Application of Probabilistic Bias-Analysis Methods to Historical Cohort Mortality Studies

A DISSERTATION SUBMITTED TO THE FACULTY OF THE UNIVERSITY OF MINNESOTA

BY

Laura L. Scott

IN PARTIAL FULTILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

George Maldonado, PhD, Advisor

December 2018

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Dr. George Maldonado, for his mentorship, enduring patience, and willingness to let me control the direction of my dissertation research. Dr. Maldonado's knowledge regarding epidemiological methods and statistics is matched only by his affable personality, and I am truly fortunate to have had the opportunity to work with him. I would also like to thank my committee members – Drs. Richard MacLehose, Jeffrey Mandel and Timothy Church – for their guidance and thought-provoking suggestions. I am equally thankful to the National Institute of Occupational Safety and Health (NIOSH) Training Program (Grant # T42/OH008434) and the University of Minnesota School of Public Health Dean's Scholars Program, which provided funding for my research.

.

DISCLAIMER

The contents of this document are solely the responsibility of the Laura L. Scott and do not necessarily represent the official views of the National Institute for Occupational Safety and Health (NIOSH)

DEDICATION

This work is dedicated to my family, particularly my husband, for their unwavering encouragement and support throughout this journey and to the memory of my uncle Thomas Ferriby and my grandfather William Arnold, both of whom will forever inspire my human experience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter II

Chapter III

Chapter IV

CHAPTER I

INTRODUCTION

A number of epidemiologic study designs are available for evaluating exposure-disease relationships in occupational settings, yet the historical cohort mortality (HCM) study is one of the most commonly used [1, 2]. Generally, the HCM study is frequently chosen over other study designs because it typically requires less time to complete, is inexpensive compared to other types of studies, and is well-suited for evaluating multiple outcomes and occurrences of rare diseases. As with all observational research, however, results from HCM studies are vulnerable to various types of systematic error, including confounding, selection bias, and information bias such as disease and exposure misclassification.

The effect of systematic study error on findings from observational research, particularly with regard to causal inference, has long been a concern of epidemiologists and research methodologists [3-14]. Because biased estimates of the exposure-disease relationship can critically influence the direction of future research as well as changes to public health policy, researchers have developed a variety of methods to adjust for different types of systematic error. One such method is probabilistic bias analysis (also known as probabilistic uncertainty analysis or probabilistic sensitivity analysis) [3-6, 8-12, 15].

Although detailed methods for conducting bias analyses have been published in recent years [5, 9, 11, 15-19], application of this method to adjust for different sources of systematic error in HCM studies has not been previously described. This is not unexpected, however, given that conducting quantitative bias analyses can be demanding and also repeatedly presents distinct challenges. This is particularly true for historical cohort mortality studies. Specifically, to conduct a bias analysis of an HCM study one must use a measure which does not rely on person-time (*i.e.*, odds ratio, incidence proportion ratio, *etc.*), which results in disease misclassification due to loss to follow-up. In addition, when evaluating a disease specific mortality as the outcome, sources of non-cases must be separated so that deceased individuals are not reclassified as alive. Lastly, most occupational exposures are not binary, but probabilistic bias analysis methods for studies with polytomous exposure variables have not been fully developed.

The research presented here focuses on the development of probabilistic bias analysis methods to adjust results from HCM studies with multiple levels of exposure for three types of systematic error: 1) disease misclassification due to loss to follow-up, 2) disease misclassification due to incorrect death data, and 3) exposure misclassification. While these errors do not occur as mutually exclusive events, a bias-analysis method was developed separately for each type of error to simplify application and ensure proper implementation. Specifically, using published mortality data on a historical cohort of trichlorophenol workers from

New Zealand, this research provides a general framework for conducting these types of analyses on HCM studies and describes each probabilistic bias analysis in detail to illustrate how differential and non-differential misclassification can impact an estimate of effect.

CHAPTER II

QUANTIFYING AND ADJUSTING FOR DISEASE MISCLASSIFICATION DUE TO LOSS TO FOLLOW-UP IN HISTORICAL COHORT MORTALITY STUDIES

Laura L. F. Scott and George Maldonado

Division of Environmental Health Sciences, University of Minnesota School of Public Health, Minneapolis, MN, 55455

Summary

The purpose of this analysis was to quantify and adjust for disease misclassification from loss to follow-up in a historical cohort mortality study of workers where exposure was categorized as a multi-level variable. Disease classification parameters were defined using 2008 mortality data for the New Zealand population and the proportions of known deaths observed for the cohort. The probability distributions for each classification parameter were constructed to account for potential differences in mortality due to exposure status, gender, and ethnicity. Probabilistic uncertainty analysis (bias analysis), which uses Monte Carlo techniques, was then used to sample each parameter distribution 50,000 times, calculating adjusted odds ratios ($OR_{DM\text{-}LTF}$) that compared the mortality of workers

with the highest cumulative exposure to those that were considered never-exposed. The geometric mean $OR_{DM\text{-}LTF}$ ranged between 1.65 (certainty interval (CI): 0.50–3.88) and 3.33 (CI: 1.21–10.48), and the geometric mean of the disease-misclassification error factor ($\mathcal{E}_{DM\text{-}LTF}$), which is the ratio of the observed odds ratio to the adjusted odds ratio, had a range of 0.91 (CI: 0.29–2.52) to 1.85 (CI: 0.78–6.07). Only when workers in the highest exposure category were more likely than those never-exposed to be misclassified as non-cases did the $OR_{DM\text{-}LTF}$ frequency distributions shift further away from the null. The application of uncertainty analysis to historical cohort mortality studies with multi-level exposures can provide valuable insight into the magnitude and direction of study error resulting from losses to follow-up.

Keywords: probabilistic bias analysis; Monte Carlo; disease misclassification; loss to follow-up; historical cohort mortality

Introduction

Epidemiologists have several means available with which to evaluate the exposure-disease relationship in occupational settings. One of the most frequently used methods is the historical cohort mortality study [1, 2]. In general, this type of study offers several benefits in that it typically requires less time to complete, is inexpensive compared to other types of studies, and is well-suited for evaluating multiple outcomes and occurrences of rare diseases. However, historical cohort studies are also vulnerable to loss to follow-up, with one method of addressing this being to withdraw lost individuals from the analysis at the time of loss [1, 20, 21].

Detailed methods for conducting bias analysis (probabilistic uncertainty analysis) of epidemiologic studies have been described by a number of researchers and methodologist [5, 9, 11, 15-19]. Historical cohort mortality studies, however, provide distinct challenges for quantifying study error. First, these types of studies commonly use person-time as the denominator of disease frequency measures; yet bias analysis methods for adjusting an estimate with a person-time denominator have not been described in the peer-reviewed literature. Consequently, to conduct a bias analysis of a cohort mortality study one must use a measure which does not rely on person-time (*i.e.*, odds ratio, incidence proportion ratio, *etc.*). Although using one of these measures would seem to be a simple solution, it introduces another issue unique to historical cohort mortality studies: disease misclassification due to loss to follow-up, which can result from counting those lost to follow-up

who have died as alive at the end of a study. Bias analysis methods to account for this type of disease misclassification have not been previously described. Here, we describe such a method and illustrate how it can be applied to a historical cohort mortality study of New Zealand trichlorophenol workers.

Methods

Error Term for Disease Misclassification Due to Losses

In 2008, Maldonado [5] detailed the mathematical relationship between a causal relative risk, an observed relative risk, and error terms for study bias. We have provided a modification of this relationship (Equation 1), where $OR_{DM\text{-}LTF}$ is the odds ratio adjusted for disease misclassification due to loss to follow-up, $OR_{observed}$ is the observed crude odds ratio, and ε_i are the terms which quantify the systematic error in a study. Because in this manuscript only one error is being evaluated, the denominator has been simplified to $\varepsilon_{DM\text{-}LTF}$, the error term for disease misclassification due to loss to follow-up. $\varepsilon_{DM\text{-}LTF}$ is calculated by taking the ratio of the observed odds ratio to the adjusted odds ratio.

$$OR_{DM-LTF} = \frac{OR_{observed}}{\prod_{i=1}^{n} \varepsilon_i} = \frac{OR_{observed}}{\varepsilon_{DM-LTF}} \quad (1)$$

Crude Odds Ratio

Using the mortality data described by McBride *et al.* [22], we calculated a crude odds ratio for the association between ischemic heart disease (IHD) mortality and exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD). The odds of IHD death for the group with the highest TCDD exposure was 14/148 = 0.0946, and the odds of IHD death for "never-exposed" workers was 14/451 = 0.0310, giving an observed odds ratio of 3.05 and a 95% confidence interval of 1.42–6.54 (Table 1).

Table 1. Cell counts used to estimate the crude odds ratio and 95% confidence limits for the association between occupational TCDD exposure and ischemic heart disease using data reported by McBride *et al.* [22].

	TCDD Exposure		
Outcome	**≥2085.8 ppt-mo**	**0–2085.7 ppt-mo**	**Never-Exposed**
IHD Cases	14	47	14
Non-cases	148	925	451
Alive	112	826	414
Deceased[a]	36	99	37

TCDD, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin; ppt-mo, parts per trillion-month; IHD, ischemic heart disease

[a] From causes of death other than IHD.

Number of All-Cause Deaths among Losses to Follow-up

To estimate the number of workers lost that could have died from IHD for each exposure category, we used a multi-step process (Figure 1). First, we defined a probability distribution for the total number of those lost to follow-up that may have died from any cause for all exposure levels combined. A total of 338 individuals (~21% of the cohort) were lost to follow-up in the cohort mortality study published by McBride and colleagues [22]. We assumed that anywhere from

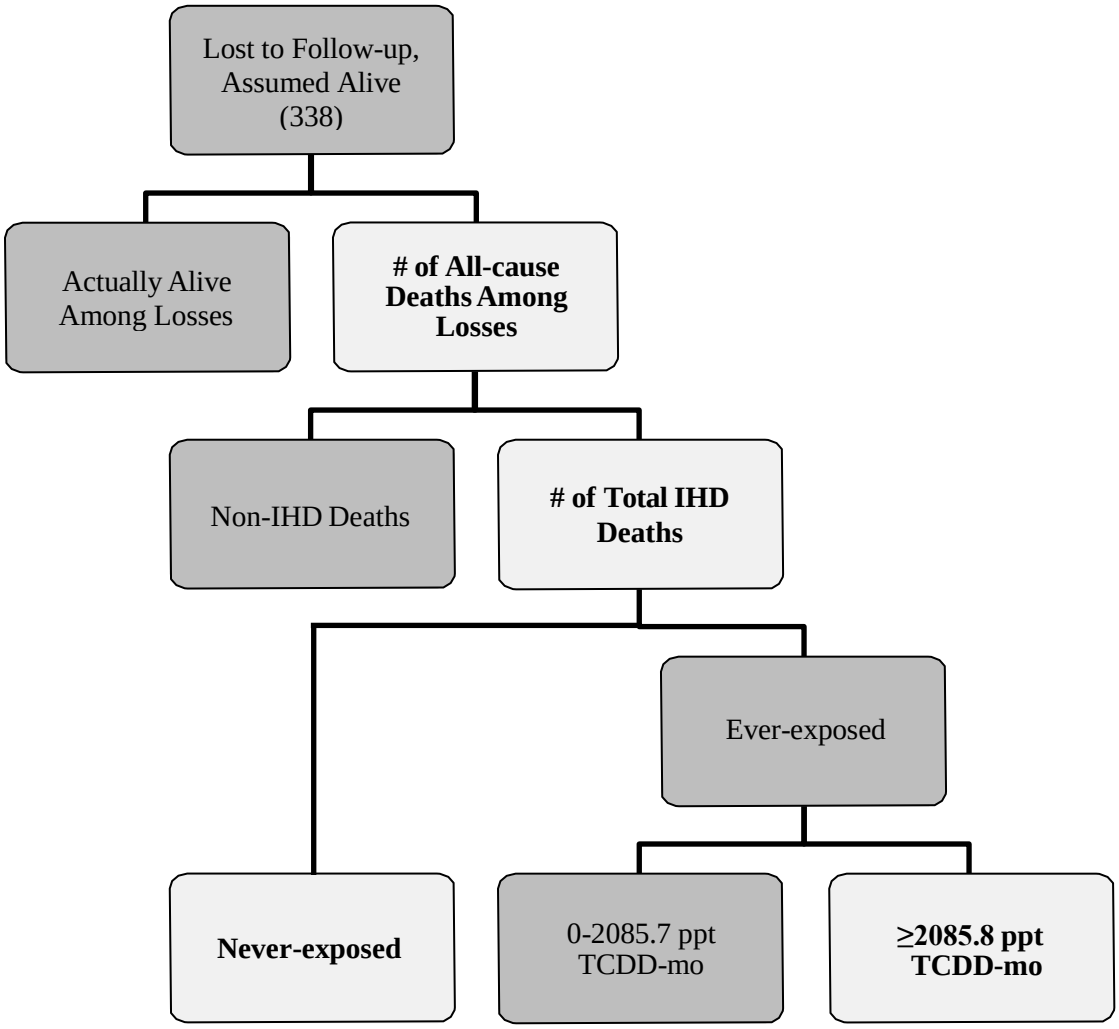


Figure 1. Flow diagram describing how losses to follow-up in McBride *et al.* [22] could result in outcome misclassification. Lighter shapes with bolded text indicate the parameters that were specified in our bias analysis.

zero to 338 individuals might have died from any cause. Therefore, the minimum and maximum of the probability distribution were set to zero and 338, respectively. We specified the peak of this probability distribution by using the proportions of known deaths observed for the cohort. These ranged from 11.0%, the observed proportion of all deaths in the never-exposed category, to 30.9%, the observed

proportion of all deaths in the highest exposure category. The peak number of all-cause deaths for this probability distribution was estimated by multiplying each proportion by the total number of individuals lost. More specifically, 30.9% of 338 provided a peak value of 104 (Scenarios 1–4 in Table 2) and 11.0% of 338 provided a peak value of 37 (Scenarios 5–8 in Table 2). We chose a negative binomial distribution — a discrete distribution with more flexibility than the Poisson distribution for providing the desired shape of the probability distribution — with lower and upper truncation points of 0 and 338, respectively, to describe the spread of the number of all-cause deaths in those workers that were lost. These minimum, maximum and peak values determined the probability and shape input of all-cause deaths for each bias-analysis scenario.

Number of Total IHD Deaths among Losses to Follow-Up

Next, we used 2008 mortality data for the New Zealand population [23] to estimate the total number of deaths from IHD for all exposure levels combined. The proportion of New Zealanders who died from IHD varied by both gender and ethnicity, with the proportion of IHD deaths the highest in non-Maori males (20.4%) and the lowest in Maori females (13.9%). The BetaPERT (*i.e.*, PERT) distribution, which is derived from the beta distribution and is a smoother alternative to the triangular distribution, was specified as the probability distribution for this disease-classification parameter. We selected this distribution over the

Table 2. Bias-analysis scenarios: description of probability distributions for classification parameters used to estimate the number of workers lost to follow-up that could have died from IHD and corresponding geometric mean errors ($\mathcal{E}_{DM\text{-}LTF}$), adjusted odds ratios ($OR_{DM\text{-}LTF}$) and 95% bias-analysis certainty intervals.

