referent
>0-0.332	15 (11,070)	0.79	0.37, 1.71	14 (10,189)	0.81	0.36, 1.85	13 (9,324)	0.81	0.35, 1.89	9 (8,488)	0.60	0.22, 1.63
>0.332	15 (7,366)	0.83	0.33, 2.09	12 (6,437)	0.72	0.26, 2.00	11 (5,529)	0.76	0.27, 2.19	7 (4,669)	0.55	0.16, 1.88
Synthetic												
0	56 (50,721)	1.00	Referent	50 (48,481)	1.00	Referent	46 (46,207)	1.00	Referent	39 (43,914)	1.00	Referent
>0-0.09	8 (5,059)	1.04	0.37, 2.96	8 (4,601)	1.16	0.39, 3.48	8 (4,155)	1.42	0.46, 4.39	5 (3,725)	1.33	0.34, 5.28
>0.09	8 (2,003)	2.40	0.88, 6.54	7 (1,729)	2.71	0.92, 7.98	6 (1,453)	2.78	0.88, 8.73	5 (1,193)	3.76	1.04, 13.51
Continuous[b]												
Straight		0.81	0.51, 1.27		0.87	0.57, 1.35		0.87	0.54, 1.39		0.82	0.42, 1.58
Soluble		0.90	0.64, 1.27		0.91	0.59, 1.39		0.84	0.51, 1.38		0.75	0.39, 1.44
Synthetic		1.06	0.95, 1.17		1.13	0.99, 1.30		1.15	1.00, 1.32		1.17	1.01, 1.35

Abbreviations: UAW-GM, United Autoworkers-General Motors; CI, confidence interval; HR, hazard ratio.

[a] Each Cox regression models included cumulative exposure to the three fluid types, used age as time scale, and adjusted for year of hire, calendar year, race, and manufacturing plant.

[b] HR per increase of one interquartile range among exposed subjects age 55 or younger; for the three fluid types these are 0.157, 0.474, and 0.092 mg/m³-years, respectively.

straights. The percentages of subjects ever exposed to straight, soluble, and synthetic fluids were 53.7, 84.1, and 38.0, respectively. Cumulative exposure concentrations were highest for soluble and lowest for synthetic fluids (Table 3.1).

We observed elevated rates of incident breast cancer among those with greater straight MWF exposure after reducing bias due to left truncation. When we examined the full study population using continuous exposure, results were null for all fluids types (Table 3.2). But in the analyses using sub-cohorts restricted by year of hire, the HR per increase of one interquartile range in straight MWF exposure were increasingly farther from the null with more restrictive sub-cohorts (i.e., with decreased time between hire and the start of follow-up). Examining the exposure-response relation using categorized exposure revealed elevated HR, but with wide confidence intervals, for straight fluids in the full study population (Table 3.2). Results were stronger and presented a more positive exposure-response in the sub-cohort analyses. No elevated HR were observed for either soluble or synthetic fluids. Models with penalized spline coding for straight MWF exposure showed an approximately linear exposure-response relationship with the hazard ratio on the log scale, up to the 99th percentile of exposure among cases for both the 1959 and 1969 year of hire restricted sub-cohorts (data not shown).

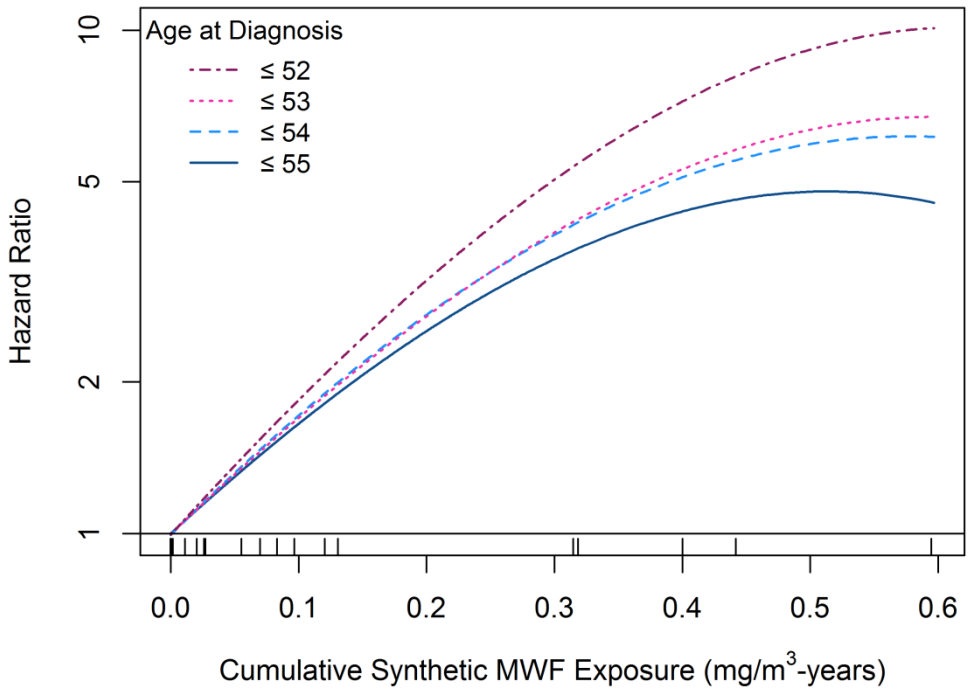


Figure 3.2 Adjusted hazard ratios for premenopausal breast cancer incidence as a smoothed function of 20-year lagged cumulative synthetic metalworking fluid (MWF) as estimated in a Cox regression model using penalized splines (2 degrees of freedom) based on a cohort of female workers in the United Autoworker-General Motor incidence cohort who were alive in 1985, Michigan, 1985-2013. Premenopausal breast cancer was defined as diagnosis by four select age cut points. Model used age as the time scale and adjusted for cumulative exposure to straight and soluble MWF, year of hire, calendar year, race, and manufacturing plant. Graph truncated at the 99th percentile of synthetic fluid exposure among cases (0.60 mg/m^3-years).

Results from the premenopausal breast cancer incidence analyses suggest an increased hazard associated with synthetic fluid exposure, but null associations for straight and soluble MWF. The HR per increase of one interquartile range in exposure to synthetic fluids was elevated when using age 54 as the cut point to define premenopausal cases and increased in magnitude with lower age cut points (Table 3.3). When using a categorical exposure metric, a strong exposure-response was observed for synthetic MWF, particular with younger age cut points; the maximum HR was 3.8 (95% confidence interval (CI): 1.0, 13.5) when age 52 was the cut point for premenopausal breast cancer. Models with penalized spline coding for synthetic MWF exposure showed a positive exposure-response relationship regardless of age cut point used through the 99th percentile of exposure among cases (Figure 3.2). As with the linear and categorical models, HR were higher when the premenopausal case definition was based on younger age cut points.

3.5 Discussion

In a prospective cohort of 4,503 female autoworkers from the UAW-GM study we examined MWF exposure and its association with incident breast cancer. Exposure to straight MWF, but neither soluble nor synthetic, was found to be positively associated with breast cancer. This was evident in the analyses using sub-cohorts restricted by year of hire. When restricted to subjects hired 1959 or later, the HR for straight MWF exposure was elevated in both the continuous and categorical analyses. Results became stronger when using the subcohort restricted to subjects hired 1969 or later. This more restrictive sub-cohort was designed to reduce bias due to left truncation, and may explain the stronger results observed in these analyses.

Though power was adequate for a pooled analysis of pre and post-menopausal cases, when we restricted to cases presumed to be premenopausal, the exposed categories were sparse. Results however, were modestly suggestive of an increased risk associated with higher synthetic MWF exposure. Because of a lack of menopausal status data, age at diagnosis was instead used to define premenopausal breast cancer cases. A discernibly positive exposure-response, however, was found for all four case definitions examined. More pronounced associations were observed when we used a younger age cut point to define cases. Case definitions with these younger cut points were more specific for premenopausal breast cancer, which, when the outcome is rare, is more important than sensitivity and will reduce bias.[32] Among this younger subset we observed the strongest results for an association between synthetic MWF exposure and incident breast cancer. Our interpretation of these analyses, however, was constrained by the small number of cases.

Two previous breast cancer studies relied on data from the UAW-GM study.[17,22] Neither study included only registry-identified incident breast cancer cases. The first study combined incident and mortal cases, and imputed date of diagnosis if missing.[17] The modest risk increase for cumulative soluble MWF exposure observed was inconsistent with the literature on latency for breast cancer, suggesting higher risk associated with exposure in the decade preceding diagnosis. The present analysis was restricted to incident cases and we imposed a 20-year lag on cumulative exposures based on the latency period for

breast cancer.[33-35] The second study, based on 43 breast cancer deaths, found a HR of 1.4 (95% CI: 0.7, 2.5) associated with straight MWF exposure above the median.[22] Mortality, however, is a poor surrogate for breast cancer incidence given a five-year survival rate of 88%.[36] After accounting for left truncation, we observed a similar but marginally larger HR associated with straight fluid exposure above the median. Use of the more appropriate breast cancer outcome measure of incidence rather than mortality may have led to the slightly higher results.