	Total All-Cause Deaths	Total IHD Deaths	IHD Deaths by Exposure Status			$\mathcal{E}_{DM\text{-}LTF}$		$OR_{DM\text{-}LTF}$	
				Never-exposed	≥2085.8 ppt-mo		95%		95%
Scenario	Distribution (Parameters)	Distribution (Parameters)	Direction of Misclassification	Distribution (Parameters)	Distribution (Parameters)	GM	Certainty Interval	GM	Certainty Interval
1	Negative Binomial[a] (0.02, 3)	BetaPERT[b] (0, 0.204[c] × AD, AD)	Differential A[d]	BetaPERT (0, 3/4 × ID, ID)	BetaPERT (0, 1/2 × IDE, IDE)[e]	1.85	0.78–6.07	1.65	0.50–3.88
3	Negative Binomial[a] (0.02, 3)	BetaPERT (0, 0.139[c] × AD, AD)	Differential A	BetaPERT (0, 3/4 × ID, ID)	BetaPERT (0, 1/2 × IDE, IDE)	1.62	0.74–4.93	1.88	0.62–4.11
5	Negative Binomial[f] (0.027, 2)	BetaPERT (0, 0.204 × AD, AD)	Differential A	BetaPERT (0, 3/4 × ID, ID)	BetaPERT (0, 1/2 × IDE, IDE)	1.50	0.87–3.92	2.03	0.78–3.51
7	Negative Binomial[f] (0.027, 2)	BetaPERT (0, 0.139 × AD, AD)	Differential A	BetaPERT (0, 3/4 × ID, ID)	BetaPERT (0, 1/2 × IDE, IDE)	1.43	0.87–3.65	2.13	0.83–3.50
2	Negative Binomial[a] (0.02, 3)	BetaPERT (0, 0.204 × AD, AD)	Differential B[g]	BetaPERT (0, 1/4 × ID, ID)	BetaPERT (0, 1/2 × IDE, IDE)	0.91	0.29–2.52	3.33	1.21–10.48
4	Negative Binomial[a] (0.02, 3)	BetaPERT (0, 0.139 × AD, AD)	Differential B	BetaPERT (0, 1/4 × ID, ID)	BetaPERT (0, 1/2 × IDE, IDE)	0.92	0.32–2.37	3.31	1.29–9.65
6	Negative Binomial[f] (0.027, 2)	BetaPERT (0, 0.204 × AD, AD)	Differential B	BetaPERT (0, 1/4 × ID, ID)	BetaPERT (0, 1/2 × IDE, IDE)	0.95	0.42–1.98	3.20	1.54–7.26
8	Negative Binomial[f] (0.027, 2)	BetaPERT (0, 0.139 × AD, AD)	Differential B	BetaPERT (0, 1/4 × ID, ID)	BetaPERT (0, 1/2 × IDE, IDE)	0.96	0.46–1.86	3.18	1.64–6.64

AD, number of total all-cause deaths; ID, number of total IHD deaths; IDE, number of IHD deaths for those workers ever-exposed;

[a] Negative binomial distribution (probability, shape)—probability and shape were determined based on minimum, likeliest and maximum counts of (0, 104, 338).

[b] BetaPERT distribution (minimum, likeliest, maximum).

[c] 0.204: proportion of all-cause deaths due to IHD among non-Maori males; 0.139: proportion of all-cause deaths due to IHD among Maori females.

[d] Never-exposed more likely to be misclassified as alive than highest exposed.

[e] The maximum value for this distribution is capped at 112, which is the number of individuals in the highest exposure group (*i.e.*, ≥2085.8 ppt-mo) that were classified as living non-cases.

[f] Negative binomial distribution (probability, shape)—probability and shape were determined based on minimum, likeliest and maximum counts of (0, 37, 338).

[g] Never-exposed less likely to be misclassified as alive than highest exposed.

negative binomial and Poisson distributions because (1) it is considered to be ideal for modeling expert opinion of a variable [24], (2) it was much more flexible than the Poisson distribution, and (3) the maximum and likeliest values of the distribution, which were dependent on the total number of all-cause deaths selected in the first step, could easily be varied. For example, if the number of all-cause deaths was 100 for a bias-analysis simulation trial, then the distribution for the number of deaths from IHD would range from 0 to 100 with a likeliest value of 13.9, assuming 13.9% of all-cause deaths were due to IHD. Since the BetaPERT distribution is continuous and we are interested in estimating discrete counts, we used the TRUNC function in Excel to remove the decimal portion of each bias-analysis simulation trial value. The difference in the adjusted odds ratios and error terms estimated with and without use of the TRUNC function was negligible.

Number of IHD Deaths among Losses to Follow-Up: Never-Exposed

For the next step, we specified a probability distribution for the number of IHD deaths for the "never-exposed" group. The BetaPERT distribution, along with the TRUNC function in Excel, was also used for this parameter. The maximum and likeliest values were adjusted in a similar manner to that for the total number of IHD deaths. The maximum was set to equal the total number of IHD deaths selected in the previous bias-analysis simulation step. Since the total number that died of IHD could be categorized into one of three exposure groups (*i.e.*, "never-

exposed", "0–2085.7 ppt-mo", "≥2085.8 ppt-mo"), we used simple fractions to determine the likeliest value of this probability distribution depending on whether IHD deaths among the "never-exposed" were (1) more likely to be misclassified as those in the highest exposure group (Differential A) or (2) less likely to be misclassified as those in the highest exposure group (Differential B). When it was assumed that the "never-exposed" were more likely to be misclassified compared to those in the highest exposure group, the likeliest value was set to equal 3/4 the total number of IHD deaths. Under the second assumption, the likeliest value was set to equal 1/4 the total number of IHD deaths. For example, if the total number of IHD deaths selected in the second step of a bias-analysis simulation trial is 24, then the likeliest values for the "never-exposed" group would be 18 and six, respectively.

Number of IHD Deaths among Losses to Follow-Up: ≥2085.8 ppt TCDD-mo

Last, a probability distribution for the number of IHD deaths among workers with the highest exposure (*i.e.*, ≥2085.8 ppt TCDD-mo) was specified. For this parameter, we again chose to use the BetaPERT distribution, truncating each trial value at the decimal point to obtain a whole number. Using the IF function in Excel, the distribution maximum was set to equal the number of IHD deaths for those "ever-exposed" up to 112, which is the number of individuals in the highest exposure group that were classified as living non-cases, and equal to 112 when the total

number of "ever-exposed" that were potentially misclassified as alive (*i.e.*, died of IHD) was greater than 112. The likeliest value of the probability distribution for the highest exposure group was set to equal 1/2 the number of IHD deaths for the "ever-exposed" workers so that approximately 3/8 (*i.e.*, 3/4 × 1/2) and 1/8 (*i.e.*, 1/4 × 1/2) of the total number of IHD deaths would fall into the highest exposure category.

For example, the probability distribution input used for Scenario 1 is shown in Figure 2. In this scenario, the distribution for the total number of all-cause deaths among those lost to follow-up ranges from 0 to 338 and peaks at 104 (30.9% of the number lost to follow-up). Assuming 104 all-cause deaths, the distribution for the total number of IHD deaths would then range from 0 to 104, with the highest probability at 21.2 or 20.4% of 104. Given 21 total IHD deaths, the likeliest number of workers misclassified as alive would then be 15 for the "never-exposed" category (3/4 × # IHD deaths = 3/4 × 21 ≈ 15). For the highest exposure group, the likeliest number of workers lost to follow-up that may have died from IHD would be three ((# total IHD deaths – # never-exposed IHD deaths) × 1/2 = (21 − 15) × 1/2 = 3).

Scenarios and Monte Carlo Simulation Methods

Combining the different distributions for each classification parameter resulted in eight scenarios (Table 2). The distributions of the classification parameters were sampled with Crystal Ball software [25] to generate adjusted

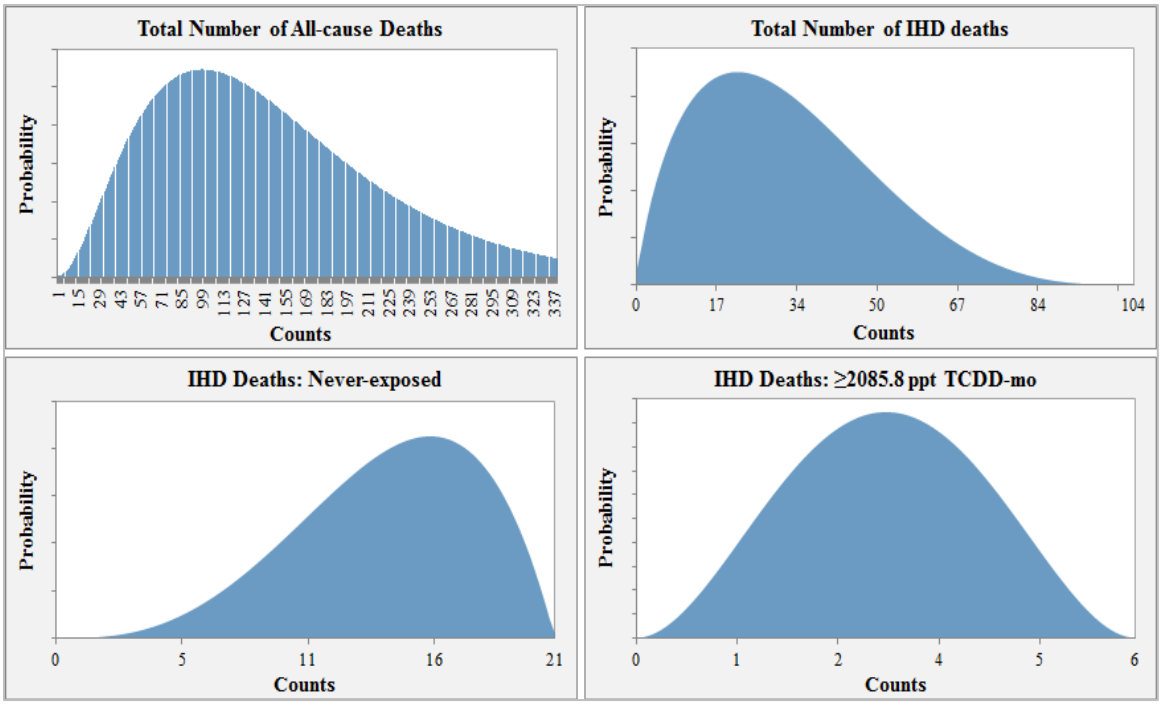


Figure 2. Example of parameter distribution input for Scenario 1.

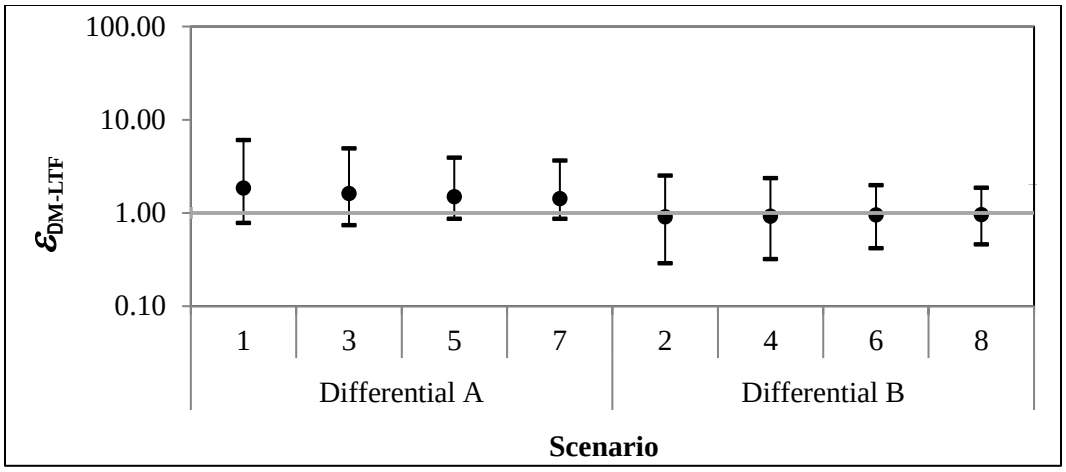
counts of IHD cases and non-cases by exposure group, which we used to calculate adjusted odds ratios and study error. Since Crystal Ball will not allow the program to proceed when zero (or a number truncated to zero) is selected for a given classification parameter, we created a conditional action—using the IF function in Excel—whereby the program would automatically set the ensuing classification parameter values to zero anytime this occurred. For each of the eight scenarios, we conducted 50,000 trials to generate frequency distributions for $OR_{DM\text{-}LTF}$ and $\mathcal{E}_{DM\text{-}LTF}$ as well as 95% certainty intervals. Under specific conditions, a 95% certainty interval may approximate a 95% Bayesian posterior probability interval, such that there is a 95% chance that the true estimate for the sample population will fall within the

interval [10, 26, 27]. This interpretation is different from that of a 95% confidence interval, which is defined as a range of values that will include the true parameter value 95% of the time.

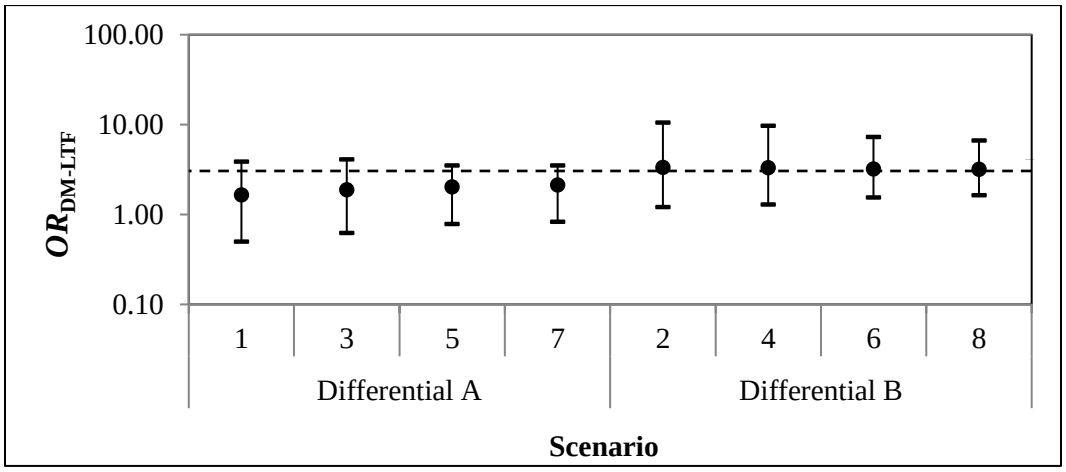
Results

Results for each simulation of the probabilistic uncertainty analysis are summarized in Table 2 and Figure 3. The geometric mean of the error term for disease misclassification due to loss to follow-up ($\mathcal{E}_{DM\text{-}LTF}$) had a range of 0.91 to 1.85. The geometric mean adjusted odds ratio ($OR_{DM\text{-}LTF}$) ranged between 1.65 and 3.33. Estimated certainty intervals (CI) for the geometric mean $OR_{DM\text{-}LTF}$ excluded the null for all four scenarios in which those categorized as "never-exposed" were less likely to be misclassified as alive than workers in the highest exposure category.

The direction of the error was primarily determined by the exposure classification parameters. In the four scenarios where workers with the highest exposure were more likely than those never-exposed to be misclassified as alive, the adjusted OR moved away from the null (*i.e.*, the crude OR was biased toward the null). In contrast, when the never-exposed group had a greater proportion misclassified as non-cases than the highest exposure group (Scenarios 1, 3, 5, and 7), adjustment for study bias due to loss to follow-up resulted in a shift of the $OR_{DM\text{-}LTF}$ frequency distributions toward the null, lessening the observed effect for the



(A)



(B)

Figure 3. Geometric mean errors ($\mathcal{E}_{DM\text{-}LTF}$) (**A**), adjusted odds ratios ($OR_{DM\text{-}LTF}$) (**B**) and 95% certainty intervals by scenario. The dashed horizontal black line in (**B**) indicates the crude odds ratio ($OR_{observed}$) of 3.05. In the Differential A scenarios, the "never-exposed" were more likely to be misclassified as alive than the highest exposed. In the Differential B scenarios, the "never-exposed" were less likely to be misclassified as alive than the highest exposed.

exposure-disease relationship (Figure 4).