The contrasting results in the pooled analysis versus the sub-analysis restricted to premenopausal breast cancer, suggest a possibly distinct biologic mechanisms for this type of breast cancer. Though, the distribution of cumulative straight MWF exposure was markedly lower among premenopausal cases, with a mean of 0.02 mg/m^3-years (standard deviation (SD): 0.05) based on an age cut point of 55 years compared to all cases at 0.29 (SD: 1.08), and may also have contributed to a lack of association for straight fluids in the premenopausal analyses. The polycyclic aromatic hydrocarbons (PAH) content of straight MWF may be the causal factor in the elevated risk. Researchers examined air pollution exposure, as a proxy for PAH, and found associations with postmenopausal breast cancer. Total suspended particulate exposure was associated with postmenopausal, but not premenopausal, breast cancer in a population-based, case-control study in Western New York State.[9] A later study in the same population modeled total PAH exposure and similarly found exposure at first birth to be associated with postmenopausal, but not premenopausal, breast cancer.[10] The only study to examine occupational PAH exposure found an increased odds of premenopausal breast cancer among those exposed.[34] Because benzene was a co-exposure in the study, the authors additionally evaluated exclusive PAH exposed and found null results. Studies of the Long Island Breast Cancer Study Project have investigated PAH-DNA adducts, which serve as a short-term PAH biomarkers of exposure, and found them to be positively associated with incident breast cancer.[37,38]

There is less comparable evidence for synthetic fluids and premenopausal breast cancer risk. A class of potentially hazardous contaminants found in synthetic fluids are nitrosamines which form due to the co-occurrence of amines and nitrites or nitrates.[39] N-nitrosodiethanolamine, a type of nitrosamine, has been found in synthetic fluids[40,41] and in post-shift urine samples of workers exposed to water-based MWF.[42] It is classified as possibly carcinogenic to humans and is an animal carcinogen.[43] It has been demonstrated to induce DNA damage in both animal and human cells[44] and metal workers exposed to higher levels of *N*-nitrosodiethanolamine were found to have increased levels of DNA single strand breaks in blood cells.[45] While there is no specific research on this compound and breast cancer risk, these studies lend to some biological plausibility for synthetic fluids being associated with increased premenopausal breast cancer risk.

Bias due to the healthy worker survivor effect is a concern in occupational epidemiology.[46-48] We recently assessed the presence of the healthy worker survivor effect in the UAW-GM incidence cohort based on three necessary underlying conditions: 1) leaving work predicts future exposure, 2) leaving work is associated with disease outcome, and 3) prior exposure increases probability of leaving work.[49] The first condition is a given, since those subjects who leave work are no longer exposed. We found prior soluble and synthetic, but not

straight, MWF exposure to be associated with leaving work among female subjects. This supports the third condition for the water-based soluble and synthetic fluids, but not for straight fluids. While breast cancer incidence was not examined as part of that assessment, we used the same statistical method to evaluate it here. We found that having left work was not associated with breast cancer incidence, indicating a lack of evidence for the second condition. Overall, these results imply that healthy worker survivor bias did not influence our estimates of the association between MWF exposure and breast cancer risk, particularly the analyses of straight fluids.

The main limitation of this study is unmeasured potential confounders. While we do control for age and race, we do not account for several other breast cancer risk factors including social economic status, family history of breast cancer, age at menarche or menopause, age at first-full term pregnancy, parity, breastfeeding history, use of oral contraceptives or postmenopausal hormones, and alcohol consumption. Because these are all blue-collar workers employed at the same three plants, we do not expect a large difference in social economic status. Additionally, we do not expect family history of breast cancer, age at menarche or menopause, or use of oral contraceptives or postmenopausal hormones to be related to MWF exposure. Lastly, prior studies of this cohort have found that cirrhosis death, a proxy for alcohol consumption, was not associated with MWF exposure.[50] Therefore, the potential confounders of most concern for the present study are those related to childbearing: age at first full-term pregnancy, parity, and breastfeeding history. Women with a younger age at first full-term pregnancy, higher parity, and who breastfed longer have a reduced risk of breast cancer and are likely to have lower MWF exposure due to taking maternity leave, entering the workforce later, and/or leaving the workforce earlier. This would produce positive confounding and bias results upward. In the above HWSE evaluation, however, we did not see evidence of leaving work being associated with reduced breast cancer risk. Furthermore, we would expect such a bias to affect the results for all three fluid types, not just the straight fluids in the pooled analyses or the synthetic fluids in the premenopausal analyses. Although unmeasured confounding may account for some portion of the results observed, this study makes a contribution to the literature.

3.6 Conclusion

This study adds to the limited literature regarding quantitative chemical exposures and breast cancer risk in humans. This study utilized a well-characterized occupational cohort of blue collar women with long cancer incidence follow-up and provided evidence of an increased risk of breast cancer associated with cumulative straight MWF exposure. The observation of suggestive results for higher synthetic fluid exposure and premenopausal breast cancer presents a new avenue for future research.

3.7 References

1. U.S. Cancer Statistics Working Group. United States Cancer Statistics: 1999–2011 Incidence and Mortality Web-based Report. Atlanta, GA: Department of Health and

Human Services, Center for Disease Control and Prevention, and National Cancer Institute, 2014.
2. SEER. Cancer Statistics Factsheets: Female Breast Cancer. http://seer.cancer.gov/statfacts/html/breast.html Accessed November 7, 2016.
3. Madigan MP, Ziegler RG, Benichou J, Byrne C, Hoover RN. Proportion of breast cancer cases in the United States explained by well-established risk factors. *J Natl Cancer Inst* 1995;87(22):1681-5.
4. Committee on Breast Cancer and the Environment. Breast Cancer and the Environment: A Life Course Approach. Insititute of Medicine, 2012;468.
5. Rudel RA, Attfield KR, Schifano JN, Brody JG. Chemicals causing mammary gland tumors in animals signal new directions for epidemiology, chemicals testing, and risk assessment for breast cancer prevention. *Cancer* 2007;109(12 Suppl):2635-66.
6. Neale M. *Tribiology Handbook*. New York: John Wiley and Sons, 1973.
7. International Agency for Research on Cancer (IARC). Mineral Oils, Untreated or Mildly Treated. *Monographs, Chemical Agents and Related Occupations, vol 100F*. Vol. 100F. Lyon, France: IARC, 2012.
8. Catchpole WM, Macmillan E, Powell H. Specifications for cutting oils with special reference to carcinogenicity. *Ann Occup Hyg* 1971;14(2):171-9.
9. Bonner MR, Han D, Nie J, Rogerson P, Vena JE, Muti P, Trevisan M, Edge SB, Freudenheim JL. Breast cancer risk and exposure in early life to polycyclic aromatic hydrocarbons using total suspended particulates as a proxy measure. *Cancer Epidemiol Biomarkers Prev* 2005;14(1):53-60.
10. Nie J, Beyea J, Bonner MR, Han D, Vena JE, Rogerson P, Vito D, Muti P, Trevisan M, Edge SB, Freudenheim JL. Exposure to traffic emissions throughout life and risk of breast cancer: the Western New York Exposures and Breast Cancer (WEB) study. *Cancer Causes Control* 2007;18(9):947-55.
11. U.S. Environmental Protection Agency. Endocrine Disruptor Screening Program: Universe of Chemicals for Potential Endocrine Disruptor Screening and Testing. In: Office of Chemical Safetly and Pollution Prevention, Office of Water, Office of Research and Development, eds, 2012.
12. Steenland K, Burnett C, Lalich N, Ward E, Hurrell J. Dying for work: The magnitude of US mortality from selected causes of death associated with occupation. *Am J Ind Med* 2003;43(5):461-82.
13. International Labour Office. Global Employment Trends for Women: March 2009. Geneva, Switzerland: ILO, 2009.
14. Frost and Sullivan. Analysis of the Global Metalworking Fluids Market. Mountain View, CA, 2014.
15. Delzell E, Beall C, Macaluso M. Cancer mortality among women employed in motor vehicle manufacturing. *J Occup Med* 1994;36(11):1251-9.
16. Beall C, Delzell E, Macaluso M. Mortality patterns among women in the motor vehicle manufacturing industry. *Am J Ind Med* 1995;28(3):325-37.
17. Thompson D, Kriebel D, Quinn MM, Wegman DH, Eisen EA. Occupational exposure to metalworking fluids and risk of breast cancer among female autoworkers. *Am J Ind Med* 2005;47(2):153-60.
18. Brophy JT, Keith MM, Gorey KM, Luginaah I, Laukkanen E, Hellyer D, Reinhartz A, Watterson A, Abu-Zahra H, Maticka-Tyndale E, Schneider K, Beck M, Gilbertson M.

Occupation and breast cancer: a Canadian case-control study. *Ann N Y Acad Sci* 2006;1076:765-77.
19. Hansen J. Breast cancer risk among relatively young women employed in solvent-using industries. *Am J Ind Med* 1999;36(1):43-7.
20. Villeneuve S, Fevotte J, Anger A, Truong T, Lamkarkach F, Gaye O, Kerbrat P, Arveux P, Miglianico L, Imbernon E, Guenel P. Breast cancer risk by occupation and industry: analysis of the CECILE study, a population-based case-control study in France. *Am J Ind Med* 2011;54(7):499-509.
21. Brophy JT, Keith MM, Watterson A, Park R, Gilbertson M, Maticka-Tyndale E, Beck M, Abu-Zahra H, Schneider K, Reinhartz A, Dematteo R, Luginaah I. Breast cancer risk in relation to occupations with exposure to carcinogens and endocrine disruptors: a Canadian case-control study. *Environ Health* 2012;11:87.
22. Friesen MC, Betenia N, Costello S, Eisen EA. Metalworking fluid exposure and cancer risk in a retrospective cohort of female autoworkers. *Cancer Causes Control* 2012;23(7):1075-82.
23. Eisen EA, Tolbert PE, Monson RR, Smith TJ. Mortality studies of machining fluid exposure in the automobile industry I: A standardized mortality ratio analysis. *Am J Ind Med* 1992;22(6):809-24.
24. Eisen EA, Bardin J, Gore R, Woskie SR, Hallock MF, Monson RR. Exposure-response models based on extended follow-up of a cohort mortality study in the automobile industry. *Scand J Work Environ Health* 2001;27(4):240-9.
25. Michigan Cancer Surveillance Program. Cancer Statistics. http://www.michigan.gov/mdch/0,4612,7-132-2945_5221-16586--,00.html Accessed May 5, 2015.
26. Dividion of Cancer Prevention and Control National Center for Chronic Disease Prevention and Health Promotion. National Program of Cancer Registries. http://www.cdc.gov/cancer/npcr/ Accessed November 5, 2014.
27. Cowper DC, Kubal JD, Maynard C, Hynes DM. A primer and comparative review of major US mortality databases. *Ann Epidemiol* 2002;12(7):462-8.
28. Hallock MF, Smith TJ, Woskie SR, Hammond SK. Estimation of historical exposures to machining fluids in the automotive industry. *Am J Ind Med* 1994;26(5):621-34.
29. Woskie SR, Smith TJ, Hallock MF, Hammond SK, Rosenthal F, Eisen EA, Kriebel D, Greaves IA. Size-selective pulmonary dose indices for metal-working fluid aerosols in machining and grinding operations in the automobile manufacturing industry. *Am Ind Hyg Assoc J* 1994;55(1):20-9.
30. De Boor C. *A practical guide to splines*. Applied mathematical sciences: 27 New York : Springer-Verlag, c1978., 1978.
31. Applebaum KM, Malloy EJ, Eisen EA. Left truncation, susceptibility, and bias in occupational cohort studies. *Epidemiology* 2011;22(4):599-606.
32. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology*. 3rd ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2008.
33. Aschengrau A, Paulu C, Ozonoff D. Tetrachloroethylene-contaminated drinking water and the risk of breast cancer. *Environ Health Perspect* 1998;106 Suppl 4:947-53.
34. Petralia SA, Vena JE, Freudenheim JL, Dosemeci M, Michalek A, Goldberg MS, Brasure J, Graham S. Risk of premenopausal breast cancer in association with occupational