In all the scenarios we examined, uncertainty about the amount of disease misclassification due to assuming that lost subjects were alive resulted in uncertainty about the magnitude of the TCDD-IHD association estimate.

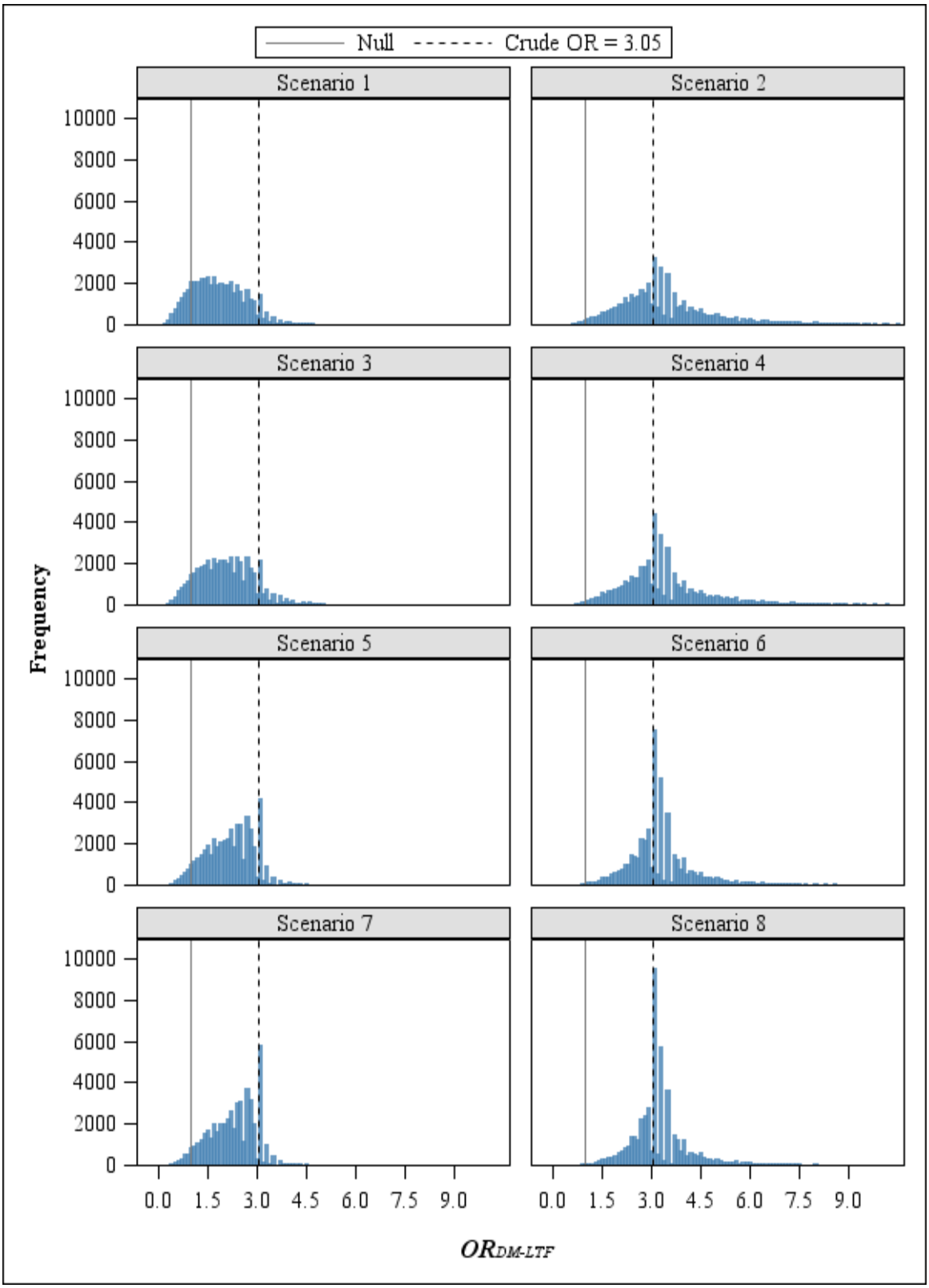


Figure 4. Frequency distributions of $OR_{DM\text{-}LTF}$ by scenario.

Discussion

Here we demonstrated that losses to follow-up in a historical cohort mortality study can cause disease misclassification. This source of disease misclassification is different from the misclassification that may result from using death certificate or other mortality data and is unique to historical cohort mortality studies. Currently, methods using person-time to conduct probabilistic uncertainty analyses have not been published. As such, one must use either the incidence proportion ratio or the odds ratio as the estimator when conducting bias analyses of these studies. When using either of these measures, however, problems those lost to follow-up are consequentially counted as alive even though some may have died from the disease of interest. For both measures, study subjects contribute to the denominator but have no opportunity to contribute to the numerator. The only way to prevent this would be to exclude anyone who was lost, but this creates other potential problems and has never been an accepted method for addressing losses to follow-up.

Since we had no reason to assume that disease misclassification caused by losses to follow-up would not affect the association measure for occupational exposure to TCDD and IHD mortality, we developed a probabilistic bias analysis method to adjust for this source of systematic error using Monte Carlo simulations. In developing this method, we encountered several issues that challenged the implementation of our approach. First, very little information was available from the original study [22] on how losses were distributed across the

exposure categories. This would likely be easier to address when the exposure variable is dichotomous, but can become particularly complicated with multi-level exposures. We believed that the best method of making the process more manageable was to divide it into multiple steps, which also allowed us to check the accuracy of our process at different points. In addition, it was necessary to calculate the number of living subjects in each exposure group and set this as the maximum distribution value when the total number for a given step exceeded the number of living that could have been misclassified as alive. For example, 148 workers were categorized as non-cases with exposure ≥2085.8 ppt TCDD-mo, but only 112 of these employees were classified as alive at the end of the study. The other 36 subjects died of causes other than IHD and could not have been potentially misclassified as alive due to losses to follow-up. Accordingly, we used the IF function in Excel to set the distribution maximum for the number of IHD deaths in the exposure group "≥2085.8 ppt TCDD-mo" equal to 112 when the total number of "ever-exposed" that were potentially misclassified as alive (*i.e.*, died of IHD) was greater than 112. Otherwise, the maximum value for this parameter equaled the number of IHD deaths among those lost that were "ever-exposed."

We also addressed a variety of concerns with regard to the simulation software used and the distributions available. For the number of all-cause deaths among losses, we chose a negative binomial distribution because it was discrete and provided the flexibility we needed to describe the distribution for this

parameter. Unfortunately, the negative binomial distribution is determined by probability and shape, which are not easily varied, so it was not an ideal distribution for the remaining three classification parameters — each of which are dependent on the preceding step. We initially considered the Poisson distribution because the peak (λ) could be determined based on the trial value of another classification parameter. However, this distribution was not as flexible as we had hoped. For example, if the number of all-cause deaths among losses was 104, the distribution for the total number of IHD deaths would need to peak at 21, assuming a proportion of 20.4%, and have a range of 0 to 104. With a Poisson distribution, the probability of the simulation choosing a count less than 9 or greater than 37 was zero, effectively excluding the majority of potential values. Rather, we chose to use the continuous BetaPERT distribution — which is renormalized over a finite range other than (0,1), re-parameterized by the minimum, maximum, and mode [24], and very flexible — and truncate the selected trial values to whole numbers. Comparison of the adjusted odds ratios and error terms estimated with and without use of the TRUNC function demonstrated any differences were negligible. Last, it should be noted that the simulation software we used would suspend a run anytime zero was selected as a trial value for the first two parameters or when the total number "ever-exposed" that were potentially misclassified as alive was zero. To address this, we used the IF function in Excel to set the subsequent classification parameter values to zero when

any of the described conditions occurred (*e.g.*, if a trial value for the total number of IHD deaths among losses was zero, then all successive parameter values for that trial would be zero). We believe this is a superior alternative to excluding these trials.

In the example described here, we demonstrated that the magnitude of the error from disease misclassification due to loss to follow-up can be substantial. The certainty intervals for $OR_{DM\text{-}LTF}$ in the simulations for this bias analysis were quite wide, with the distance between the upper and lower bounds ranging from 2.67 to 9.27. While the exposure parameters ("IHD deaths: never-exposed" and "IHD deaths: ≥2085.8 ppt-mo") were the main determinants of the location of the geometric mean adjusted odds ratios, the width of the certainty intervals was likely influenced more by the degree of misclassification.

By using a probabilistic method that specifies a range of values for the classification parameters under a variety of thoughtfully constructed scenarios, our approach allows one to estimate intervals that may better represent the level of uncertainty from systematic study error in a given exposure-disease relationship. Compared to other techniques such as simple sensitivity analysis and inverse probability weighting (IPW), use of probabilistic bias analysis may be more advantageous because it does not rely on conditional weighting estimates or single sensitivity and specificity values. One of the primary limitations of IPW is that the model estimating the weights is built as a function of subject characteristics (*e.g.*,

age, gender). Anything that might be related to follow-up should be part of the model building process. Yet, when characteristics that may be predictive of follow-up are not measured or collected as part of the study, use of this method may actually introduce bias. Our method, however, allows one to use external data to adjust for any characteristic — whether or not measured as part of the initial study — that may affect follow-up or the outcome of interest.

Our primary objectives were to illustrate how to adjust an odds ratio for disease misclassification resulting from losses to follow-up in a historical cohort mortality study and to evaluate the effect of this source of error. It is likely, therefore, that uncertainty about the magnitude of other study limitations; for example, exposure misclassification would further increase the uncertainty about the TCDD-IHD association. Additionally, the method described here is contingent on the distributions constructed for each of the classification parameters, with the usefulness of the results limited by the accuracy of those distributions. For example, we used New Zealand mortality data from 2008, which assumes the proportion of deaths from IHD was constant over the entire study period evaluated by McBride *et al.* [22]. However, the New Zealand Ministry of Health reported that the percentage of IHD deaths declined over the 1980–2008 time frame [23] and we, therefore, may have underestimated the number lost to follow-up that died from IHD and the level of uncertainty.

Conclusions

In summary, our method can be employed to quantify and adjust for disease misclassification due to losses to follow-up in any historical cohort mortality study with a defined polytomous exposure variable. Use of our probabilistic bias method to adjust for this source of systematic error may well be a considerable improvement over the standard conjecture that, most often, incorrectly assumes such error would be non-differential and have little or no effect on the observed exposure-disease relationship.

CHAPTER III

A PROBABILISTIC BIAS-ANALYSIS METHOD FOR EVALUATING DISEASE MISCLASSIFICATION IN A HISTORICAL COHORT MORTALITY STUDY

Laura L. F. Scott, Jeffrey Mandel, George Maldonado

Division of Environmental Health Sciences, University of Minnesota School of Public Health, Minneapolis, MN, 55455

Summary

Occupational epidemiologists have long considered death certificate inaccuracies – and the subsequent disease misclassification they cause in epidemiologic studies – a critical issue. However, the vast majority of historical cohort mortality studies do not include a quantitative assessment of the impact of disease misclassification on study results. We developed a probabilistic bias-analysis method to evaluate the effect of disease misclassification due to incorrect cause-of-death data on results from historical cohort mortality studies. The method is illustrated by applying classification proportions to count data from a historical cohort mortality study of New Zealand workers. When misclassification was differential, the geometric mean adjusted odds ratio ranged from 1.9 to 4.9 with

study error resulting in bias both away from and toward the null. Under the assumption of non-differential misclassification, the geometric mean adjusted odds ratio was slightly smaller than the unadjusted estimate. Probabilistic bias analysis can be a useful tool for evaluating disease misclassification.

Key words: probabilistic bias analysis, Monte Carlo, cause-of-death data, disease misclassification, study error, historical cohort mortality

Introduction

Death certificates are a vital source of cause-of-death data and are commonly used for surveillance, epidemiologic research, and developing public health policy. Death certificate data, however, is known to be highly variable and inaccurate for several reasons. Chiefly, records are often incomplete or vague [3, 28-30]. In addition, changes in terminology, coding rules and diagnosis criteria, along with the progression of multiple chronic diseases in an individual, can all result in inaccuracies [31-33]. Finally, cause of death may also be miscoded when sudden accidental deaths are a direct result of a chronic condition (*e.g.*, an individual has a fatal heart attack which leads to a car crash).

Concerns regarding inaccurate death certificate data, and therefore disease misclassification, have long been acknowledged by epidemiologists and occupational physicians. Indeed, Dr. Irving Selikoff noted in 1992 that the problem was not only an international issue, but one that particularly affects epidemiology studies of occupational hazards and diseases [32]. Although use of medical records and proxy reports have been suggested as alternatives to using death certificate data [34], this would require spending a substantial amount of resources and time identifying and interviewing proxies and/or obtaining and reviewing potentially hundreds to thousands of documents from multiple sources for each deceased subject. Without a time- and cost-effective option, occupational epidemiologists are likely to continue utilizing cause-of-death data from death

certificates for conducting historical cohort mortality (HCM) studies. Nonetheless, a vast majority of these types of studies will not include a quantitative assessment of the impact of disease misclassification from using incorrect death certificate data.

While a number of efforts have been made in recent years to develop techniques for addressing outcome misclassification [35-38], a probabilistic bias-analysis method to adjust for disease misclassification resulting from use of cause-of-death data in HCM studies has not yet been published. This is not unexpected, however, given that conducting quantitative bias analyses – both basic and probabilistic – can be demanding and often also presents distinct challenges. For example, Jurek *et al*. [18] demonstrated the need to account for case-control sampling and other forms of outcome-related selection when adjusting for outcome misclassification, and Scott and Maldonado [39] recognized that, in historical cohort mortality studies, using disease frequency measures with count denominators can result in disease misclassification due to loss to follow-up. Another situation unique to bias analyses occurs when evaluating the impact of disease misclassification in historical cohort mortality studies: non-cases can include both living subjects and those who died of other causes. More specifically, in a standard bias-analysis of a 2×2 contingency table, any non-cases – including living subjects – could be re-classified as (deceased) cases. However, only non-cases who died from a disease other than the disease under study should be

considered for possible outcome re-classification. Here, we demonstrate the use of a probabilistic bias-analysis method we developed for evaluating the effect of potential disease misclassification in HCM studies that addresses this issue.

Methods

Development of Bias-analysis Method

The bias-analysis method developed in the current study combines Monte Carlo simulation techniques with an extension of the matrix-algebra correction method described by Greenland and Kleinbaum [40]. The data used to illustrate application of this method is from a historical cohort mortality study of New Zealand trichlorophenol workers exposed to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) [22]. Although the study [22] examined multiple diseases, the outcome of interest for this bias-analysis was ischemic heart disease mortality.

To begin, we created a 3×2 disease-exposure contingency table using the three outcome categories (*i.e.*, IHD cases, non-cases who died from causes other than IHD, and alive non-cases) and two exposure groups (*i.e.*, never-exposed and ≥2085.8 parts per trillion-month) of interest (Table 1).

Table 1. 3×2 Disease-exposure contingency table.

	TCDD Exposure	
Outcome	**Never-exposed**	**≥2085.8 ppt-mo**
Cases: Deceased, IHD	a_0	a_1
Non-cases: Deceased, non-IHD	b_0	b_1
Non-cases: Alive	c_0	c_1

TCDD, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin; ppt-mo, parts per trillion-month; IHD, ischemic heart disease

Of particular importance, non-cases in Table 1 have been displayed in two different rows: (i) those who died from causes other than IHD, and (ii) those who were alive. When performing a bias analyses of disease misclassification in studies that use death certificates to determine cause of death (as the current manuscript does), these two types of non-cases must be separated to prevent non-cases who are known to be alive from being reclassified as (deceased) cases.

Next, we reshaped Table 1 into a column vector of six cell counts (C matrix):

$$C = \begin{bmatrix} a_0 \\ a_1 \\ b_0 \\ b_1 \\ c_0 \\ c_1 \end{bmatrix}$$

This vector of cell counts takes two forms: the vector of *observed* cell counts (*i.e.*, the cell counts with disease misclassification) and the vector of cell counts *corrected for disease misclassification* (*i.e.*, the cell counts that would have been observed in the absence of disease misclassification). To differentiate the two, an asterisk is used to denote the vector of observed cell counts with misclassification. More specifically, Table 2 shows the count data from McBride *et al*. [22] in the 3×2 cross tabulation format presented in Table 1.

Table 2. Count data from McBride *et al*. [22]

	TCDD Exposure	
Outcome	**Never-exposed**	**≥2085.8 ppt-mo**
Cases: Deceased, IHD	14	14
Non-cases: Deceased, non-IHD	37	36
Non-cases: Alive	414	112

TCDD, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin; ppt-mo, parts per trillion-month; IHD, ischemic heart disease

Using these counts, C^* is as follows:

$$C^* = \begin{bmatrix} 14 \\ 14 \\ 37 \\ 36 \\ 414 \\ 112 \end{bmatrix}$$

We then specified a column vector of classification proportions for each cell of Table 1; these proportions describe the proportion of people truly in that one cell that were classified into each of the six cells of the table. This resulted in a 6×6 matrix of classification proportions (P matrix; Table 3).

Note that the diagonal elements of the P matrix (*i.e.*, p_{11}, p_{22}, p_{33}, p_{44}, p_{55}, p_{66}) indicate the degree of *correct* classification; all other (*i.e.*, non-diagonal) elements in the matrix indicate the degree of exposure and/or disease *misclassification*. For example, p_{11} is the proportion of subjects who (i) truly were never-exposed and truly died from IHD and (ii) were (correctly) classified in the study data as never-exposed cases of IHD. In contrast, p_{13} is the proportion of subjects who (i) truly were never-exposed and truly died of something other than

Table 3. Notation for matrix of classification proportions (P matrix).

		True Exposure and Disease Status					
		Cases: Deceased, IHD		*Non-cases: Deceased, non-IHD*		*Non-cases: Alive*	
Classified Disease Status	**Classified Exposure Status**	Never-exposed	≥2085.8 ppt-mo	Never-exposed	≥2085.8 ppt-mo	Never-exposed	≥2085.8 ppt-mo
Cases: Deceased, IHD	Never-exposed	p_{11}	p_{12}	p_{13}	p_{14}	p_{15}	p_{16}
	≥2085.8 ppt-mo	p_{21}	p_{22}	p_{23}	p_{24}	p_{25}	p_{26}
Non-cases: Deceased, non-IHD	Never-exposed	p_{31}	p_{32}	p_{33}	p_{34}	p_{35}	p_{36}
	≥2085.8 ppt-mo	p_{41}	p_{42}	p_{43}	p_{44}	p_{45}	p_{46}
Non-cases: Alive	Never-exposed	p_{51}	p_{52}	p_{53}	p_{54}	p_{55}	p_{56}
	≥2085.8 ppt-mo	p_{61}	p_{62}	p_{63}	p_{64}	p_{65}	p_{66}
	Total	1	1	1	1	1	1

IHD, ischemic heart disease; ppt-mo, parts per trillion-month

IHD but (ii) were (incorrectly) classified in the study data as never-exposed cases of IHD.

According to Greenland and Kleinbaum [40] and the rules of matrix multiplication (*e.g.*, $a_0^* = 14 = a_0 p_{11} + a_1 p_{12} + b_0 p_{13} + b_1 p_{14} + c_0 p_{15} + c_1 p_{16}$), the relationship between the column vector of misclassified cell counts (C^*), the column vector of correctly classified cell counts (C), and the P matrix (P) is

$$P \times C = C^* \tag{1}$$

Expansion of Equation (1) produces the matrix form

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & p_{26} \\ p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & p_{36} \\ p_{41} & p_{42} & p_{43} & p_{44} & p_{45} & p_{46} \\ p_{51} & p_{52} & p_{53} & p_{54} & p_{55} & p_{56} \\ p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & p_{66} \end{bmatrix} \times \begin{bmatrix} a_0 \\ a_1 \\ b_0 \\ b_1 \\ c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 14 \\ 14 \\ 37 \\ 36 \\ 414 \\ 112 \end{bmatrix}$$

which highlights the implicit assumption that, for a typical analysis, all the diagonal elements of the P matrix equal one and all other elements equal zero. Stated more explicitly, standard analyses of epidemiologic data assume no misclassification. The bias-analysis method described here relaxes this strong assumption for historical cohort mortality studies in an effort to account for disease misclassification.

The intent of this study was to focus exclusively on the impact of disease misclassification and the unique challenge posed by living non-cases when conducting a bias analysis of historical cohort mortality data, resulting in three

fundamental assumptions: (i) study subjects were not misclassified with regard to exposure, (ii) no living subjects were incorreclty classified as deceased, and (iii) those who were truly deceased were not misclassified as being alive. Table 4 presents a simplified form of the P matrix based on these assumptions.

To obtain the vector of cell counts adjusted for disease misclassification (C), we multiplied each side of Equation (1) by the inverse of the P matrix (P^{-1}):

$$\begin{aligned} P^{-1} \times P \times C &= P^{-1} \times C^{*} \\ C &= P^{-1} \times C^{*} \end{aligned} \tag{2}$$

where the expanded form of the matrix is given by

$$\begin{bmatrix} a_0 \\ a_1 \\ b_0 \\ b_1 \\ c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & p_{26} \\ p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & p_{36} \\ p_{41} & p_{42} & p_{43} & p_{44} & p_{45} & p_{46} \\ p_{51} & p_{52} & p_{53} & p_{54} & p_{55} & p_{56} \\ p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & p_{66} \end{bmatrix}^{-1} \times \begin{bmatrix} 14 \\ 14 \\ 37 \\ 36 \\ 414 \\ 112 \end{bmatrix}$$

From the cell counts a_0, a_1, b_0, b_1, c_0, and c_1, we calculated an unadjusted odds ratio corrected for disease misclassification:

$$OR_{DM-COD} = \frac{a_1 \big/ (b_1 + c_1)}{a_0 \big/ (b_0 + c_0)} \tag{3}$$

Finally, we estimated $\varepsilon_{DM\text{-}COD}$, the error factor for the impact of disease misclassification, by dividing $OR_{DM\text{-}COD}$ into the observed odds ratio ($OR_{observed}$) [5].

Table 4. P matrix with assumptions for fixed parameters.

		True Exposure and Disease Status					
		Cases: Deceased, IHD		*Non-cases: Deceased, non-IHD*		*Non-cases: Alive*	
Classified Disease Status	**Classified Exposure Status**	Never-exposed	≥2085.8 ppt-mo	Never-exposed	≥2085.8 ppt-mo	Never-exposed	≥2085.8 ppt-mo
Cases: Deceased, IHD	Never-exposed	p_{11}	0	$1-p_{33}$	0	0	0
	≥2085.8 ppt-mo	0	p_{22}	0	$1-p_{44}$	0	0
Non-cases: Deceased, non-IHD	Never-exposed	$1-p_{11}$	0	p_{33}	0	0	0
	≥2085.8 ppt-mo	0	$1-p_{22}$	0	p_{44}	0	0
Non-cases: Alive	Never-exposed	0	0	0	0	1	0
	≥2085.8 ppt-mo	0	0	0	0	0	1
	Total	1	1	1	1	1	1

IHD, ischemic heart disease; ppt-mo, parts per trillion-month

Probability Distributions of Classification Proportions for Monte Carlo Simulations

Among deaths, p_{11} and p_{22} are the sensitivities and p_{33} and p_{44} are the specificities for the use of death certificates to classify study subjects as cases (*i.e.*, IHD deaths). To identify sensitivity and specificity values for constructing probability distributions of these parameters, we searched the peer-reviewed literature for death certificate validation studies of ischemic heart disease mortality in New Zealand and other high-income countries as defined by the World Bank [41]. Various combinations of keywords were used to search for relevant English-language articles in PubMed including "ischemic," "coronary," "heart disease," "mortality," "validation," "death certificate," "diagnosis," "ICD," "coding," and "cause of death." A general internet search of non-peer-reviewed publications and review of secondary references from applicable sources were also completed. Each publication was reviewed to identify the country, study population, ICD-9 and -10 codes evaluated (see Table A1 in the Appendix for additional information on ICD codes), and the standard used to calculate sensitivity and specificity. Only one article, a publication from 1988, evaluated IHD death certificate diagnoses in New Zealand [42], with a sensitivity of 0.92 and specificity of 0.47. We identified 13 additional studies from other high-income countries that either reported sensitivity and specificity explicitly for IHD mortality or provided enough information to calculate these values [43-55]. For these studies, death certificate sensitivity and specificity ranged from 0.77 to 0.95 and 0.59 to 0.87, respectively

(Table 5).

We defined the parameter distributions for a non-differential (in expectation) disease misclassification scenario first, with the distribution peaks equal to the sensitivity and specificity values from the New Zealand study and the shape of the distributions constructed using the aggregate sensitivity and specificity data from the remaining studies. A beta distribution was used for the sensitivity parameters and a maximum extreme distribution for the specificity parameters. To evaluate the potential effects of differential misclassification (in expectation), we also varied the distribution peaks for the highest (p_{22}, p_{44}) and lowest (p_{11}, p_{33}) exposure categories (Table 6). For these differential scenarios (*i.e.*, 2-9), the peak sensitivity was set to equal either the sensitivity from the New Zealand study (*i.e.*, 0.92) or the average sensitivity of all 14 studies (*i.e.*, 0.89), and the peak specificity was set to equal either the specificity from the New Zealand study (*i.e.*, 0.47) or the average specificity of all 14 studies (*i.e.*, 0.71). As before, a beta distribution was used for all the sensitivity parameters. However, in addition to using the maximum extreme distribution, the minimum extreme distribution was utilized when the peak specificity was set to 0.71. Figure 1 illustrates the probability input required for any given simulation using the parameter distributions constructed for Scenario 6. To limit the degree of differentiality within simulation trials [56], we designated a correlation of 0.75 between the sampled proportions for each exposure group. For each scenario, the

Table 5. Validation studies of IHD death certificate diagnoses in high income countries.[a]

Reference	Country	Population	Year(s)	ICD Codes	Standard	Sensitivity	Specificity
Jackson *et al*. [42]	New Zealand	Permanent residents of the Auckland area	1983-1984	ICD-9 410-414	WHO criteria[b]	0.92	0.47
Harriss *et a*l. [48]	Australia	Melbourne residents aged 40 to 81 years	1990-2002	ICD-9 410-414; ICD-10 I20-I25	ANBP2[c] and LIPID[d] mortality definitions, WHO criteria	0.77	0.87
Sexton *et al*. [55]	Australia	Tasmanian males aged 25 to 74 years	1987-1988	ICD-9 410-414	WHO criteria	0.94	0.59
Martin *et al*. [52]	Australia	Residents of Perth and the surrounding urban areas aged 25 to 64 years	1971-1982	ICD-9 410-414	WHO criteria	0.91, 0.93	NA
Dobson *et al*. [45]	Australia	Permanent residents of the Hunter region of New South Wales aged 20 to 69 years	1979	ICD-8/9 410-414	WHO criteria	0.91	NA
Mahonen *et al*. [51]	Finland	Residents of North Karelia, Kuopio, and the Turku/ Loimaa region aged 35 to 64 years	1983-1992	ICD-9 410-414	WHO criteria	0.95	0.69
McIlwaine *et al*. [53]	Ireland	Residents of Belfast	1981-1982	ICD-9 410-414	WHO criteria	0.89	0.67
Saito *et al*. [54]	Japan	Residents of Oita City aged 25 to 74 years	1997-1998	ICD-10 I20-I25	WHO criteria	0.87	0.65
Agarwal *et al*. [43]	U.S.	New York City residents who died in one of the City's 70 hospitals	2003	ICD-10 I20-I25, I51.6	Gillum[e] and CCSP[f] criteria	0.87	NA
Lloyd-Jones *et al*. [50]	U.S.	Residents of Framingham, MA 45 years of age and older	1948-1988	ICD-9 410-414	Physician panel	0.84	0.84

Coady *et al*. [44]	U.S.	Residents 35 to 74 years of age living within the county boundaries of Forsyth County, NC, or the metropolitan boundaries for Jackson, MS, or suburban Minneapolis, MN	1987-1995	ICD-9 410-414, 429.2	Gillum and CCSP criteria	0.81	0.72
Goraya *et al*. [47]	U.S.	10% random sample of Olmstead County, MN residents who died out of hospital	1981-1994	ICD-9 410-414	Gillum criteria	0.91	0.86
Iribarren *et al*. [49]	U.S.	Residents of Mankato, MN; North Mankato, MN; Winona, MN; Bloomington, MN; Roseville, MN; Maplewood, MN; North St. Paul, MN; and Moorhead, MN who were 30 to 74 years of age and died out of hospital	1985-1990	ICD-9 410-414	Physician review using unknown criteria	0.87	0.66
Folsom *et al*. [46]	U.S.	33% random sample of Minneapolis, MN and St. Paul, MN residents who were 30 to 74 years of age and died out of hospital	1979	ICD-9 410-414, 427	Gillum criteria	0.90	0.83

[a]Using the World Bank income groups of low, lower-middle, upper-middle, and high, which are based on the World Bank list of analytical income classification of economies [41].

[b]Established by the World Health Organization for the MONICA Project [57].

[c]Mortality definitions established for the Second Australian National Blood Pressure Study [58, 59].

[d]Mortality definitions established for the Long-term Intervention with Pravastatin in Ischaemic Disease Study [60, 61].

[e]Published by Gillum *et al*. [62]

[f]Established by the Community Cardiovascular Surveillance Program [63].

Table 6. Description of probability distributions used for classification parameters**.**

Scenario	p_{11}	p_{22}	p_{33}	p_{44}
Non-differential misclassification[a]				
1	Beta (23.9, 3)[b]	Beta (23.9, 3)	Maximum Extreme (0.47, 0.17)[c]	Maximum Extreme (0.47, 0.17)
Differential misclassification				
2	Beta (23.9, 3)	Beta (23.9, 3)	Maximum Extreme (0.47, 0.17)	Minimum Extreme (0.71, 0.25)[d]
4	Beta (23.9, 3)	Beta (23.9, 4.3)[e]	Maximum Extreme (0.47, 0.17)	Maximum Extreme (0.47, 0.17)
6	Beta (23.9, 3)	Beta (23.9, 4.3)	Maximum Extreme (0.47, 0.17)	Minimum Extreme (0.71, 0.25)
8	Beta (23.9, 4.3)	Beta (23.9, 3)	Maximum Extreme (0.47, 0.17)	Minimum Extreme (0.71, 0.25)
Differential misclassification				
3	Beta (23.9, 3)	Beta (23.9, 3)	Minimum Extreme (0.71, 0.25)	Maximum Extreme (0.47, 0.17)
5	Beta (23.9, 4.3)	Beta (23.9, 3)	Maximum Extreme (0.47, 0.17)	Maximum Extreme (0.47, 0.17)
7	Beta (23.9, 4.3)	Beta (23.9, 3)	Minimum Extreme (0.71, 0.25)	Maximum Extreme (0.47, 0.17)
9	Beta (23.9, 3)	Beta (23.9, 4.3)	Minimum Extreme (0.71, 0.25)	Maximum Extreme (0.47, 0.17)

[a]In expectation- the individual trials may only be approximately non-differential, but the average difference in classification proportions across all simulation trials is 0.

[b]Beta distribution (alpha, beta) for minimum, maximum and peak proportions of 0, 1 and 0.92, respectively.