exposure to polycyclic aromatic hydrocarbons and benzene. *Scand J Work Environ Health* 1999;25(3):215-21.
35. Kelsey JL, Berkowitz GS. Breast cancer epidemiology. *Cancer Res* 1988;48(20):5615-23.
36. ACS. Cancer Facts & Figures 2014. Atlanta, GA: American Cancer Society, 2014.
37. Gammon MD, Sagiv SK, Eng SM, Shantakumar S, Gaudet MM, Teitelbaum SL, Britton JA, Terry MB, Wang LW, Wang Q, Stellman SD, Beyea J, Hatch M, Kabat GC, Wolff MS, Levin B, Neugut AI, Santella RM. Polycyclic aromatic hydrocarbon-DNA adducts and breast cancer: a pooled analysis. *Arch Environ Health* 2004;59(12):640-9.
38. Gammon MD, Santella RM, Neugut AI, Eng SM, Teitelbaum SL, Paykin A, Levin B, Terry MB, Young TL, Wang LW, Wang Q, Britton JA, Wolff MS, Stellman SD, Hatch M, Kabat GC, Senie R, Garbowski G, Maffeo C, Montalvan P, Berkowitz G, Kemeny M, Citron M, Schnabel F, Schuss A, Hajdu S, Vinceguerra V. Environmental toxins and breast cancer on Long Island. I. Polycyclic aromatic hydrocarbon DNA adducts. *Cancer Epidemiol Biomarkers Prev* 2002;11(8):677-85.
39. NIOSH. Criteria for a Recommended Standard: Occupational Exposure to Metalworking Fluids. Vol. DHHS Publication No. 98-102. Cincinnati, OH: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health,, 1998.
40. Fan TY, Morrison J, Rounbehler DP, Ross R, Fine DH, Miles W, Sen NP. N-Nitrosodiethanolamine in synthetic cutting fluids: a part-per-hundred impurity. *Science* 1977;196(4285):70-1.
41. Keefer LK, Goff U, Stevens J, Bennett EO. Persistence of N-nitrosodiethanolamine contamination in American metal-working lubricants. *Food Chem Toxicol* 1990;28(7):531-4.
42. Ducos P, Gaudin R. N-nitrosodiethanolamine urinary excretion in workers exposed to aqueous metalworking fluids. *Int Arch Occup Environ Health* 2003;76(8):591-7.
43. International Agency for Research on Cancer (IARC). Some Industrial Chemicals. *Monographs on the Evaluation of the Carcinogenic Risks to Humans, vol 77*. Vol. 77. Lyon, France: IARC, 2000.
44. Holzer J, Voss B, Karroum S, Hildmann H, Wilhelm M. A comparative study of chemically induced DNA damage in isolated nasal mucosa cells of humans and rats assessed by the alkaline comet assay. *J Toxicol Environ Health A* 2008;71(13-14):936-46.
45. Fuchs J, Burg J, Hengstler JG, Bolm-Audorff U, Oesch F. DNA damage in mononuclear blood cells of metal workers exposed to N-nitrosodiethanolamine in synthetic cutting fluids. *Mutat Res* 1995;342(1-2):95-102.
46. Arrighi HM, Hertz-Picciotto I. The evolving concept of the healthy worker survivor effect. *Epidemiology* 1994;5(2):189-96.
47. Buckley JP, Keil AP, McGrath LJ, Edwards JK. Evolving methods for inference in the presence of healthy worker survivor bias. *Epidemiology* 2015;26(2):204-12.
48. Picciotto S, Hertz-Picciotto I. Commentary: healthy worker survivor bias: a still-evolving concept. *Epidemiology* 2015;26(2):213-5.
49. Garcia E, Picciotto S, Costello S, Bradshaw PT, Eisen EA. Assessment of the healthy worker survivor effect in cancer studies of the United Autoworkers-General Motors cohort. *Occup Environ Med* 2017.

50. Eisen EA, Tolbert PE, Hallock MF, Monson RR, Smith TJ, Woskie SR. Mortality studies of machining fluid exposure in the automobile industry. III: A case-control study of larynx cancer. *Am J Ind Med* 1994;26(2):185-202.

Chapter 4 Metalworking Fluids Exposure and Lung Cancer Mortality; Controlling for the Healthy Worker Survivor Effect

4.1 Abstract

Metalworking fluids (MWF) used to cool and lubricate industrial machining and grinding operations are linked with several cancers. Studies of lung cancer, however, are inconsistent. Applying methods to control for the healthy worker survivor effect, we examined the relationship between lung cancer mortality and exposure to straight and synthetic MWF, as well as to biocide added to water-based fluids to control microbial growth, in a cohort of autoworkers. Using the parametric g-formula, we estimated risk ratios (RR) comparing cumulative lung cancer mortality under hypothetical interventions: always high exposure while at work and always unexposed. In high-exposure interventions, MWF exposure was set to the 99th percentile of annual average daily exposure (straight fluids: 1.36 mg/m^3; synthetic fluids: 0.48 mg/m^3) or biocides set to always exposed (specified as a time-varying binary variable). We also specified interventions on synthetic MWF and biocides simultaneously to estimate independent effects. We observed a RR=1.10 (95% confidence interval: 0.79-1.43) for lung cancer mortality at age 86 due to straights fluids, which increased slightly at younger ages: 1.15 (0.84-1.55) at age 55. For synthetic MWF exposure, the risk of lung cancer mortality was RR=1.03 (0.68-1.38), which increased marginally when biocide interventions were considered. Biocide exposure was inversely associated with lung cancer mortality (RR=0.83 (0.64-1.02)). Results from this study suggest slightly elevated lung cancer mortality related to straight MWF exposure, albeit with wide confidence intervals. Our results do not support a negative association for synthetic fluids reported in earlier studies; instead, biocide in the fluid, a marker for endotoxin, was associated with decreased lung cancer.

4.2 Introduction

In the United States, 4.4 million workers are potentially exposed to metalworking fluids (MWF), which are coolants and lubricants used in industrial machining and grinding operations.[1] MWF are classified into three categories based on composition: straight, soluble, and synthetic. Straight fluids are complex mixtures of paraffinic, naphthenic, and aromatic compounds refined from mineral oil.[2] Soluble fluids are composed of oils emulsified in water. Synthetic fluids are water-soluble chemical lubricants without oil. While MWF have been linked with several cancers, including skin,[3-7] laryngeal,[7-12] rectal,[7,11-14] pancreatic,[7,15,16] esophageal,[7,16,17] and bladder,[18-21] their relationship with lung cancer is less clear.

Straight MWF potentially contain polycyclic aromatic hydrocarbons—including benzo(a)pyrene, a lung carcinogen[22]—but most cohort-based studies have reported a null association with lung cancer.[11,12,23-25] Only one study has reported elevated lung cancer mortality statistically significantly associated with straight fluids.[26] In contrast, two studies found decreased risk associated with the water-based synthetic fluids.[27,28] These negative

results were suggested to be due to endotoxins in the synthetic MWF. Endotoxins are components of the outer membrane of gram-negative bacteria cell walls that are released when bacteria lyse.[29] They have reported antitumor activity possibly through release of tumor necrosis factor α in response to exposure and are thought to be responsible for reduced lung cancer rates in cotton textile and agricultural workers.[27,29] Bacterial contamination of water-based MWF, particular by gram-negative species, can be considerable and lead to fluid deterioration and foul odors.[30] Biocides are routinely added to the fluids to counteract (but in the long-term do not eliminate) microbial growth and thus serve to indicate bacterial and endotoxin contamination.[27,28,30-33]

The United Autoworkers-General Motors (UAW-GM) study is an occupational cohort of 46,316 hourly workers in automotive manufacturing and is considered to be the most comprehensive cohort study of MWF-exposed workers.[12,24,25,34] We took advantage of the cohort's long follow-up and extensive quantitative exposure assessment to examine the relationship between lung cancer mortality and exposure to straight and synthetic MWF. Prior analyses demonstrated the presence of the healthy worker survivor effect[35] for lung cancer in the UAW-GM cohort,[36] indicating the need to employ a method that can adjust for time-varying confounding affected by prior exposure.[37,38] Therefore, we employed the parametric g-formula,[39] which uses standardization to control time-varying confounders affected by prior exposure,[40] to evaluate lung cancer mortality under multiple hypothetical interventions on straight and synthetic fluids. Biocide exposure was considered a proxy for endotoxin contamination, and was additionally intervened upon in order to observe the independent effect of synthetic fluids on lung cancer mortality.

4.3 Methods

Study Population

The UAW-GM cohort has been described in detail previously.[7,12] Briefly, the original study included all hourly workers hired between 1938 and 1982 who worked at least three years at any of three automobile manufacturing plants in Michigan, U.S.A.[12] We restricted the cohort to those missing less than half of their employment history and hired no younger than age 16. The final study population comprised 38,560 workers. Information on date of birth, year of hire, race, sex, and plant was ascertained from employment records. For our analysis, follow-up began three years after hire and ended at death, age 86 (age of oldest case), or the end of 1994, whichever occurred first.

Outcome ascertainment

Data on vital status were obtained through linkage with the Social Security Administration, the National Death Index, plant records, and state mortality files. Death certificates, state vital records, and the National Death Index were used to determine cause of death. The outcome of interest was lung cancer mortality (ICD-9 162; ICD-10 C34). We identified 873 lung cancer deaths during follow-up from 1941 through 1994.

Exposure Assessment

Exposure estimates for each MWF fluid were calculated for each subject based on detailed employment records and a time-varying job-exposure-matrix (JEM). The JEM was based on an extensive retrospective exposure assessment.[41,42] MWF concentrations characterized by particle-size fraction were estimated for homogenous exposure groups as an 8-hour time-weighted average (mg/m^3) based on measurements collected during the mid-1980s. Scale factors were developed to adjust MWF concentrations for temporal trends. This study uses the thoracic size fraction (<9.8 μm) of the exposure estimates, which deposit mostly in the tracheobronchial and alveolar regions of the lung. The presence of biocide was determined for each homogenous exposure group over time by reviewing Material Safety Data Sheets and historical records of lubricant specification. The JEM was combined with employment records to estimate time-varying annual average daily and cumulative exposure to straight, soluble, and synthetic MWF (mg/m^3 and mg/m^3-years, respectively) and to categorize subjects into a time-varying indicator of past biocide exposure (never or ever exposed). Gaps in employment history information were interpolated by averaging exposures from previous and subsequent jobs. To account for lung cancer latency, cumulative MWF exposures were lagged by 15 years.

Statistical Methods

We applied the parametric g-formula to estimate the cumulative risk ratio of several hypothetical MWF exposure interventions on lung cancer mortality. In addition to the simulated natural course (no intervention), we evaluated separate sets of interventions on three exposures: straight fluids, synthetic fluids, and biocides, as well as joint interventions on synthetic fluids and biocides. The interventions are listed in table 4.1.

Table 4.1. Exposure interventions simulated in the UAW-GM cohort study. [MWF=metalworking fluid]

Intervention variable	Intervention	MWF intervention [a]	Biocide intervention [b]
(1) Straight MWF	Always unexposed	0	·
	Always exposed while at work	99th%tile (1.36 mg/m^3)	·
(2) Synthetic MWF	Always unexposed	0	·
	Always exposed while at work	99th%tile (0.48 mg/m^3)	·
(3 and 4) Synthetic MWF and biocides	Always unexposed to synthetic fluids and always exposed to biocides	0	Always exposed
	Always exposed while at work and always exposed to biocides	99th%tile (0.48 mg/m^3)	Always exposed
	Always unexposed to synthetic fluids and always unexposed to biocides	0	Always unexposed
	Always exposed while at work and always unexposed to biocides	99th %tile (0.48 mg/m^3)	Always unexposed
(5) Biocides	Always exposed to biocides	·	Always exposed
	Always unexposed to biocides	·	Always unexposed

· No intervention
[a] annual average daily concentration
[b] ever/never

The parametric g-formula is a generalization of standardization that can be applied to estimation of measures of association in the presence of time-varying exposures and covariates under the following three assumptions: 1) no unmeasured confounders (conditional exchangeability), 2) counterfactual consistency, and 3) correct model specification.[37,43,44] It has had several applications in epidemiologic studies, including a number of instructive articles.[43,45-47] We describe it briefly. This method allowed us to estimate the risk of lung cancer mortality under each intervention as a weighted sum of the probability of lung cancer death, conditional on past exposure and covariate histories. We first developed parametric models for the outcome, competing risk, time-varying covariates, and exposures based on the observed data. These models were conditional on prior covariate and exposure histories, and on baseline covariates. Next, pseudo-samples based on the observed sample population were used in Monte Carlo simulations along with the estimates from the parametric models to generate covariate and exposure histories. Lastly, cumulative risk of lung cancer mortality under each hypothetical intervention was quantified, and interventions were compared using risk ratios.

Parametric models were fitted for the outcome (lung cancer mortality), competing risk (non-lung cancer death), the three separate MWF exposures (annual average daily exposure to straight, soluble, and synthetic fluids lagged 15 years), and the time-varying covariate employment status (lagged 15 years), all conditional on calendar year, age, prior covariate and exposure histories and baseline covariates (sex, race, plant, and year of hire). Cumulative MWF exposures for each of the three fluid types were calculated based on the annual exposure histories. A model was fit for biocide exposure, conditional on calendar year, age, prior covariate and exposure histories and baseline covariates as well as prior biocide exposure and the current year's MWF exposure level for each fluid type. Cumulative MWF exposures and the biocide covariate were used only in the prediction of the outcome and competing risk, not in the prediction of MWF exposures or employment status. See appendix for additional details.

Workers in the pseudo-sample, drawn to be of the same size as the cohort (n=38,560), were followed from their age at the beginning of follow-up through age 86. Parameters from the models for exposures and covariates described above were used to simulate exposure and covariate levels at each age. After employment termination, all exposures were set to 0. For interventions on exposure, the value of the exposure of interest was intervened on by changing it from the predicted value to the intervention value (either 0 or 99th percentile of exposure for MWF or exposed/unexposed for biocide), conditional on being employed. Subsequent values of exposure, covariates, and risk of non-lung cancer and lung cancer mortality were then similarly simulated at each age through the end of follow-up. For the simulated natural course, no variables were intervened upon and the simulation predicted risk under the natural course of events. For each intervention, lung cancer mortality was evaluated using a cumulative mortality estimator for the sub-distribution of the event of interest in the presence of competing risks.[48] We then calculated cumulative incidence ratios for lung cancer mortality by ages 86, 75, 65, and 55 years. We ascertained 95% confidence intervals from 500 bootstraps each the same size as the cohort (n=38,560).

The analysis was performed in SAS software version 9.4 (SAS Institute, Cary, NC) based on the GFORMULA macro available at https://www.hsph.harvard.edu/causal/software/. Use of human subjects data in this study was reviewed and approved by the Office for the Protection of Human Subjects at the University of California, Berkeley.

4.4 Results

Table 4.2. Demographic characteristics of the UAW-GM cohort at baseline. [IQR=interquartile range]

	N	%	Median	IQR
No. subjects	38,560			
No. person-years	1,122,160			
Characteristic				
Age			24	19–32
Male	33,801	87.7		
Female	4,759	12.3		
White	31,426	81.5		
African American	7,134	18.5		
Plant				
1	9,090	23.6		
2	17,094	44.3		
3	12,376	32.1		
Calendar Year			1965	1952–1973
Number of events				
Lung cancer deaths	873			
Other deaths	8,537			

Table 4.2 presents descriptive characteristics of the 38,560 autoworkers included in the study population. Figure 4.1 shows the number of workers being followed, indicating person-time, and the number of lung cancer deaths by age.

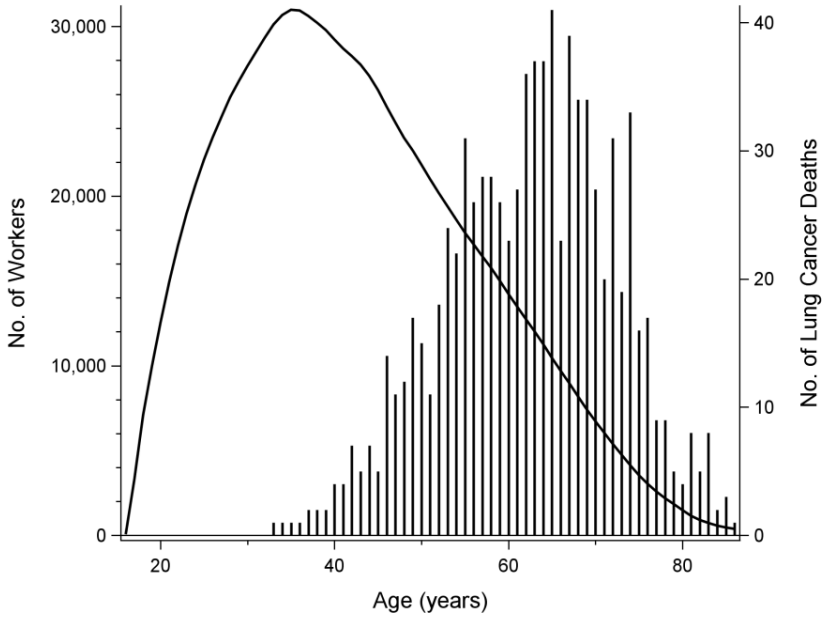


Figure 4.1. Number of workers under follow-up (curved line) and number of lung cancer deaths (vertical lines) by age among 38,560 autoworkers during the 1,122,160 person-years of follow-up between January 1, 1941 and December 31, 1994.

Table 4.3. Exposure characteristics over person-years in the observed data and simulated interventions [a] in the UAW-GM cohort. [MWF=metalworking fluid; Med=median; IQR=interquartile range]

	Observed			Simulated Natural Course			Straight Fluid Interventions						Synthetic Fluid Interventions					
							Exposed			Unexposed			Exposed			Unexposed		
	%	Med	IQR	%	Med	IQR	%	Med	IQR	%	Med	IQR	%	Med	IQR	%	Med	IQR
Employed [b]	76.7			63.8			73.9			64.5			65.1			64.2		
MWF [b c]																		
Straight	21.7	0.08	0.04–0.43	13.9	0.09	0.03–0.35	73.9	1.36	1.36–1.36	0.0	·	·	14.9	0.04	0.02–0.13	12.5	0.11	0.03–0.40
Soluble	47.4	0.48	0.20–0.88	29.3	0.26	0.12–0.65	37.8	0.32	0.14–0.76	28.1	0.28	0.12–0.69	26.1	0.20	0.09–0.49	28.8	0.27	0.12–0.66
Synthetic	10.9	0.09	0.01–0.13	10.8	0.02	0.01–0.08	10.5	0.03	0.01–0.09	9.3	0.02	0.01–0.10	65.1	0.48	0.48–0.48	0.0	·	·
Biocide																		
Ever Exposed	25.4			32.0			30.9			32.4			36.1			32.0		

[a] Exposed interventions: MWF exposure set to the 99th percentile of exposure (1.36 mg/m^3 for straights and 0.48 mg/m^3 for synthetics) while employed; Unexposed interventions: MWF exposure set to always unexposed.