[c]Maximum extreme distribution (likeliest, scale) for minimum, maximum and peak proportions of 0, 1, and 0.47, respectively.

[d]Minimum extreme distribution (likeliest, scale) for minimum, maximum and peak proportions of 0, 1, and 0.71, respectively.

[e]Beta distribution (alpha, beta) for minimum, maximum and peak proportions of 0, 1 and 0.89.

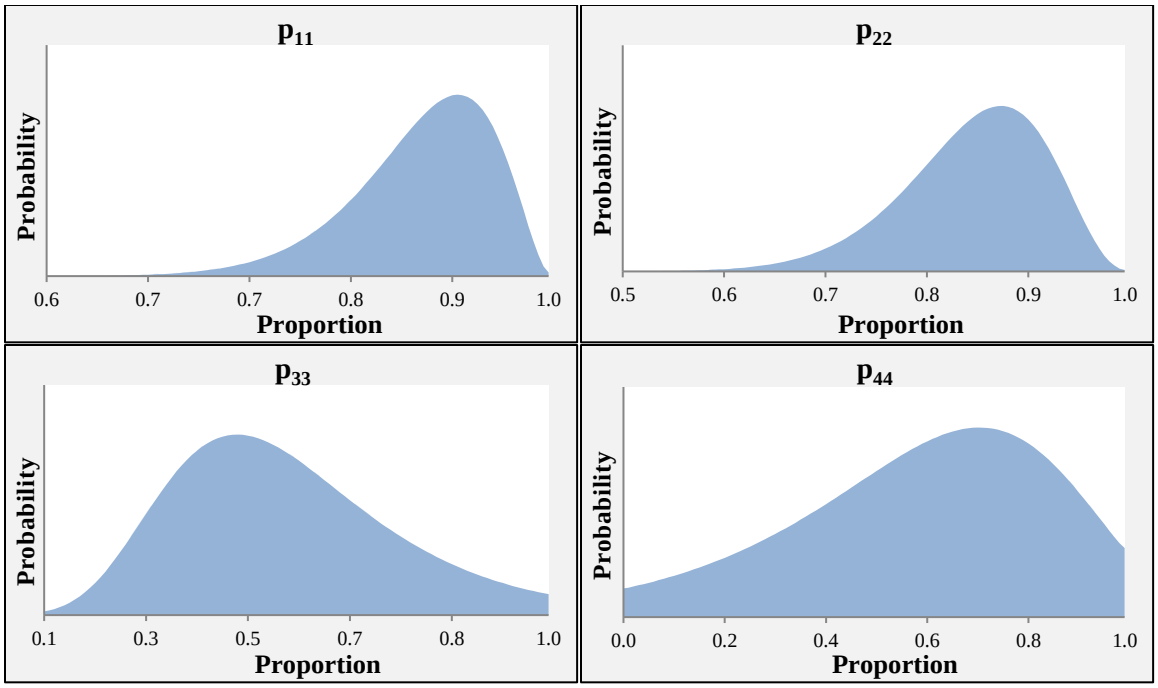


Figure 1. Example of paramter distribution input for Scenario 6.

specified parameter distributions were sampled using Oracle Crystal Ball 11.1 (Redwood Shores, CA), and all calculations were conducted in Microsoft Office Excel 2010 (Redmond, WA).

Any Monte Carlo simulation trial with a negative count was excluded as it is not possible to have negative counts of subjects in a study. Parameter distributions were sampled enough times to obtain a minimum of 10,000 valid simulation trials. Each analysis was conducted separately to generate a frequency distribution for $OR_{DM\text{-}COD}$ and $\mathcal{E}_{DM\text{-}COD}$. The lower 2.5 and upper 97.5 percentiles of each frequency distribution were used to estimate 95% certainty intervals.

Results

Results for each scenario are illustrated in Figure 2. The geometric mean (GM) of the error term for disease misclassification ($\varepsilon_{DM\text{-}COD}$) ranged from 0.7 to 1.6, and the GM adjusted odds ratio ($OR_{DM\text{-}COD}$) had a range of 1.9 to 4.9. Estimated certainty intervals (CI) for the geometric mean $OR_{DM\text{-}COD}$ did not exclude the null in any of the nine scenarios.

Under the assumption of non-differential misclassification, the GM $OR_{DM\text{-}COD}$ was 3.0, just slightly smaller than the observed odds ratio of 3.05. In the three scenarios where we specified less misclassification of deceased non-cases in the highest exposure group (*i.e.*, scenarios 2, 6, and 8), the adjusted OR shifted away from the null (*i.e.*, $OR_{observed}$ was biased toward the null). More notably, the direction of the effect was preserved regardless of whether cases in the highest exposure category had a higher (scenario 6) or lower (scenario 8) amount of misclassification than cases in the never-exposed group. Greater misclassification of the highest exposed cases (scenario 4) also produced a geometric mean $OR_{DM\text{-}COD}$ greater than the $OR_{observed}$, although to a lesser extent. As expected a similar pattern, but in the opposite direction, was observed when a greater degree of misclassification was specified for deceased non-cases with the highest exposures (scenarios 3, 7, 9) and a smaller amount of misclassification was specified for the highest exposed cases (scenario 5). In these scenarios, misclassification of deceased non-cases in the never-exposed category (*i.e.*, parameter p_{33}) had the greatest influence on the directional shift of the distribution.

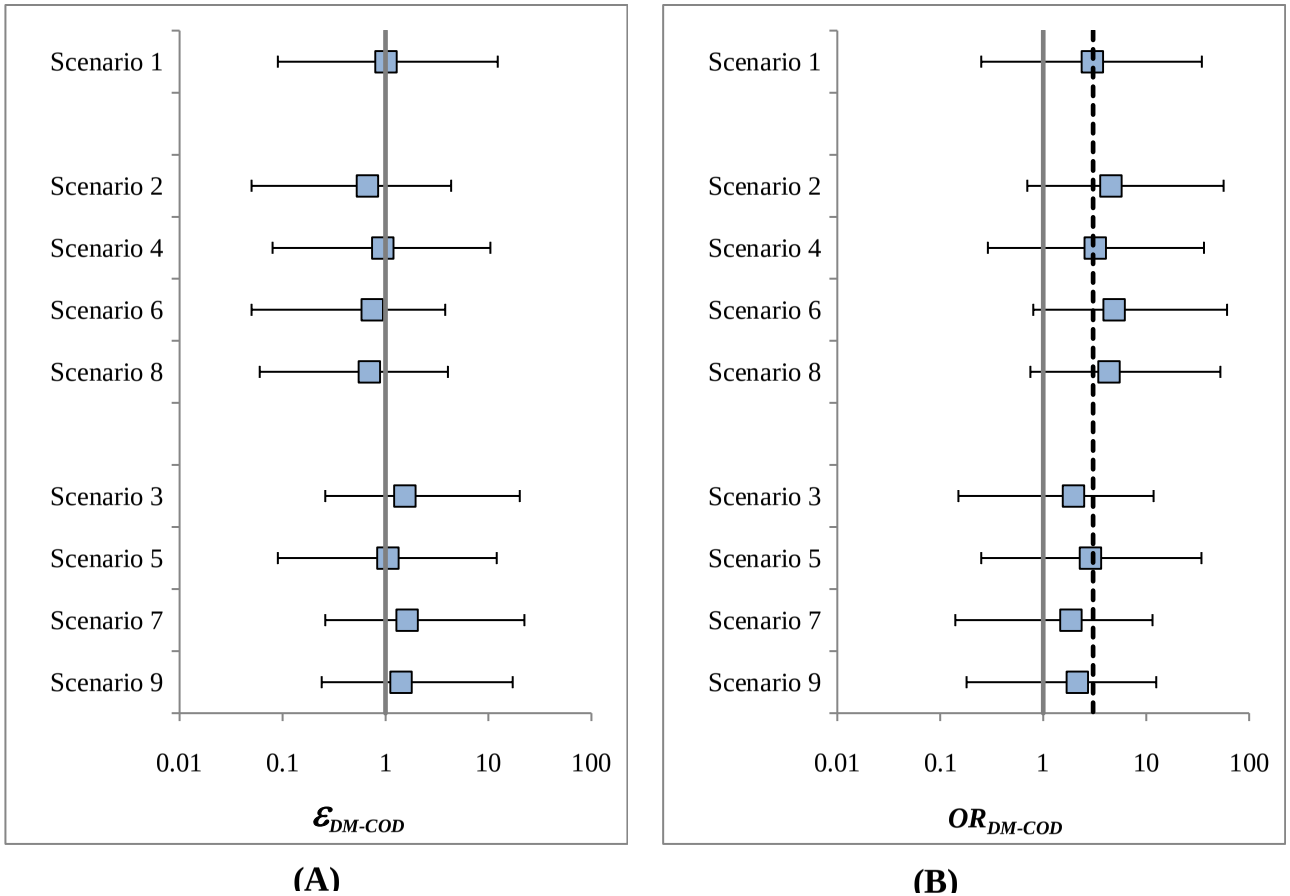


Figure 2. Geometric mean errors ($\mathcal{E}_{DM\text{-}COD}$) (A), adjusted odds ratios ($OR_{DM\text{-}COD}$) (B), and 95% certainty intervals by scenario.

Discussion

Death certificates are undoubtedly an important resource for monitoring population health and conducting research. Yet, information they capture on cause of death can be ambiguous and erroneous. In their study, McBride and colleagues [22] identified employee deaths and the causes of death using the New Zealand Health Information Service (NZHIS), which determines the underlying cause of death after obtaining medical certificates and coroners' findings from the Registrar-General of Births, Deaths and Marriages [64]. When insufficient information is available to assign an underlying cause of death, NZHIS will seek information from other sources such as hospitals, the courts, the New Zealand Cancer Registry, and the Land Transport Safety Authority, among others. However, while this process is likely adequate for most deaths, it is not without limitations and there have been past concerns about the accuracy of New Zealand death certificates [65, 66].

To adjust for potential disease misclassification from the use of death certificate data, we developed a probabilistic bias-analysis method specific for historical cohort mortality studies. Whereas bias analysis for assessing outcome misclassification is much more straightforward when cases are defined as having the disease and non-cases as not having the disease, it is more complex when the outcome being evaluated is a specific cause of death since non-cases can include living individuals as well as those that died from other causes. In these situations,

bias-analysis corrections for disease misclassification must only be applied to cases and non-cases who have died (*i.e.*, a living individual could not have been misclassified with regard to cause of death).

The impact of study error on a measure of effect is dependent on a number of factors – including the type of error, levels of the exposure variable, and degree of misclassification – and can, therefore, be difficult to anticipate. In the present study, probabilistic application of plausible sensitivity and specificity values to the counts for each outcome category demonstrated that the unadjusted TCDD-IHD association could be biased in either direction when misclassification was differential in expectation. In particular, less misclassification of deceased non-cases in the highest exposure group shifted the frequency distribution of adjusted estimates to the right of $OR_{observed}$ while less misclassification of deceased non-cases in the never-exposed category shifted the distribution to the left, with misclassification of IHD cases in either exposure category having only minor influence. As important, when misclassification was non-differential in expectation, the GM $OR_{DM\text{-}COD}$ of 3.00 was actually smaller than the $OR_{observed}$ of 3.05, reinforcing more contemporary understanding that non-differential misclassification does not always bias an estimate toward the null [67].

Only one IHD validation study of New Zealand death certificate data was identified through a literature search. To supplement the sensitivity and specificity reported by Jackson *et al*. [42], we included estimates from other validation

studies of IHD in high-income countries, which we assumed would have similar infrastructure for identifying deaths and recording their causes. For transparency, we provided the relevant characteristics – including country, study population and years, ICD codes, and the validation standard – of each study used to inform our probability distributions. While these studies reported sufficient data to construct the distributions for each disease classification parameter, the small number of cases in the two exposure categories of interest (*i.e.*, never-exposed and ≥2085.5 ppt-mo) greatly limited the number of valid sensitivity and specificity combinations. In addition, the accuracy with which ischemic heart disease is assigned as the cause of death has been reported to vary by age, gender, and race [31, 43, 44, 50, 54, 63, 68], with age having a particularly notable influence on disease classification in occupational studies [32]. Exposure groups indeed have the potential to be differentially misclassified, not because of exposure itself, but as a result of dissimilarities in the demographic characteristics of each group (*e.g.*, those with no occupational exposure may be predominantly younger, white women whereas those with the highest exposures may be mostly older, minority men). However, details on age, gender, and race were not available for the different exposure groups in the study by McBride *et al*. [22]; such information, though not necessary to conduct the analysis, would have been helpful in refining the sensitivity and specificity distributions in the eight scenarios with differential misclassification.

The methods and analyses described here focus specifically on disease misclassification resulting from the use of cause-of-death data. To simplify application of this method, we purposely assumed there was no exposure misclassification, disease misclassification from loss to follow-up (*i.e.*, study subjects classified as living who truly died of IHD), or other disease misclassification (*i.e.*, study subjects classified as deceased from IHD who were truly alive). Although it is doubtful the latter type of misclassification would result in profound study error [69-73], Scott and Maldonado [39] previously demonstrated – using a different probabilistic bias-analysis method – that losses to follow-up in historical cohort mortality studies can cause appreciable disease misclassification. In the future, the combined effect of these sources of misclassification will need to be examined along with determining the model that best represents their mathematical relationship. Logically, adjustment for the type of disease misclassification described in the current manuscript must be done prior to adjustment for disease misclassification resulting from loss to follow-up. The order of adjustment for other study errors (*i.e.*, exposure misclassification), however, will need to be determined.

Probabilistic sensitivity methods, like the example provided here, can be useful for evaluating study error and improving causal inference. While the technique we described relies on the availability of external validation data and specifying well-defined parameters, it can be applied to mortality studies with a

multi-level exposure variable using any number of Monte Carlo software programs and does not require access to the original data. With the continued development and refinement of methods to account for sources of study error, we hope the practice of conducting bias analyses will become more commonplace.

CHAPTER IV

APPLICATION OF PROBABILISTIC BIAS ANALYSIS TO ADJUST FOR EXPOSURE MISCLASSIFICATION IN HISTORICAL COHORT MORTALITY STUDIES WITH MULTI-LEVEL EXPOSURE VARIABLES: AN EXAMPLE USING A COHORT OF TRICHLOROPHENOL WORKERS

Laura L. F. Scott and George Maldonado

Division of Environmental Health Sciences, University of Minnesota School of Public Health, Minneapolis, MN, 55455

Summary

The objective of developing this method was to demonstrate the application of probabilistic bias analysis to quantify and adjust for exposure misclassification in a historical cohort mortality study of trichlorophenol workers with a multi-level exposure variable. Published exposure information available for this cohort of workers was used to specify the initial classification parameter distributions, which were then varied to assess the potential impacts of different levels of misclassification and both non-differential and differential exposure misclassification. Each parameter distribution was sampled 50,000 times using Monte Carlo simulation techniques, calculating odds ratios adjusted for exposure

misclassification (OR_{EM}) and the associated exposure misclassification error terms (ε_{EM}). Given the specified assumptions, the geometric mean (GM) OR_{EM} had a range of 2.89 to 5.05, and the GM ε_{EM} ranged from 0.60 to 1.05. In all non-differential scenarios and scenarios in which non-cases had greater proportions of misclassification, the $OR_{observed}$ of 3.05 was closer to the null (*i.e.*, 1) than the GM OR_{EM}. For the differential simulations where cases had higher proportions of misclassification, the direction of the error was dependent on the level of misclassification error, with smaller proportions of misclassification resulting in the $OR_{observed}$ being farther away from the null than the GM OR_{EM}. These findings demonstrate that probabilistic bias analysis of historical cohort mortality studies can be an effective tool for understanding trends in study error stemming from exposure misclassification and confirm the importance of quantifying potential sources of systematic error prior to forming conclusions regarding causality.