[b] Lagged 15 years

[c] Percent of person-years exposed and annual exposure (mg/m^3) distribution among exposed

Over the entire study period, the percentages of subjects ever exposed to straight, soluble, and synthetic MWF were 58, 90, and 36, respectively. Estimated annual average daily exposure estimates were highest for soluble and lowest for synthetic fluids. In a given year an average of 12% of subjects were exposed to biocides, while over the entire study period 36% were ever exposed to biocides. Synthetic MWF exposure did not always entail biocide exposure. In our study population, workers were unexposed to biocide 22% of the time they worked a job with synthetic exposure. Subjects could also be, though rarely were, exposed to biocide when working with soluble fluids: 7% of the time spent in jobs with soluble MWF exposure also had biocide exposure. Table 4.3 shows the distributions of employment status and exposure over person-time in the observed data and in the pseudo-populations for the simulated natural course and for four of the simulated interventions. The natural course simulation was somewhat consistent with the observed data. The simulation under-predicted both *whether* a subject was exposed to straight or soluble fluids in a given year and the exposure *levels* among the exposed for soluble and synthetic fluids. Because biocide exposure was not used in the prediction models for leaving work or MWF exposure, the distributions reported for the synthetic fluid interventions in table 4.3 were essentially the same as those when we also intervened on biocide exposure (data not shown).

Figure 4.2 shows the cumulative mortality curve for lung cancer in the observed data compared to the simulated natural course. The cumulative lung cancer mortality under the natural course simulation was nearly identical to the observed cumulative risk at age 86; though at younger ages the fit was not as close.

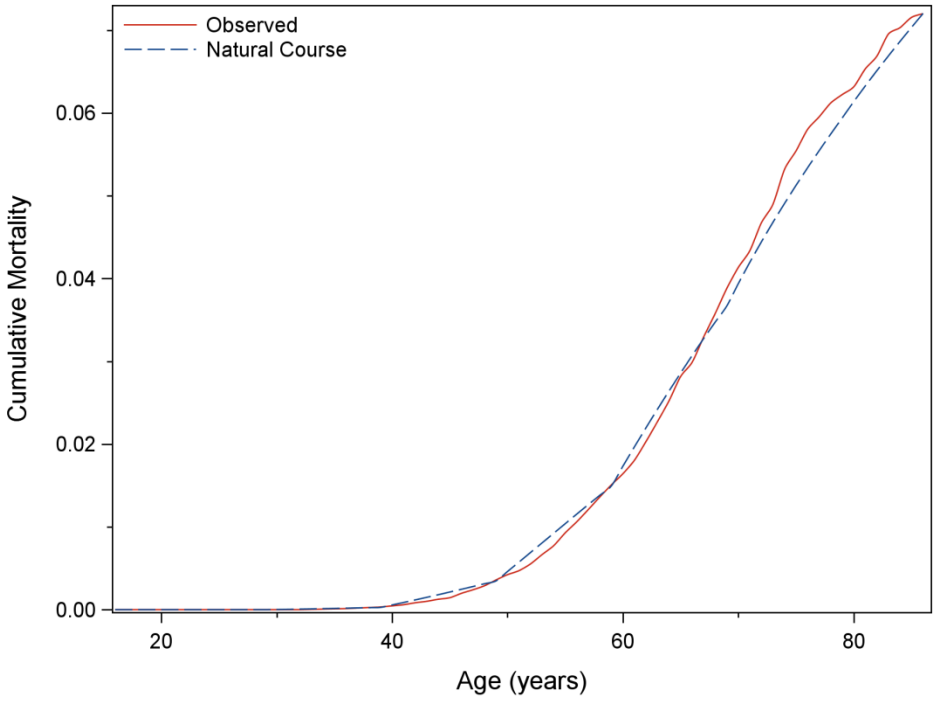


Figure 4.2. Observed (solid line) and simulated (dashed line) cumulative lung cancer mortality in the UAW-GM cohort. under the natural course.

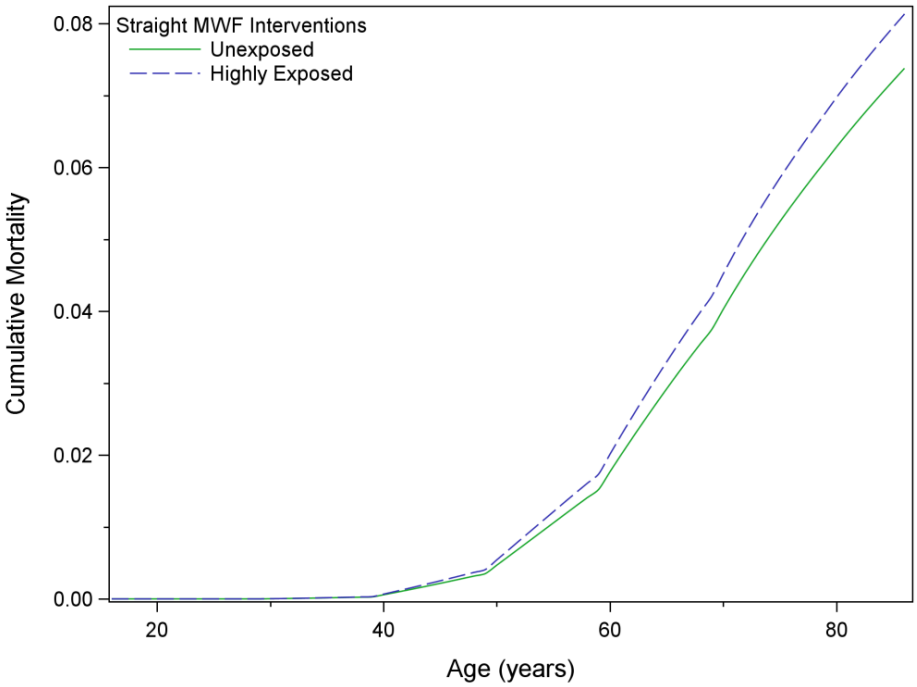


Figure 4.3. Simulated cumulative lung cancer mortality in the UAW-GM cohort under two straight MWF interventions: always unexposed (solid line) and always exposed to the 99th percentile of exposure (1.36 mg/m^3) while employed (dashed line).

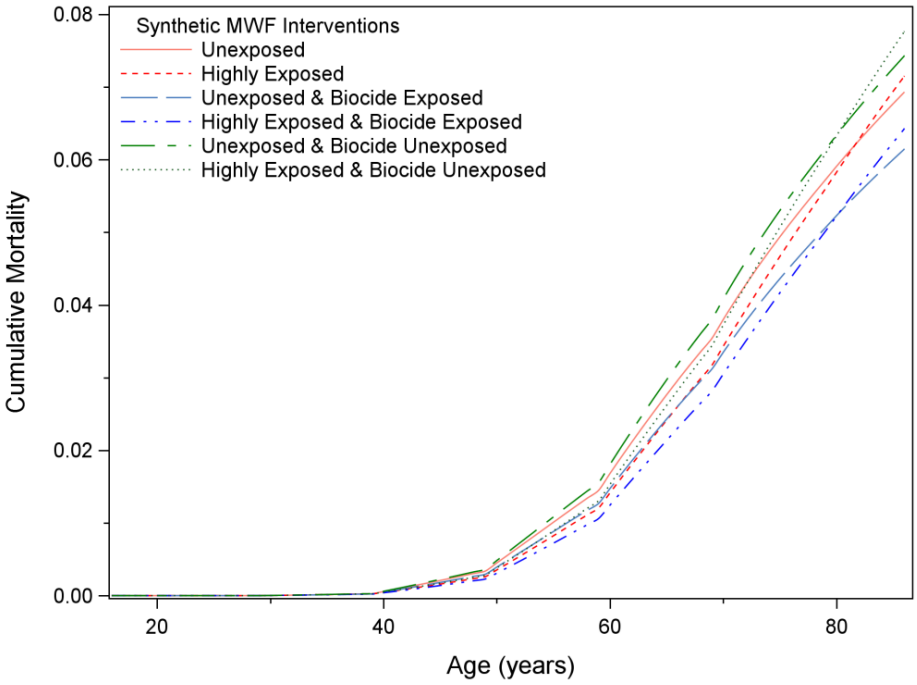


Figure 4.4. Simulated cumulative lung cancer mortality in the UAW-GM cohort under six synthetic MWF interventions: 1) always unexposed (solid line), 2) always exposed to the 99th percentile of exposure (0.48 mg/m^3) while employed, interventions 1) and 2) repeated while additionally intervening on biocide exposure and setting subjects to always exposed (interventions 3) and 4)), and interventions 1) and 2) repeated while additionally intervening on biocide exposure and setting subjects to always unexposed (interventions 5) and 6)).

Table 4.4. Cumulative lung cancer mortality by select ages under several simulated interventions [a] in the UAW-GM cohort. [MWF=metalworking fluid; Exp=exposed; Unexp=unexposed; RR=risk ratio; 95% CI=95% confidence interval]

	86 years-old				75 years-old				65 years-old				55 years-old			
Intervention	Exp Risk	Unexp Risk	RR	95% CI	Exp Risk	Unexp Risk	RR	95% CI	Exp Risk	Unexp Risk	RR	95% CI	Exp Risk	Unexp Risk	RR	95% CI
(1) Straight MWF	8.1	7.4	1.10	0.78–1.47	5.9	5.2	1.12	0.77–1.51	3.3	2.9	1.13	0.77–1.54	1.2	1.1	1.15	0.78–1.58
(2) Synthetic MWF only	7.2	6.9	1.03	0.67–1.43	4.7	4.9	0.95	0.58–1.40	2.4	2.8	0.87	0.51–1.34	0.8	1.0	0.82	0.49–1.26
(3) Synthetic MWF, exposed to biocides	6.4	6.1	1.05	0.68–1.46	4.2	4.4	0.96	0.58–1.42	2.1	2.4	0.88	0.52–1.37	0.7	0.9	0.82	0.49–1.28
(4) Synthetic MWF, unexposed to biocides	7.8	7.4	1.04	0.68–1.45	5.1	5.3	0.96	0.59–1.42	2.6	3.0	0.88	0.52–1.36	0.9	1.1	0.82	0.49–1.28
(5) Biocide only	6.4	7.7	0.83	0.67–1.01	4.5	5.5	0.82	0.66–1.01	2.5	3.1	0.82	0.65–1.01	0.9	1.1	0.81	0.65–1.00

[a] Exposed interventions: MWF exposure set to the 99th percentile of exposure (1.36 mg/m^3 for straights and 0.48 mg/m^3 for synthetics) while employed, biocide exposure set to always ever exposed; Unexposed interventions: exposure set to always unexposed.