Key words: probabilistic uncertainty analysis, Monte Carlo, exposure misclassification, historical cohort mortality

Introduction

It is well known that systematic study error in observational research can severely impact study findings, resulting in potentially incorrect inferences [3-14]. Because biased estimates of the exposure-disease relationship can have serious implications – for not only determining the direction of future research but also for both risk assessment practices and the development of public health policy – researchers have developed a variety of methods to address this issue. One such method is probabilistic bias analysis (also referred to as probabilistic uncertainty analysis) [3-6, 8-12, 15]. Using this approach, an investigator can evaluate the effect of systematic error by specifying bias parameters and then adjusting an effect estimate for biases. This method is different from simple sensitivity analysis, however, because it uses probability distributions for bias parameters and Monte Carlo techniques to incorporate uncertainty about the parameter value(s) into the adjusted estimate of effect.

Historical cohort mortality studies are one of the most common tools used by epidemiologists to evaluate associations between occupational exposure and disease mortality [2, 74-77], yet the impact of exposure misclassification is infrequently quantified in these studies. To our knowledge, the magnitude of exposure misclassification has only once been quantitatively estimated in a cohort mortality study using probabilistic bias analysis [78]. Although significant advancements have been made in exposure assessment methods over the last several decades, there is still the potential for considerable exposure

misclassification in historical cohort mortality studies. First, job histories, and thereby exposure proxies, are almost always partially ascertained from administrative and human resources records that may be incomplete, inaccurate and lack specificity. Even recent record-keeping systems that provide more detailed information may not adequately describe the frequency of the tasks employees perform for each job title and class, the materials and processes with which they work, the use of personal protective equipment and any non-routine operations or changes in operations. Occassionally, the judgment of current or former employees (or both) is used to help identify tasks, exposures to specific agents, and temporal variability in exposure, but this can also introduce systematic error in exposure estimates. Exposure misclassification error can also occur with quantitative measurements such as biological, personal, and area monitoring that may be measured years after exposure and then adjusted using mathematical models to reflect historical exposures. Biomonitoring data, in particular, can lead to misclassification if concentrations were measured infrequently and in only a non-randomly selected subpopulation of the total cohort, but are assumed to represent all workers with similar job histories. It is also often obtained without any insight as to an individual's elimination kinetics.

Bias analysis methods for studies with polytomous exposure variables have not been fully developed. Thus, we designed a probabilistic bias analysis model to adjust results from historical cohort mortality studies with multiple levels of

exposure. With the use of published data on a cohort of trichlorophenol workers from New Zealand, we describe the process of adjusting for exposure misclassification when exposure is classified into three or more categories and also illustrate how differential and non-differential misclassification can change an estimate of effect.

Materials and Methods

The data used here are from a historical cohort mortality study [22]. The authors examined multiple causes of mortality and occupational exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) among 1599 workers at a trichlorophenol plant in New Zealand. Employee deaths and underlying causes were identified using the New Zealand Health Information Service Mortality Collection. In this paper, we focused on the association between TCDD exposure and ischemic heart disease (IHD) mortality.

Based on work history records, individuals were categorized as "ever-exposed" (n = 1134) or "never-exposed" (n = 465) to TCDD. Workers in the "ever-exposed" category were further classified by first developing cumulative TCDD exposure estimates (in units of parts per trillion times months [ppt-mo]) using work history files, current serum TCDD levels of 346 cohort members, and linear regression modeling of pharmacokinetic equations to estimate dose rates for each job exposure group (JEG). Individuals were then categorized into one of four groups based on their cumulative exposure estimate: "0-68.3 ppt-mo" (n = 1061);

"68.4-475.0 ppt-mo" (*n* = 668); "475.1-2085.7 ppt-mo" (*n* = 363); and "≥2085.8 ppt-mo" (*n* = 162).

The study described by McBride *et al*. [22] reported no statistical increase in the standardized mortality ratio (SMR) for death from IHD in either the "ever-exposed" (SMR: 1.1, 95% confidence interval (CI): 0.9-1.5) or the "never-exposed" (SMR: 0.9, 95% CI: 0.5-1.5) categories when compared to the New Zealand general population. Additionally, no statistically significant trend was observed ($p = 0.50$) when IHD mortality in the three highest cumulative exposure categories was compared with that of the lowest cumulative TCDD exposure group. Although there was no significant increase in IHD death, there was also no clear evidence of a prominent healthy worker effect for cardiovascular disease mortality, as is often observed in cohorts of industrial workers [79-82].

Multiple potential sources of systematic error – including uncontrolled confounding, selection bias, and disease and exposure misclassification – could have influenced the reported findings for this cohort of workers. Although the study by McBride and coworkers [22] is one of the few studies to use serum dioxin levels to determine cumulative occupational exposure to TCDD and disease risks, there were several steps in estimating cumulative exposures where misclassification could have occurred. In the present bias analysis, we focused on the impact of exposure misclassification, which could have arisen in several ways. First, TCDD serum concentrations were measured in only a small proportion

(21.6%) of the total cohort, and employee serum was sampled between 17 and 19 years after occupational exposure ceased. Given the age- and concentration-dependent half-life of TCDD ranges from less than three years to over 10 years [83], the use of biomonitoring data could have resulted in underestimation of exposure for some workers and overestimation for others. In addition, exposure misclassification could also have resulted from a) using paper work history records and automated payroll records to determine job assignments for each worker, b) using interviews of current and former employees to qualitatively categorize departments and jobs with similar potential for exposure, and c) the multiple linear regression model used to estimate past exposures for each job exposure group. In addition, McBride *et al*. [22] compared IHD mortality in the higher cumulative exposure groups with that of the lowest cumulative exposure group, rather than with "never-exposed" workers. Consequently, the reported findings may not provide an estimate of the effect of TCDD exposure on IHD mortality that accurately corresponds to the true relationship [21, 77, 84].

In equation 19 of his 2008 journal article, Maldonado [5] described the mathematical relationship between a causal relative risk (*e.g.*, causal odds ratio or incidence proportion ratio), an observed relative risk from a typical epidemiologic analysis, and error terms (for study imperfections and random error) as

$$E(RR_{causal}) = \frac{RR_{adjusted}}{\prod_{i=1}^{n} \varepsilon_i} \quad (1)$$

where $E(RR_{causal})$ is the expected value of a causal relative risk, $RR_{adjusted}$ is an observed relative risk from a typical epidemiologic analysis (perhaps adjusted for covariates), ε_i are error terms that describe the impact of study imperfections and random error on $RR_{adjusted}$, and $\prod_{i=1}^{n}$ denotes that the n error terms multiply.

We modified Equation (1) to fit the topic of this manuscript, where the odds ratio is the relative risk of interest, and the only study bias under examination is that due to possible exposure misclassification. In this situation, Equation (1) simplifies to Equation (2), where OR_{EM} is an odds ratio adjusted for exposure misclassification, $OR_{observed}$ is an observed crude odds ratio, and ε_{EM} is the error term for exposure misclassification.

$$OR_{EM} = \frac{OR_{observed}}{\varepsilon_{EM}} \quad (2)$$

Count data extracted from McBride *et al*. [22] were used to calculate a crude odds ratio ($OR_{observed}$) and 95% confidence interval for the association between TCDD exposure and ischemic heart disease mortality (Table 1). In the analysis presented here, in contrast to the analysis by McBride *et al*. [22], we chose to compare the IHD mortality of those TCP production workers that had the highest cumulative TCDD exposure with those at the facility who were considered never-exposed because 1) using an internal reference group may lessen the potential for confounding and healthy worker bias inherent in standardized

mortality ratios (SMRs) and 2) this contrast is the most extreme and so avoids dilution of the effect from combining all exposed workers, including those with much lower exposures, into one group [77, 84-87].

Table 1. Cell counts, crude odds ratio and 95% confidence interval for the association between TCDD exposure and ischemic heart disease using data reported by McBride *et al*. [22].

	TCDD Exposure		**OBSERVED ODDS RATIO**	**95% CONFIDENCE LIMITS**
Outcome	Never-exposed	≥2085.8 ppt-mo		
IHD Cases	14	14	3.05	1.42, 6.54
Non-cases	451	148		

TCDD, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin; ppt-mo, parts per trillion-month; IHD, ischemic heart disease

To estimate the proportions of workers misclassified, we used a multi-step process. First, we defined a matrix of classification proportions for the exposure categories "never-exposed" and "ever-exposed" (E_1 matrix) (Table 2). We then stratified the group classified in Table 2 as “ever-exposed” into two categories, specifying a second matrix of classification proportions for the exposures “0-2085.7 ppt-mo” and “≥2085.8 ppt-mo” (E_2 matrix) (Table 3).

Table 2. E_1 matrix of classification proportions for "never-exposed" and "ever-exposed" categories.

	True Exposure Status					
	IHD Cases			*Non-cases*		
Classified Exposure Status	Never-exposed	≥2085.8 ppt-mo	0-2085.7 ppt-mo	Never-exposed	≥2085.8 ppt-mo	0-2085.7 ppt-mo
Never-exposed	e_{11}	e_{12}	e_{13}	e_{14}	e_{15}	e_{16}
Ever-exposed	e_{21}	e_{22}	e_{23}	e_{24}	e_{25}	e_{26}
Total	1	1	1	1	1	1

IHD, ischemic heart disease; ppt-mo, parts per trillion-month

Table 3. E_2 matrix of classification proportions for "0-2085.7 ppt-mo" and "≥2085.8 ppt-mo" categories.

	IHD Cases Classified as Ever-exposed			*Non-cases Classified as Ever-exposed*		
Classified Exposure Status	True Never-exposed (e_{21})	True ≥2085.8 ppt-mo (e_{22})	True 0-2085.7 ppt-mo (e_{23})	True Never-exposed (e_{24})	True 2085.8 ppt-mo (e_{25})	True 0-2085.7 ppt-mo (e_{26})
≥2085.8 ppt-mo	e_{31}	e_{32}	e_{33}	e_{34}	e_{35}	e_{36}
0-2085.7 ppt-mo	e_{41}	e_{42}	e_{43}	e_{44}	e_{45}	e_{46}
Total	1	1	1	1	1	1

IHD, ischemic heart disease; ppt-mo, parts per trillion-month

Probability distributions for the 12 exposure-classification parameters (e_{21}, e_{12}, e_{13}, e_{24}, e_{15}, e_{16}, e_{31}, e_{42}, e_{33}, e_{34}, e_{45}, e_{36}) were constructed using published exposure data available for the same cohort of New Zealand TCP workers [88, 89]. Although the studies of this cohort by Aylward *et al*. [88] and Collins *et al*. [89] did not provide any exposure validation estimates, we were able to use the data described in these studies to specify probability distributions for the exposure-classification parameters. For example, in Collins *et al*. [89] the authors note that one of the 105 workers in the "never-exposed" group had a TCDD serum concentration "clearly outside the range of New Zealand background" and, according to the employee questionnaire, worked on the incinerator, which was categorized as having high continuous exposure potential. The resulting proportion (~1%) was used to define the exposure-classification probability distributions for e_{12} (the proportion of IHD cases with exposure "≥2085.8 ppt-mo" misclassified as "never-exposed"), e_{13} (the proportion of IHD cases with exposure "0-2085.7 ppt-mo" misclassified as "never-exposed"), e_{15} (the proportion of non-cases with

exposure "≥2085.8 ppt-mo" misclassified as "never-exposed"), and e_{16} (the proportion of non-cases with exposure "0-2085.7 ppt-mo" misclassified as "never-exposed"). More specifically, for these four parameters we set the mode/likeliest value of the classification probability distributions to 0.01, the minimum to 0.005, and the maximum to 0.015. We also used data from Collins and colleagues [89] to construct the probability distributions for the classification parameters e_{21} (the proportion of IHD cases with no occupational exposure misclassified as "ever-exposed") and e_{24} (the proportion of non-cases with no occupational exposure misclassified as "ever-exposed"). The probability distributions for these parameters were determined using the minimum and maximum proportions (17%-32%) of JEG-specific serum donors that had mean TCDD serum concentrations less than 3.9 ppt lipid (*i.e.*, the lipid-adjusted serum background level for New Zealanders of comparable age) [22, 90]. Here, we set the mode/likeliest values of the classification probability to 0.25, the minimum to 0.15 and the maximum to 0.35.

The exposure classification probability distributions for e_{42}, e_{33}, e_{45}, and e_{36} were constructed using data from the Aylward *et al*. study of this cohort [88]. A total of 11 individuals out of 343 (*i.e.*, 346 individuals that donated serum for analysis minus three workers excluded from the model) had much higher measured TCDD concentrations compared to their modeled exposure. Of these 11, seven were identified by the study authors as having been misclassified, giving a

proportion of 7/343 = 0.02 for those workers who may have truly had considerable exposure (*i.e.*, "≥2085.8 ppt-mo") but would have been classified into one of the middle exposure categories (e_{42} and e_{45}). For workers who were categorized in the highest exposure group but had much lower exposures, we estimated the proportion (*i.e.*, e_{33} and e_{36}) to be 4/343, or 1.17%, using data from Figure 1 of the Aylward publication [88] and assuming that a smaller proportion of workers would have their exposures overestimated to such a magnitude that they would be classified in the highest category. Lastly, we estimated that the proportion of workers classified in the highest exposure category who were truly never occupationally exposed (e_{31} and e_{34}) at this New Zealand facility to be ~1% (1/105; 4/343). The mode/likeliest, minimum and maximum values were all determined in a similar manner to the method we used for the parameters e_{12}, e_{13}, e_{15}, and e_{16}.

Several different distributions (Tables 4A and 4B) were specified to account for varying levels of misclassification and assess differences in the impact of non-differential (in expectation) and differential misclassification. To start, we defined two basic distributions (*i.e.*, uniform and triangular) and a modified beta distribution (called a betaPERT or simply PERT) using the proportions described in the previous paragraphs (scenarios 1, 3 and 5). We chose the betaPert distribution – which is renormalized over a finite range other than [0,1] and reparameterized by the minimum, maximum, and mode or likeliest value – as a smoother alternative to the triangular distribution [24]. Because we collapsed the

Table 4A. Description of probability distributions used for classification parameters for cases.