Figure 4.3 and 4.4 show the cumulative mortality curves for lung cancer under the straight and synthetic MWF interventions, respectively. The risk ratios (RR) comparing the risk of lung cancer mortality under each exposed intervention to the unexposed intervention are presented in Table 4.4. Risk for lung cancer mortality was higher when straight MWF exposure was set to high exposure during employment compared to when subjects were always unexposed (intervention 1). The RR was slightly higher at younger ages. For synthetic fluid exposure (intervention 2), the RR for lung cancer mortality at age 86 years was only just above the null, but increased slightly when we additionally intervened on biocide exposure, regardless of whether biocide exposure was set to always exposed (intervention 3) or always unexposed (intervention 4). The RRs for mortality at younger ages, however, were null and even appear negatively associated. When synthetic fluid exposures were allowed to follow a natural course, lung cancer risk was lower when workers were always exposed to biocide compared to when they were always unexposed to biocides (intervention 5).

4.5 Discussion

Most cohort-based studies of straight MWF and lung cancer risk have found a null association.[23-25] Results from this study suggest a slightly elevated risk of lung cancer mortality related to straight MWF exposure, albeit with a wide confidence interval. Prior studies, particularly early reports from this cohort,[11,12] may have been biased due to the healthy worker survivor effect.[36] The one study to observe a positive result for straight fluids and lung cancer mortality risk applied g-estimation, an alternative g-method, in this same cohort to control for the healthy worker survivor effect,[26] estimating a hazard ratio of 1.07 (1.04-1.14) associated with five years of exposure to straight MWF. Although the specific parameter estimated in that study differs from that estimated here, both indicate a positive relationship.

A previous study in the UAW-GM cohort focusing on biocide and synthetic MWF exposure found reduced risk of lung cancer mortality associated with synthetic fluids when biocides were also present.[28] Our hypothesis was that biocide exposure—serving as a proxy for endotoxin contamination—was causing the protective effect of synthetic fluids on lung cancer, and that synthetic fluid itself did not reduce lung cancer risk. We therefore investigated the impact of the synthetic fluid intervention while 1) not intervening on biocide exposure, and intervening on biocides by either 2) setting all subjects to always exposed or 3) all subjects to always unexposed. Our results supported our hypothesis. Rather than a protective effect of synthetics, we observed a nominally elevated risk comparing always highly exposed to always unexposed. This risk increased slightly in the scenarios where we also intervened on biocide exposure, consistent with biocide exposure being negatively linked with lung cancer risk. These results suggest that exposure to synthetic fluid, per se, does not reduce lung cancer mortality, but rather it is the presence of biocide in the water-based fluid, serving as a surrogate for endotoxins, that reduces the risk.

Several occupational studies have supported the link between endotoxin exposure and reduced risk for lung cancer, particularly in the textile and agricultural sectors.[29,49,50] A 2010 meta-analysis reported the summary risk of lung cancer as 0.72 (0.57-0.90) in the cotton textile industry and 0.62 (0.52-0.75) in the agricultural industry.[49] Our results, using biocide as a surrogate for endotoxin, are consistent with this hypothesized link. In the scenario comparing the interventions always unexposed to biocide to always exposed to biocide, we found a moderate decrease in lung cancer mortality of borderline significance. Prior studies in the UAW-GM cohort are inconsistent regarding this relationship. The aforementioned UAW-GM study on biocide and synthetic MWF exposure observed a null association for years exposed to biocide.[28] An earlier study, however, observed reduced lung cancer odds ratios for workers with longer durations of exposure to biocides.[27] Workers with 8.52 or more years of biocide exposure had an odds ratio of 0.54 (0.34-0.86) compared with unexposed workers. Differences in analytical choices resulted in subjects from the earlier study having higher biocide exposures compared to subjects in the later study, a possible reason for the discrepancy. In the present analysis, we dichotomize biocide exposure as ever or never exposed.

We employed the parametric g-formula to control for the healthy worker survivor effect, the presence of which was reported for lung cancer in a prior analysis of the UAW-GM cohort.[36] A limitation of the parametric g-formula is that it is predicated on the assumption of correctly specified models; in our study, particularly for the parametric models for employment status, MWF and biocide exposures, death due to a competing risk, and death due to lung cancer (see appendix for details). Models may be biased if there are violations of the assumptions of no unmeasured confounders, no information bias, or correct model specification.[43] We optimized these models to predict cumulative lung cancer mortality at age 86; had we selected another age end point the models may have been specified differently. See figure 4.2 for deviations in the cumulative mortality curves at ages younger than 86. We also assumed the G null paradox is not the reason for our results. The G null paradox states that under the null hypothesis it may be impossible to correctly specify the parametric models for the g-formula, and consequently the parametric g-formula will reject the null when it is in fact true.[37,51] A limitation of the UAW-GM cohort was the lack of smoking data. We were unable to neither assess whether smoking was associated with exposure which would make it a potential confounder, nor control for it in our analysis. The only data available was limited to a very small number of selected male subjects collected mid-way through follow-up (1985-1986). While 1,666 workers in the present study participated in this survey, we did not use the data here because it was too selected to be representative of our entire study population. Results reported in this study may be affected by this limitation. Additionally, a lack of smoking data may have limited our models for competing risk since it is a major predictor of mortality outside of lung cancer mortality.

4.6 Conclusion

While not definitive, this study adds to the literature on MWF exposure and risk of lung cancer. By using the parametric g-formula, we controlled for the healthy worker survivor

effect previously observed to be present in this cohort.[36] Results from this study suggest a slightly elevated risk of lung cancer mortality related to straight MWF exposure, albeit with wide confidence intervals. Our results do not support a negative effect for synthetic fluids reported in earlier studies; instead, biocide in the fluid, a proxy for endotoxin contamination, appears to be associated with decreased lung cancer risk.

4.7 References

1. Steenland K, Burnett C, Lalich N, Ward E, Hurrell J. Dying for work: The magnitude of US mortality from selected causes of death associated with occupation. *Am J Ind Med* 2003;43(5):461-82.
2. Neale M. *Tribiology Handbook*. New York: John Wiley and Sons, 1973.
3. Jarvholm B, Lavenius B. Mortality and cancer morbidity in workers exposed to cutting fluids. *Arch Environ Health* 1987;42(6):361-6.
4. Jarvholm B, Fast K, Lavenius B, Tomsic P. Exposure to cutting oils and its relation to skin tumors and premalignant skin lesions on the hands and forearms. *Scand J Work Environ Health* 1985;11(5):365-9.
5. Roush GC, Kelly JA, Meigs JW, Flannery JT. Scrotal carcinoma in Connecticut metalworkers: sequel to a study of sinonasal cancer. *Am J Epidemiol* 1982;116(1):76-85.
6. Costello S, Friesen MC, Christiani DC, Eisen EA. Metalworking fluids and malignant melanoma in autoworkers. *Epidemiology* 2011;22(1):90-7.
7. Eisen EA, Bardin J, Gore R, Woskie SR, Hallock MF, Monson RR. Exposure-response models based on extended follow-up of a cohort mortality study in the automobile industry. *Scand J Work Environ Health* 2001;27(4):240-9.
8. Zagraniski RT, Kelsey JL, Walter SD. Occupational risk factors for laryngeal carcinoma: Connecticut, 1975-1980. *Am J Epidemiol* 1986;124(1):67-76.
9. Wortley P, Vaughan TL, Davis S, Morgan MS, Thomas DB. A case-control study of occupational risk factors for laryngeal cancer. *Br J Ind Med* 1992;49(12):837-44.
10. Eisen EA, Tolbert PE, Hallock MF, Monson RR, Smith TJ, Woskie SR. Mortality studies of machining fluid exposure in the automobile industry. III: A case-control study of larynx cancer. *Am J Ind Med* 1994;26(2):185-202.
11. Tolbert PE, Eisen EA, Pothier LJ, Monson RR, Hallock MF, Smith TJ. Mortality studies of machining-fluid exposure in the automobile industry. II. Risks associated with specific fluid types. *Scand J Work Environ Health* 1992;18(6):351-60.
12. Eisen EA, Tolbert PE, Monson RR, Smith TJ. Mortality studies of machining fluid exposure in the automobile industry I: A standardized mortality ratio analysis. *Am J Ind Med* 1992;22(6):809-24.
13. Decoufle P. Further analysis of cancer mortality patterns among workers exposed to cutting oil mists. *J Natl Cancer Inst* 1978;61(4):1025-30.
14. Gerhardsson de Verdier M, Plato N, Steineck G, Peters JM. Occupational exposures and cancer of the colon and rectum. *Am J Ind Med* 1992;22(3):291-303.
15. Bardin JA, Eisen EA, Tolbert PE, Hallock MF, Hammond SK, Woskie SR, Smith TJ, Monson RR. Mortality studies of machining fluid exposure in the automobile industry. V: A case-control study of pancreatic cancer. *Am J Ind Med* 1997;32(3):240-7.

16. Silverstein M, Park R, Marmor M, Maizlish N, Mirer F. Mortality among bearing plant workers exposed to metalworking fluids and abrasives. *J Occup Med* 1988;30(9):706-14.
17. Sullivan PA, Eisen EA, Woskie SR, Kriebel D, Wegman DH, Hallock MF, Hammond SK, Tolbert PE, Smith TJ, Monson RR. Mortality studies of metalworking fluid exposure in the automobile industry: VI. A case-control study of esophageal cancer. *Am J Ind Med* 1998;34(1):36-48.
18. 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(12):1471-8.
19. Park RM, Mirer FE. A survey of mortality at two automotive engine manufacturing plants. *Am J Ind Med* 1996;30(6):664-73.
20. Schifflers E, Jamart J, Renard V. Tobacco and occupation as risk factors in bladder cancer: a case-control study in southern Belgium. *Int J Cancer* 1987;39(3):287-92.
21. 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(4):701-13.
22. International Agency for Research on Cancer (IARC). Benzo[a]pyrene. *Monographs, Chemical Agents and Related Occupations, vol 100F*. Vol. 100F. Lyon, France: International Agency for Research on Cancer, World Health Organization, 2012.
23. NIOSH. Criteria for a Recommended Standard: Occupational Exposure to Metalworking Fluids. Vol. DHHS Publication No. 98-102. Cincinnati, OH: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health,, 1998.
24. Tolbert PE. Oils and cancer. *Cancer Causes Control* 1997;8(3):386-405.
25. Savitz DA. Epidemiologic evidence on the carcinogenicity of metalworking fluids. *Appl Occup Environ Hyg* 2003;18(11):913-20.
26. Chevrier J, Picciotto S, Eisen EA. A comparison of standard methods with g-estimation of accelerated failure-time models to address the healthy-worker survivor effect: application in a cohort of autoworkers exposed to metalworking fluids. *Epidemiology* 2012;23(2):212-9.
27. Schroeder JC, Tolbert PE, Eisen EA, Monson RR, Hallock MF, Smith TJ, Woskie SR, Hammond SK, Milton DK. Mortality studies of machining fluid exposure in the automobile industry. IV: A case-control study of lung cancer. *Am J Ind Med* 1997;31(5):525-33.
28. Mehta AJ, Malloy EJ, Applebaum KM, Schwartz J, Christiani DC, Eisen EA. Reduced lung cancer mortality and exposure to synthetic fluids and biocide in the auto manufacturing industry. *Scand J Work Environ Health* 2010;36(6):499-508.
29. Lundin JI, Checkoway H. Endotoxin and cancer. *Environ Health Perspect* 2009;117(9):1344-50.
30. Rossmoore HW. Antimicrobial agents for water-based metalworking fluids. *J Occup Med* 1981;23(4):247-54.
31. Linnainmaa M, Kiviranta H, Laitinen J, Laitinen S. Control of workers' exposure to airborne endotoxins and formaldehyde during the use of metalworking fluids. *AIHA J (Fairfax, Va)* 2003;64(4):496-500.

32. Mattsby-Baltzer I, Sandin M, Ahlstrom B, Allenmark S, Edebo M, Falsen E, Pedersen K, Rodin N, Thompson RA, Edebo L. Microbial growth and accumulation in industrial metal-working fluids. *Appl Environ Microbiol* 1989;55(10):2681-9.
33. Veillette M, Thorne PS, Gordon T, Duchaine C. Six month tracking of microbial growth in a metalworking fluid after system cleaning and recharging. *Ann Occup Hyg* 2004;48(6):541-6.
34. Calvert GM, Ward E, Schnorr TM, Fine LJ. Cancer risks among workers exposed to metalworking fluids: a systematic review. *Am J Ind Med* 1998;33(3):282-92.
35. Arrighi HM, Hertz-Picciotto I. The evolving concept of the healthy worker survivor effect. *Epidemiology* 1994;5(2):189-96.
36. Garcia E, Picciotto S, Costello S, Bradshaw PT, Eisen EA. Assessment of the healthy worker survivor effect in cancer studies of the United Autoworkers-General Motors cohort. *Occup Environ Med* 2017;74(4):294-300.
37. Robins J. A New Approach to Causal Inference in Mortality Studies with a Sustained Exposure Period - Application to Control of the Healthy Worker Survivor Effect. *Mathematical Modelling* 1986;7(9-12):1393-1512.
38. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15(5):615-25.
39. Robins J. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. *J Chronic Dis* 1987;40 Suppl 2:139S-161S.
40. Hernan MA, Robins JM. *Causal Inference*. Boca Raton: Chapman & Hall/CRC, forthcoming, 2017.
41. Hallock MF, Smith TJ, Woskie SR, Hammond SK. Estimation of historical exposures to machining fluids in the automotive industry. *Am J Ind Med* 1994;26(5):621-34.
42. Woskie SR, Smith TJ, Hallock MF, Hammond SK, Rosenthal F, Eisen EA, Kriebel D, Greaves IA. Size-selective pulmonary dose indices for metal-working fluid aerosols in machining and grinding operations in the automobile manufacturing industry. *Am Ind Hyg Assoc J* 1994;55(1):20-9.
43. Taubman SL, Robins JM, Mittleman MA, Hernan MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int J Epidemiol* 2009;38(6):1599-611.
44. Robins J, Hernan M. Estimation of the causal effects of time-varying exposures. In: Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G, eds. *Longitudinal Data Analysis*. CRC Handbooks of Modern Statistical Methods. Boca Raton, FL: Chapman and Hall, 2008;553-599.
45. Cole SR, Richardson DB, Chu H, Naimi AI. Analysis of occupational asbestos exposure and lung cancer mortality using the g formula. *Am J Epidemiol* 2013;177(9):989-96.
46. Neophytou AM, Picciotto S, Costello S, Eisen EA. Occupational Diesel Exposure, Duration of Employment, and Lung Cancer: An Application of the Parametric G-Formula. *Epidemiology* 2016;27(1):21-8.
47. Keil AP, Edwards JK, Richardson DB, Naimi AI, Cole SR. The parametric g-formula for time-to-event data: intuition and a worked example. *Epidemiology* 2014;25(6):889-97.
48. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol* 2009;170(2):244-56.

49. Lenters V, Basinas I, Beane-Freeman L, Boffetta P, Checkoway H, Coggon D, Portengen L, Sim M, Wouters IM, Heederik D, Vermeulen R. Endotoxin exposure and lung cancer risk: a systematic review and meta-analysis of the published literature on agriculture and cotton textile workers. *Cancer Causes Control* 2010;21(4):523-55.
50. Mastrangelo G, Fedeli U, Fadda E, Milan G, Lange JH. Epidemiologic evidence of cancer risk in textile industry workers: a review and update. *Toxicol Ind Health* 2002;18(4):171-81.
51. Eisen EA, Picciotto S, Robins JM. Healthy Worker Effect. *Encyclopedia of Environmetrics* John Wiley & Sons, Ltd, 2006.

4.8 Appendix

An assumption of the parametric g-formula is correct model specification. However, baseline variables did not require parametric models as these values were drawn from the empirical distribution. While we do intervene upon the exposure variables in the simulated datasets, we do not intervene on all the exposures all of the time. For some interventions, exposure variables are treated simply as time-varying covariates. Thus, we assumed correct specifications of the parametric models for employment status, death due to a competing risk, death due to lung cancer, and for biocide and metalworking fluid (MWF) exposures, including both annual and cumulative models for each of straight, soluble, and synthetic MWF. As an informal check of this assumption, we compared the cumulative lung cancer mortality in the observed data to the cumulative lung cancer mortality in the simulated natural course. Specification of the models described below was done with the goal of minimizing the difference between the observed and simulated natural course cumulative lung cancer mortality at age 86. The observed cumulative lung cancer mortality was 7.201%. The final models, as described below, achieved a simulated a natural course cumulative lung cancer mortality of 7.202%.

Baseline covariates were entered in all models as follows: indicator variables for sex (male, female) and race (white, African American), and plant (1, 2, 3); linear term for year of hire; categorical variable for calendar year with cut points every five years starting with 1950 through 1990; categorical variable for age with a level per decade starting with the 30s. Calendar year and age were simulated as time-varying covariates that increased by one year for each subject for each person-year record after baseline. Age was the time-scale of interest, indexed a = 16 to 86. All models were conditional on prior survival.

The parametric models were fit as follows:

1. A logistic model for the probability of employment termination status at age a, restricted to records where prior employment termination status = 0 (not terminated). Because of the 15-year lag, the first 15 records were set to 0. As an independent variable for succeeding covariate models, the prior value of employment termination status was included as an indicator variable.

2. A two-stage process for the level of annual average daily straight metalworking fluid (MWF) exposure at age a. First, a logistic model was used for the probability of straight

MWF exposure being greater than 0. Second, for records in which exposure was predicted to be greater than 0, a linear model predicted the natural log of the straight MWF exposure level. In records in which employment termination status = 1 (terminated), exposure was set to 0. Because of the 15-year lag, the first 15 records were set to 0 exposure. As an independent variable for succeeding covariate models, the prior two values of straight MWF exposure were included as categorical variables with cut points at the quintiles of annual exposure among all observed person-time: 0.0295, 0.0569, 0.164, and 0.5395 mg/m^3.

3. A model accumulating annual average daily straight MWF exposure to predict cumulative straight MWF exposure at age *a*. Because of the 15-year lag, the first 15 records were set to 0 exposure. Values of cumulative straight MWF exposure were not used as independent variables in covariate models. These were only used as independent variables in the models for lung cancer death and non-lung cancer death, specifically the prior two values of cumulative straight MWF exposure were included as restricted cubic splines with knots at the quintiles of cumulative exposure among observed lung cancer deaths: 0.0871, 0.3401, 0.9497, and 2.8895 mg/m^3-years.

4. A two-stage process for the level of annual average daily soluble metalworking fluid (MWF) exposure at age *a*. First, a logistic model was used for the probability of soluble MWF exposure being greater than 0. Second, for records in which exposure was predicted to be greater than 0, a linear model predicted the natural log of the soluble MWF exposure level. In records in which employment termination status = 1 (terminated), exposure was set to 0. Because of the 15-year lag, the first 15 records were set to 0 exposure. As an independent variable for succeeding covariate models, the prior two values of soluble MWF exposure were included as categorical variables with cut points at the quartiles of annual exposure among all observed person-time: 0.2047, 0.4767, and 0.8841 mg/m^3.

5. A model accumulating annual average daily soluble MWF exposure to predict cumulative soluble MWF exposure at age *a*. Because of the 15-year lag, the first 15 records were set to 0 exposure. Values of cumulative soluble MWF exposure were not used as independent variables in covariate models. These were only used as independent variables in the models for lung cancer death and non-lung cancer death, specifically the prior two values of cumulative soluble MWF exposure were included as restricted cubic splines with knots at the quintiles of cumulative exposure among observed lung cancer deaths: 1.4934, 3.5005, 6.9223, and 13.5493 mg/m^3-years.