Scenario	e_{21}	e_{12}	e_{13}	e_{31}	e_{42}	e_{33}
Non-differential[a]						
1	Uniform[b] (0.15, 0.35)	Uniform (0.005, 0.015)	Uniform (0.005, 0.015)	Uniform (0.005, 0.015)	Uniform (0.01, 0.03)	Uniform (0.005, 0.015)
2	Uniform (0.05, 0.15)	Uniform (0.025, 0.075)	Uniform (0.05, 0.15)	Uniform (0.01, 0.03)	Uniform (0.05, 0.15)	Uniform (0.025, 0.075)
3	Triangular[c] (0.15, 0.25, 0.35)	Triangular (0.005, 0.01, 0.015)	Triangular (0.005, 0.01, 0.015)	Triangular (0.005, 0.01, 0.015)	Triangular (0.01, 0.02, 0.03)	Triangular (0.005, 0.01, 0.015)
4	Triangular (0.05, 0.1, 0.15)	Triangular (0.025, 0.05, 0.075)	Triangular (0.05, 0.10, 0.15)	Triangular (0.01, 0.02, 0.03)	Triangular (0.05, 0.1, 0.15)	Triangular (0.025, 0.05, 0.075)
5	BetaPert[d] (0.15, 0.25, 0.35)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.01, 0.02, 0.03)	BetaPert (0.005, 0.01, 0.015)
6	BetaPert (0.05, 0.1, 0.15)	BetaPert (0.025, 0.05, 0.075)	BetaPert (0.05, 0.10, 0.15)	BetaPert (0.01, 0.02, 0.03)	BetaPert (0.05, 0.1, 0.15)	BetaPert (0.025, 0.05, 0.075)
Differential[e]						
7	Uniform (0.15, 0.35)	Uniform (0.005, 0.015)	Uniform (0.005, 0.015)	Uniform (0.005, 0.015)	Uniform (0.01, 0.03)	Uniform (0.005, 0.015)
8	Uniform (0.05, 0.15)	Uniform (0.025, 0.075)	Uniform (0.05, 0.15)	Uniform (0.01, 0.03)	Uniform (0.05, 0.15)	Uniform (0.025, 0.075)
9	Triangular[c] (0.15, 0.25, 0.35)	Triangular (0.005, 0.01, 0.015)	Triangular (0.005, 0.01, 0.015)	Triangular (0.005, 0.01, 0.015)	Triangular (0.01, 0.02, 0.03)	Triangular (0.005, 0.01, 0.015)
10	Triangular (0.05, 0.1, 0.15)	Triangular (0.025, 0.05, 0.075)	Triangular (0.05, 0.10, 0.15)	Triangular (0.01, 0.02, 0.03)	Triangular (0.05, 0.1, 0.15)	Triangular (0.025, 0.05, 0.075)
11	BetaPert[d] (0.15, 0.25, 0.35)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.01, 0.02, 0.03)	BetaPert (0.005, 0.01, 0.015)
12	BetaPert (0.05, 0.1, 0.15)	BetaPert (0.025, 0.05, 0.075)	BetaPert (0.05, 0.10, 0.15)	BetaPert (0.01, 0.02, 0.03)	BetaPert (0.05, 0.1, 0.15)	BetaPert (0.025, 0.05, 0.075)
Differential[f]						
13	Uniform (0.10, 0.25)	Uniform (0.0025, 0.0075)	Uniform (0.0025, 0.0075)	Uniform (0.0025, 0.0075)	Uniform (0.005, 0.015)	Uniform (0.0025, 0.0075)

14	Uniform (0.025, 0.075)	Uniform (0.0125, 0.0375)	Uniform (0.025, 0.075)	Uniform (0.005, 0.015)	Uniform (0.025, 0.075)	Uniform (0.0125, 0.0375)
15	Triangular (0.10, 0.175, 0.25)	Triangular (0.0025, 0.005, 0.0075)	Triangular (0.0025, 0.005, 0.0075)	Triangular (0.0025, 0.005, 0.0075)	Triangular (0.005, 0.01, 0.015)	Triangular (0.0025, 0.005, 0.0075)
16	Triangular (0.025, 0.05, 0.075)	Triangular (0.0125, 0.025, 0.0375)	Triangular (0.025, 0.05, 0.075)	Triangular (0.005, 0.01, 0.015)	Triangular (0.025, 0.05, 0.075)	Triangular (0.0125, 0.025, 0.0375)
17	BetaPert (0.10, 0.175, 0.25)	BetaPert (0.0025, 0.005, 0.0075)	BetaPert (0.0025, 0.005, 0.0075)	BetaPert (0.0025, 0.005, 0.0075)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.0025, 0.005, 0.0075)
18	BetaPert (0.025, 0.05, 0.075)	BetaPert (0.0125, 0.025, 0.0375)	BetaPert (0.025, 0.05, 0.075)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.025, 0.05, 0.075)	BetaPert (0.0125, 0.025, 0.0375)

[a]In expectation- the individual trials may only be approximately non-differential, but the average difference across all simulation trials is 0.

[b]Continuous uniform distribution (minimum, maximum).

[c]Triangular distribution (minimum, mode, maximum).

[d]BetaPERT distribution (minimum, likeliest, maximum).

[e]Cases have a greater proportion of misclassified.

[f]Non-cases have a greater proportion of misclassified.

Table 4B. Description of probability distributions used for classification parameters for non-cases.

Scenario	**e_{24}**	**e_{15}**	**e_{16}**	**e_{34}**	**e_{45}**	**e_{36}**
Non-differential[a]						
1	Uniform[b] (0.15, 0.35)	Uniform (0.005, 0.015)	Uniform (0.005, 0.015)	Uniform (0.005, 0.015)	Uniform (0.01, 0.03)	Uniform (0.005, 0.015)
2	Uniform (0.05, 0.15)	Uniform (0.025, 0.075)	Uniform (0.05, 0.15)	Uniform (0.01, 0.03)	Uniform (0.05, 0.15)	Uniform (0.025, 0.075)
3	Triangular[c] (0.15, 0.25, 0.35)	Triangular (0.005, 0.01, 0.015)	Triangular (0.005, 0.01, 0.015)	Triangular (0.005, 0.01, 0.015)	Triangular (0.01, 0.02, 0.03)	Triangular (0.005, 0.01, 0.015)
4	Triangular (0.05, 0.1, 0.15)	Triangular (0.025, 0.05, 0.075)	Triangular (0.05, 0.10, 0.15)	Triangular (0.01, 0.02, 0.03)	Triangular (0.05, 0.1, 0.15)	Triangular (0.025, 0.05, 0.075)
5	BetaPert[d] (0.15, 0.25, 0.35)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.01, 0.02, 0.03)	BetaPert (0.005, 0.01, 0.015)
6	BetaPert (0.05, 0.1, 0.15)	BetaPert (0.025, 0.05, 0.075)	BetaPert (0.05, 0.10, 0.15)	BetaPert (0.01, 0.02, 0.03)	BetaPert (0.05, 0.1, 0.15)	BetaPert (0.025, 0.05, 0.075)
Differential[e]						
7	Uniform (0.10, 0.25)	Uniform (0.0025, 0.0075)	Uniform (0.0025, 0.0075)	Uniform (0.0025, 0.0075)	Uniform (0.005, 0.015)	Uniform (0.0025, 0.0075)
8	Uniform (0.025, 0.075)	Uniform (0.0125, 0.0375)	Uniform (0.025, 0.075)	Uniform (0.005, 0.015)	Uniform (0.025, 0.075)	Uniform (0.0125, 0.0375)
9	Triangular (0.10, 0.175, 0.25)	Triangular (0.0025, 0.005, 0.0075)	Triangular (0.0025, 0.005, 0.0075)	Triangular (0.0025, 0.005, 0.0075)	Triangular (0.005, 0.01, 0.015)	Triangular (0.0025, 0.005, 0.0075)
10	Triangular (0.025, 0.05, 0.075)	Triangular (0.0125, 0.025, 0.0375)	Triangular (0.025, 0.05, 0.075)	Triangular (0.005, 0.01, 0.015)	Triangular (0.025, 0.05, 0.075)	Triangular (0.0125, 0.025, 0.0375)
11	BetaPert (0.10, 0.175, 0.25)	BetaPert (0.0025, 0.005, 0.0075)	BetaPert (0.0025, 0.005, 0.0075)	BetaPert (0.0025, 0.005, 0.0075)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.0025, 0.005, 0.0075)
12	BetaPert (0.025, 0.05, 0.075)	BetaPert (0.0125, 0.025, 0.0375)	BetaPert (0.025, 0.05, 0.075)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.025, 0.05, 0.075)	BetaPert (0.0125, 0.025, 0.0375)
Differential[f]						
13	Uniform (0.15, 0.35)	Uniform (0.005, 0.015)	Uniform (0.005, 0.015)	Uniform (0.005, 0.015)	Uniform (0.01, 0.03)	Uniform (0.005, 0.015)

14	Uniform (0.05, 0.15)	Uniform (0.025, 0.075)	Uniform (0.05, 0.15)	Uniform (0.01, 0.03)	Uniform (0.05, 0.15)	Uniform (0.025, 0.075)
15	Triangular[c] (0.15, 0.25, 0.35)	Triangular (0.005, 0.01, 0.015)	Triangular (0.005, 0.01, 0.015)	Triangular (0.005, 0.01, 0.015)	Triangular (0.01, 0.02, 0.03)	Triangular (0.005, 0.01, 0.015)
16	Triangular (0.05, 0.1, 0.15)	Triangular (0.025, 0.05, 0.075)	Triangular (0.05, 0.10, 0.15)	Triangular (0.01, 0.02, 0.03)	Triangular (0.05, 0.1, 0.15)	Triangular (0.025, 0.05, 0.075)
17	BetaPert[d] (0.15, 0.25, 0.35)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.005, 0.01, 0.015)	BetaPert (0.01, 0.02, 0.03)	BetaPert (0.005, 0.01, 0.015)
18	BetaPert (0.05, 0.1, 0.15)	BetaPert (0.025, 0.05, 0.075)	BetaPert (0.05, 0.10, 0.15)	BetaPert (0.01, 0.02, 0.03)	BetaPert (0.05, 0.1, 0.15)	BetaPert (0.025, 0.05, 0.075)

[a]In expectation- the individual trials may only be approximately non-differential, but the average difference across all simulation trials is 0.

[b]Continuous uniform distribution (minimum, maximum).

[c]Triangular distribution (minimum, mode, maximum).

[d]BetaPERT distribution (minimum, likeliest, maximum).

[e]Cases have a greater proportion of misclassified.

[f]Non-cases have a greater proportion of misclassified.

three lowest cumulative TCDD exposure groups into one category, our initial distribution assumptions were likely more representative of the misclassification between the "never-exposed" group and the upper end of the collapsed category (*i.e.*, "0-2085.7 ppt-mo") or the lower end of the collapsed category and workers with an exposure ≥2085.8 ppt-mo. To capture the potential for greater misclassification between employees who were classified as "never-exposed" and the lower end of the collapsed category, as well as the upper end of the collapsed group and the highest exposure group, we multiplied the majority of distribution parameters by either a factor of two (e_{31}, e_{34}), five (e_{12}, e_{42}, e_{33}, e_{15}, e_{45}, e_{36}), or 10 (e_{13}, e_{16}). This increased the degree of misclassification to levels we believed to be reasonable between the three exposure groups (*e.g.*, we assumed misclassification between the highest exposure group and those workers who were never exposed would be less than that between workers in the never-exposed group and those in the "0-2085.7 ppt-mo" exposure group) [91, 92]. e_{21} and e_{24} were the only classification parameters that we decreased – by a factor of 2-3 – so they would be more consistent with the other levels of misclassification.

Next, we modified the six non-differential scenarios to define differential scenarios in which a greater proportion of cases were misclassified (scenarios 7-12) and a greater proportion of non-cases were misclassified (scenarios 13-18). In the former, the same probability distributions defined for the non-differential scenarios were used for the classification parameters for cases. For non-cases, we

set the minimum, maximum and mode/likeliest values for most classification parameters equal to that for cases divided by two. The distribution values for e_{24} in scenarios 7, 9 and 11 were divided by factors of 1.4 to 1.5. These changes were expected to shift the distributions for non-cases far enough to the left to ensure differential misclassification while preserving a practical level of misclassification. For scenarios 13-18, the same strategy was applied: the probability distributions specified for the non-differential scenarios were used for the classification parameters for non-cases, with the distributions for cases shifted slightly to the left.

Figure 1 illustrates the probability input required for any given simulation using the parameter distributions constructed for Scenario 12 (see Tables 4A and 4B) by cases (A) and non-cases (B). For example, the betaPert distributions for cases range from 0.05 to 0.15 for parameters e_{21}, e_{13}, and e_{42}, 0.025 to 0.075 for parameters e_{12} and e_{33}, and 0.01 to 0.03 for e_{31}. The misclassification proportions for non-cases also had betaPert distributions, ranging from 0.025 to 0.075 for parameters e_{24}, e_{16}, and e_{45}, 0.0125 to 0.0375 for parameters e_{15} and e_{36}, and 0.005 to 0.015 for e_{34}. To limit the degree of differentiality (*i.e.*, the amount of differential exposure misclassification between cases and non-cases) for any given simulation trial [56], we designated a correlation of 0.75 between the sampled proportions for cases and non-cases. This was done because, while misclassification could be differential, we had no reason to believe there would be

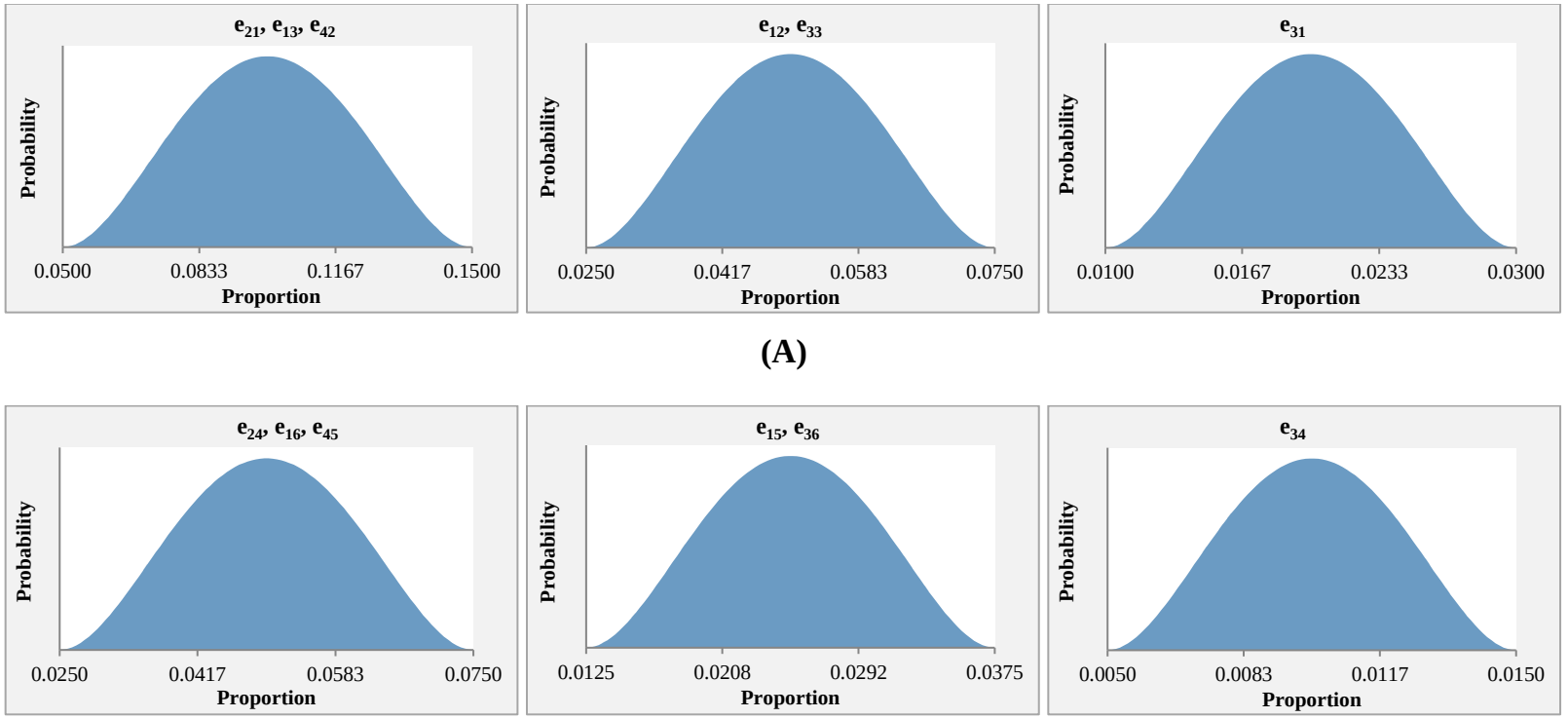


Figure 1. Parameter distributions for scenario 12 by cases (A) and non-cases (B).

extreme differences between the classification proportions for cases and non-cases in any given simulation trial.