6. A two-stage process for the level of annual average daily synthetic metalworking fluid (MWF) exposure at age *a*. First, a logistic model was used for the probability of synthetic MWF exposure being greater than 0. Second, for records in which exposure was predicted to be greater than 0, a linear model predicted the natural log of the synthetic MWF exposure level. In records in which employment termination status = 1 (terminated), exposure was set to 0. Because of the 15-year lag, the first 15 records were set to 0 exposure. As an independent variable for succeeding covariate models, the prior two values of synthetic MWF exposure were included as categorical variables with

cut points at the quintiles of annual exposure among all observed person-time: 0.0143, 0.0286, 0.0878, and 0.1831 mg/m^3.

7. A model accumulating annual average daily synthetic MWF exposure to predict cumulative synthetic MWF exposure at age *a*. Because of the 15-year lag, the first 15 records were set to 0 exposure. Values of cumulative synthetic MWF exposure were not used as independent variables in covariate models. These were only used as independent variables in the models for lung cancer death and non-lung cancer death, specifically the prior two values of cumulative synthetic MWF exposure were included as restricted cubic splines with knots at the quartiles of cumulative exposure among observed lung cancer deaths: 0.1347, 0.5216, and 1.3826 mg/m^3-years.

8. A logistic model for the probability of biocide exposure at age *a*. Restricted to records where prior biocide exposure = 0 (never exposed). Values of biocide exposure were not used as independent variables in covariate models. These were only used as independent variables in the models for lung cancer death and non-lung cancer death, specifically the prior value of biocide exposure was included as an indicator variable (never or ever exposed).

9. A logistic model for the probability of death due to causes other than lung cancer at age *a*.

10. A logistic model for the probability of lung cancer death at age *a*.

Chapter 5 Conclusions

Breast and lung cancer are the two most commonly diagnosed cancers in the United States.[1,2] Combined they account for an estimated 28% of all new cancer cases. Recognizing their significance to public health, I planned my dissertation research to evaluate their relationship with the occupational exposure, metalworking fluids (MWF). In the United States there are 4.4 million workers potentially exposed to MWF.[3] We took advantage of the United Autoworkers-General Motors (UAW-GM) cohort study, which is considered to be the most comprehensive cohort study of MWF-exposed workers, to examine the exposure-response relationships for these two outcomes.[4-7] Knowing that the Healthy Worker Survivor Effect (HWSE) may be present in the cohort and could bias results from the cancer studies, I first examined evidence for the presence of the HWSE to guide our subsequent cancer analyses. The cancer studies were then carried out appropriately based on these results.

5.1 Summary of Findings

<u>Findings of Chapter 2:</u> Evidence for the structural conditions underlying the HWSE were present in a cohort study of Michigan autoworkers exposed to MWF.

This chapter assessed the evidence for the presence of the HWSE for cancer studies in the UAW-GM cohort study. This was done by evaluating the three conditions underlying the HSWE: (1) leaving work predicts future exposure, (2) leaving work is associated with disease outcome, and (3) prior exposure increases probability of leaving work. We found evidence for the presence of all three conditions necessary for the HWSE. Evidence was strongest for leaving work by age 50, consistent with the HWSE. Lagging exposure by 15 years, in keeping with the latency period for the cancer outcomes, did not alter these results. This suggests that standard methods may underestimate the exposure-response for the cancer outcomes examined in the study and therefore a g-method should be applied to control for employment status as a time-varying confounder affected by prior exposure.[8,9] Examination of the three conditions for the HWSE should precede exposure-response analyses in other occupational cohort studies to assess whether it is present.

<u>Findings of Chapter 3:</u> Cumulative exposure to straight MWF was positively associated with incident breast cancer.

Chapter 3 examined MWF exposure and its association with incident breast cancer among 4,503 female autoworkers in the UAW-GM cohort study. Exposure to straight MWF, but neither soluble nor synthetic, was found to be positively associated with breast cancer. This was evident in the analyses using sub-cohorts restricted by year of hire. Results were stronger when using the subcohort restricted to subjects hired 1969 or later compared to the subcohort restricted to subjects hired 1959 or later. This more restrictive subcohort was designed to reduce bias due to left truncation (i.e., the loss of subjects susceptible to the effect of exposure prior to the start of follow-up in nested subcohorts), and may explain the stronger results observed in these analyses. Analyses of premenopausal breast cancer,

as defined by age at diagnosis, were modestly suggestive of an increased risk associated with higher synthetic MWF exposure. Interpretation of these analyses, however, must be constrained due to the small number of cases, but the results present a new avenue for future research. This study adds to the limited literature regarding quantitative chemical exposures and breast cancer risk.

<u>Findings of Chapter 4:</u> After controlling for the HWSE, straight MWF exposure may be associated with slightly elevated lung cancer mortality.

In this chapter, I employed the parametric g-formula to control for the healthy worker survivor effect previously observed to be present for lung cancer in the UAW-GM cohort.[10] Using this method I examined the relationship between lung cancer mortality and exposure to straight and synthetic MWF, as well as to biocide, by simulating and comparing the hypothetical interventions of always high exposure while at work and always unexposed. Results indicated a slight increase in lung cancer risk by age 86 for straight MWF exposure and risk ratios were slightly higher at younger ages. These results, however, were not statistically significant. Our results did not support a negative effect for synthetic fluids as reported in earlier studies; instead, biocide in the fluid, a proxy for endotoxin contamination, appears to be associated with decreased lung cancer risk. While not definitive, this study adds to the literature on MWF exposure and risk of lung cancer.

5.2 Conclusions and Future Directions

By first evaluating evidence for the presence of the HWSE for cancer outcomes in the UAW-GM cohort, we better informed our statistical approach for the cancer studies. It should be emphasized that the presence of HWSE does not itself dictate that results will be biased; it is the analytical method applied in the presence of the HWSE that determines whether results will be biased. Standard methods (i.e., Poisson or Cox proportional hazards models) will likely produce biased results because they cannot control for a time-varying confounder affected by prior exposure.[9,11] More advanced methods, such as g-methods, are able to adjust for this type of bias.[8,9,12,13]

While the study presented in chapter 3 provides support for an association between straight MWF exposure and increased risk of female breast cancer, more research is needed. In addition to more studies in general, evaluation of associated risk in subsets of breast cancer cases, such as by tumor hormone responsiveness, may be useful. Additionally, further exploration of the relationship between synthetic fluids and premenopausal breast cancer is recommended, preferably with more certain classification of menopausal status.

Results from chapter 4 suggest a slightly elevated risk of lung cancer related to straight MWF exposure, albeit with a wide confidence interval. Most cohort-based studies of this relationship have found a null association.[5,6,14] The one study to observe a positive result applied g-estimation, an alternative g-method, in this same cohort also to control for the HWSE.[12] Some of the prior studies may have been biased due to the HWSE, and only when the appropriate analytical method is applied is the modestly positive association observed.

Future investigators examining the relationship between straight MWF exposure and lung cancer risk are highly recommended to first evaluate the presence of the HWSE in their cohort and then decide upon statistical methods for the main analysis in order to reduce likelihood of bias from this source.

5.3 References

1. SEER. Cancer Statistics Factsheets: Female Breast Cancer. http://seer.cancer.gov/statfacts/html/breast.html Accessed November 7, 2016.
2. SEER. Cancer Statistics Factsheets: Lung and Bronchus Cancer. http://seer.cancer.gov/statfacts/html/lungb.html Accessed November 7, 2016.
3. Steenland K, Burnett C, Lalich N, Ward E, Hurrell J. Dying for work: The magnitude of US mortality from selected causes of death associated with occupation. *Am J Ind Med* 2003;43(5):461-82.
4. Eisen EA, Tolbert PE, Monson RR, Smith TJ. Mortality studies of machining fluid exposure in the automobile industry I: A standardized mortality ratio analysis. *Am J Ind Med* 1992;22(6):809-24.
5. Tolbert PE. Oils and cancer. *Cancer Causes Control* 1997;8(3):386-405.
6. Savitz DA. Epidemiologic evidence on the carcinogenicity of metalworking fluids. *Appl Occup Environ Hyg* 2003;18(11):913-20.
7. Calvert GM, Ward E, Schnorr TM, Fine LJ. Cancer risks among workers exposed to metalworking fluids: a systematic review. *Am J Ind Med* 1998;33(3):282-92.
8. Robins J. A New Approach to Causal Inference in Mortality Studies with a Sustained Exposure Period - Application to Control of the Healthy Worker Survivor Effect. *Mathematical Modelling* 1986;7(9-12):1393-1512.
9. Robins J. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. *J Chronic Dis* 1987;40 Suppl 2:139S-161S.
10. Garcia E, Picciotto S, Costello S, Bradshaw PT, Eisen EA. Assessment of the healthy worker survivor effect in cancer studies of the United Autoworkers-General Motors cohort. *Occup Environ Med* 2017;74(4):294-300.
11. Buckley JP, Keil AP, McGrath LJ, Edwards JK. Evolving methods for inference in the presence of healthy worker survivor bias. *Epidemiology* 2015;26(2):204-12.
12. Chevrier J, Picciotto S, Eisen EA. A comparison of standard methods with g-estimation of accelerated failure-time models to address the healthy-worker survivor effect: application in a cohort of autoworkers exposed to metalworking fluids. *Epidemiology* 2012;23(2):212-9.
13. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11(5):550-60.
14. NIOSH. Criteria for a Recommended Standard: Occupational Exposure to Metalworking Fluids. Vol. DHHS Publication No. 98-102. Cincinnati, OH: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health,, 1998.

Dissertation Publications

1. Assessment of the healthy worker survivor effect in cancer studies of the United Autoworkers-General Motors cohort. (Chapter 2)

   Garcia E, Picciotto S, Costello S, Bradshaw PT, Eisen EA.

   Occup Environ Med 2017;74(4):294-300.

2. Breast Cancer Incidence and Metalworking Fluid Exposure in a Cohort of Female Autoworkers. (Chapter 3)

   Garcia E, Bradshaw PT, Eisen EA.

   Submitted to American Journal of Epidemiology, March 3, 2017