For each scenario, the parameter distributions specified in the E_1 and E_2 matrices were sampled using Crystal Ball software (Oracle Corporation, Redwood Shores, CA) and then combined into a matrix of disease- and exposure-classification proportions (P matrix; Table 5). In this example, the classification proportions for disease misclassification, which is included in the P matrix for functional purposes, were set to one and zero. The inverse of the P matrix was then multiplied by the vector of observed cell counts using the matrix method of Greenland and Kleinbaum [40] (Figure 2) to generate adjusted counts of cases and non-cases by exposure group. These counts were then used to estimate odds ratios adjusted for exposure misclassification. Each analysis was conducted separately and was based on 50,000 re-samplings of the specified parameter distributions. We used the simulation output for each scenario to create frequency distributions for OR_{EM} and ε_{EM}. The lower 2.5 and upper 97.5 percentiles of each frequency distribution were used to estimate 95% certainty intervals for the distribution of adjusted odds ratios and error terms.

Results

Results for each probabilistic simulation are summarized in Table 6. None of the simulation trials produced negative numbers of adjusted cases or non-cases.

Table 5. P matrix as a function of exposure (E_1 and E_2) and disease[a] classification matrices.

		True Exposure and Disease Status					
		IHD Cases			*Non-cases*		
Classified Disease Status	**Classified Exposure Status**	Never-exposed	≥2085.8 ppt-mo	0-2085.7 ppt-mo	Never-exposed	≥2085.8 ppt-mo	0-2085.7 ppt-mo
IHD Cases	Never-exposed	$p_{11} = e_{11}$	$p_{12} = e_{12}$	$p_{13} = e_{13}$	$p_{14} = 0$	$p_{15} = 0$	$p_{16} = 0$
	≥2085.8 ppt-mo	$p_{21} = e_{21} \cdot e_{31}$	$p_{22} = e_{22} \cdot e_{32}$	$p_{23} = e_{23} \cdot e_{33}$	$p_{24} = 0$	$p_{25} = 0$	$p_{26} = 0$
	0-2085.7 ppt-mo	$p_{31} = e_{21} \cdot e_{41}$	$p_{32} = e_{22} \cdot e_{42}$	$p_{33} = e_{23} \cdot e_{43}$	$p_{34} = 0$	$p_{35} = 0$	$p_{36} = 0$
Non-cases	Never-exposed	$p_{41} = 0$	$p_{42} = 0$	$p_{43} = 0$	$p_{44} = e_{14}$	$p_{45} = e_{15}$	$p_{46} = e_{16}$
	≥2085.8 ppt-mo	$p_{51} = 0$	$p_{52} = 0$	$p_{53} = 0$	$p_{54} = e_{24} \cdot e_{34}$	$p_{55} = e_{25} \cdot e_{35}$	$p_{56} = e_{26} \cdot e_{36}$
	0-2085.7 ppt-mo	$p_{61} = 0$	$p_{62} = 0$	$p_{63} = 0$	$p_{64} = e_{24} \cdot e_{44}$	$p_{65} = e_{25} \cdot e_{45}$	$p_{66} = e_{26} \cdot e_{46}$
	Total	1	1	1	1	1	1

IHD, ischemic heart disease; ppt-mo, parts per trillion-month

[a]Disease classification parameters are set to zero and one, which assumes no disease misclassification.

		True Exposure and Disease Status								
		IHD Cases			*Non-cases*					
Classified Disease Status	**Classified Exposure Status**	Never-exposed	≥2085.8 ppt-mo	0-2085.7 ppt-mo	Never-exposed	≥2085.8 ppt-mo	0-2085.7 ppt-mo		Observed Counts	Adjusted Counts
IHD Cases	Never-exposed	1.1253484	-0.051486528	-0.128909879	0	0	0		14	8.97530
	≥2085.8 ppt-mo	0.0055745	1.176215547	-0.062558068	0	0	0		14	13.60483
	0-2085.7 ppt-mo	-0.130923	-0.124729018	1.191467947	0	0	0		47	52.41987
IHD Non-cases	Never-exposed	0	0	0	1.0555799	-0.025526115	-0.056326857	X	451	= 420.18634
	≥2085.8 ppt-mo	0	0	0	0.0008778	1.081059854	-0.027766869		148	134.70840
	0-2085.7 ppt-mo	0	0	0	-0.056458	-0.055533738	1.084093727		925	969.10525

Figure 2. Example inverse classification proportion matrix multiplied by the vector of observed cell counts to estimate adjusted counts. Values presented were taken from a single simulation trial under Scenario 12.

Table 6. Geometric mean (GM) and 95% certainty intervals for OR_{EM} and ε_{EM} after 50,000 simulation trials by scenario.

	OR_{EM}			ε_{EM}	
Scenario	**% of Trials with $OR_{EM} > OR_{observed}$**	**GM**	**95% Certainty Interval**	**GM**	**95% Certainty Interval**
Non-differential[a]					
1	83.0	3.21	2.85 – 3.61	0.95	0.84 – 1.07
2	99.9	5.05	3.58 – 7.60	0.60	0.40 – 0.85
3	90.6	3.21	2.97 – 3.47	0.95	0.88 – 1.03
4	100	4.96	3.90 – 6.61	0.61	0.46 – 0.78
5	92.1	3.21	2.99 – 3.44	0.95	0.88 – 1.02
6	100	4.94	3.96 – 6.41	0.62	0.47 – 0.77
Differential A[b]					
7	16.3	2.89	2.60 – 3.20	1.05	0.95 – 1.17
8	100	4.87	3.55 – 7.30	0.63	0.42 – 0.86
9	7.5	2.89	2.69 – 3.10	1.05	0.98 – 1.13
10	100	4.80	3.78 – 6.54	0.64	0.47 – 0.81
11	6.0	2.89	2.70 – 3.08	1.05	0.99 – 1.13
12	100	4.78	3.84 – 6.32	0.64	0.48 – 0.79
DifferentialB[c]					
13	99.4	3.47	3.13 – 3.86	0.88	0.79 – 0.97
14	95.3	3.81	2.96 – 5.00	0.80	0.61 – 1.03
15	100	3.47	3.23 – 3.74	0.88	0.82 – 0.94
16	99.4	3.80	3.17 – 4.64	0.80	0.66 – 0.96
17	100	3.47	3.25 – 3.71	0.88	0.82 – 0.94
18	99.7	3.80	3.21 – 4.55	0.80	0.67 – 0.95

OR_{EM}, odds ratio adjusted for exposure misclassification; $OR_{observed}$, odds ratio for observed data; ε_{EM}, error term for exposure misclassification

[a]In expectation- individual simulation trials may be differential.

[b]Cases have a greater proportion misclassified.

[c]Non-cases have a greater proportion misclassified.

The geometric mean of the odds ratio adjusted for exposure misclassification (OR_{EM}) ranged from 2.89 (certainty interval (CI): 2.60-3.20) to 5.05 (CI: 3.58-7.60). The geometric mean error factor had a range of 0.60 (CI: 0.40-0.85) to 1.05 (CI: 0.95-1.17). Varying the shape of the parameter distribution

had no impact on the geometric mean for either the error factor or adjusted odds ratio. However, we observed greater precision of the certainty intervals for scenarios where we specified the betaPert or triangular distributions. Not unexpectedly, the width of the certainty intervals was also affected by the degree of misclassification such that those scenarios with a greater degree of misclassification had much wider intervals.

For 14 of the 18 scenarios, greater than 90% of the simulation trials yielded adjusted ORs greater than the unadjusted $OR_{observed}$, producing a slight shift of the adjusted OR frequency distributions away from the crude OR and the null (Figures 3A-3C). In all non-differential scenarios (scenarios 1-6) and differential scenarios in which non-cases had greater proportions of misclassification (scenarios13-18), the $OR_{observed}$ was biased toward the null, although the impact of exposure misclassification was slightly attenuated in the differential scenarios compared to the non-differential scenarios.

The results for the differential simulations where cases had higher proportions of misclassification were more complex. Adjustment for misclassification in scenarios 8, 10, and 12, which ranged between 1% and 15% for cases and 0.5% and 7.5% for non-cases, demonstrated that the unadjusted odds ratio of 3.05 was also biased toward the null. In contrast, for scenarios 7, 9, and 11, where we specified only a minimal degree of misclassification for most classification parameters (*i.e.*, cases ranging from 0.5% to 1.5% and non-cases

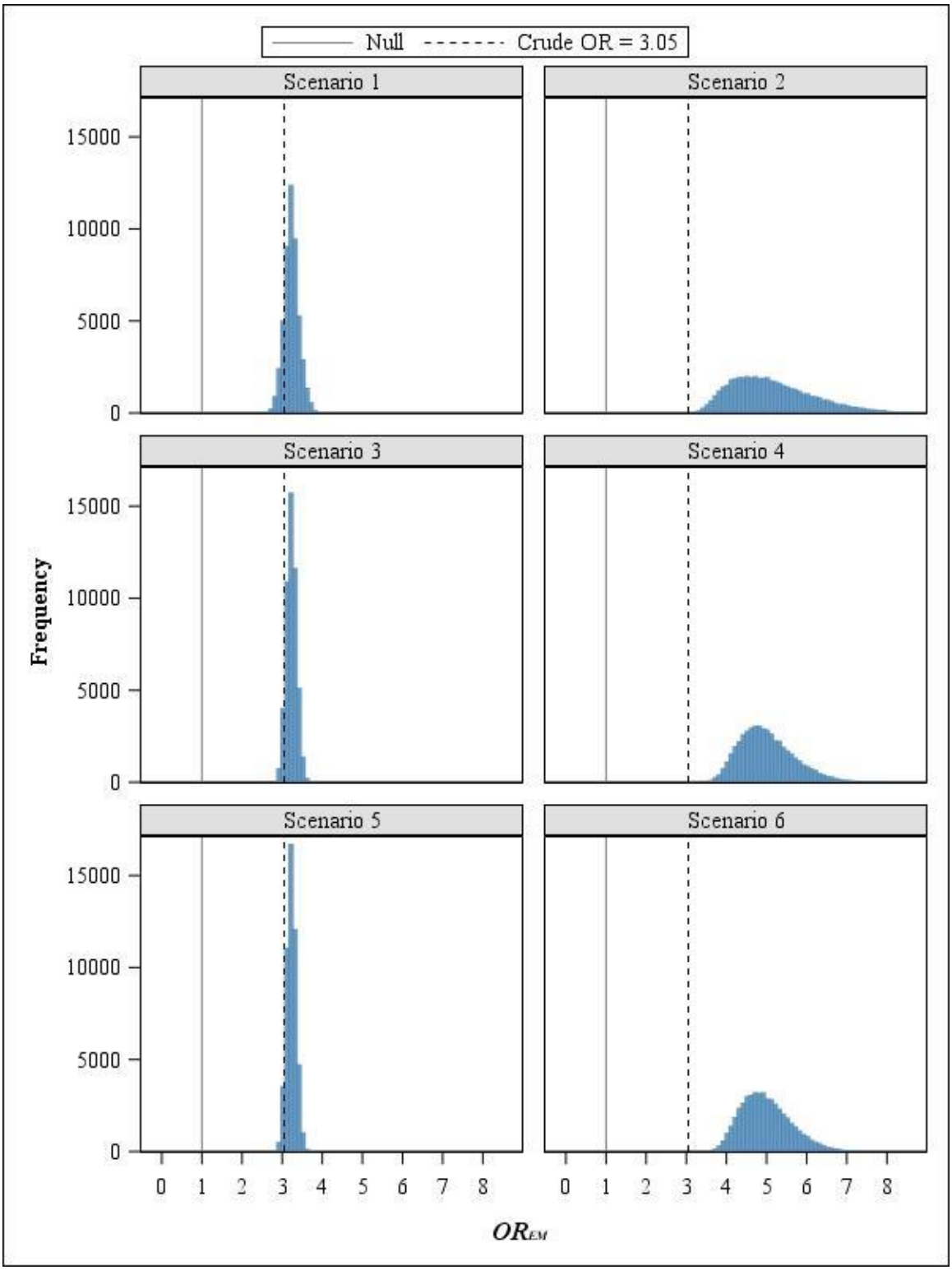


Figure 3A. Frequency distributions of OR_{EM} for non-differential scenarios.

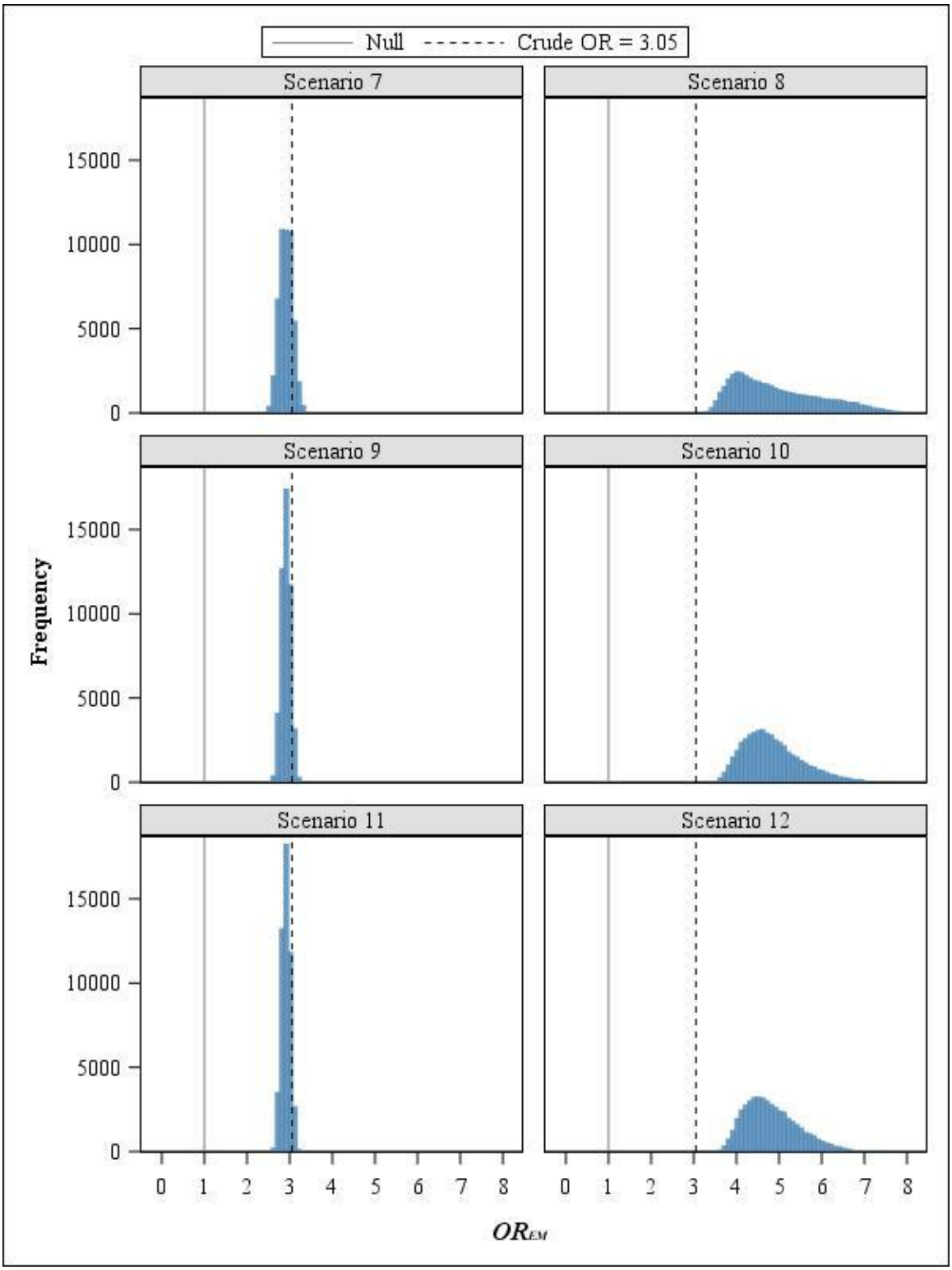


Figure 3B. Frequency distributions of OR_{EM} for differential A scenarios.

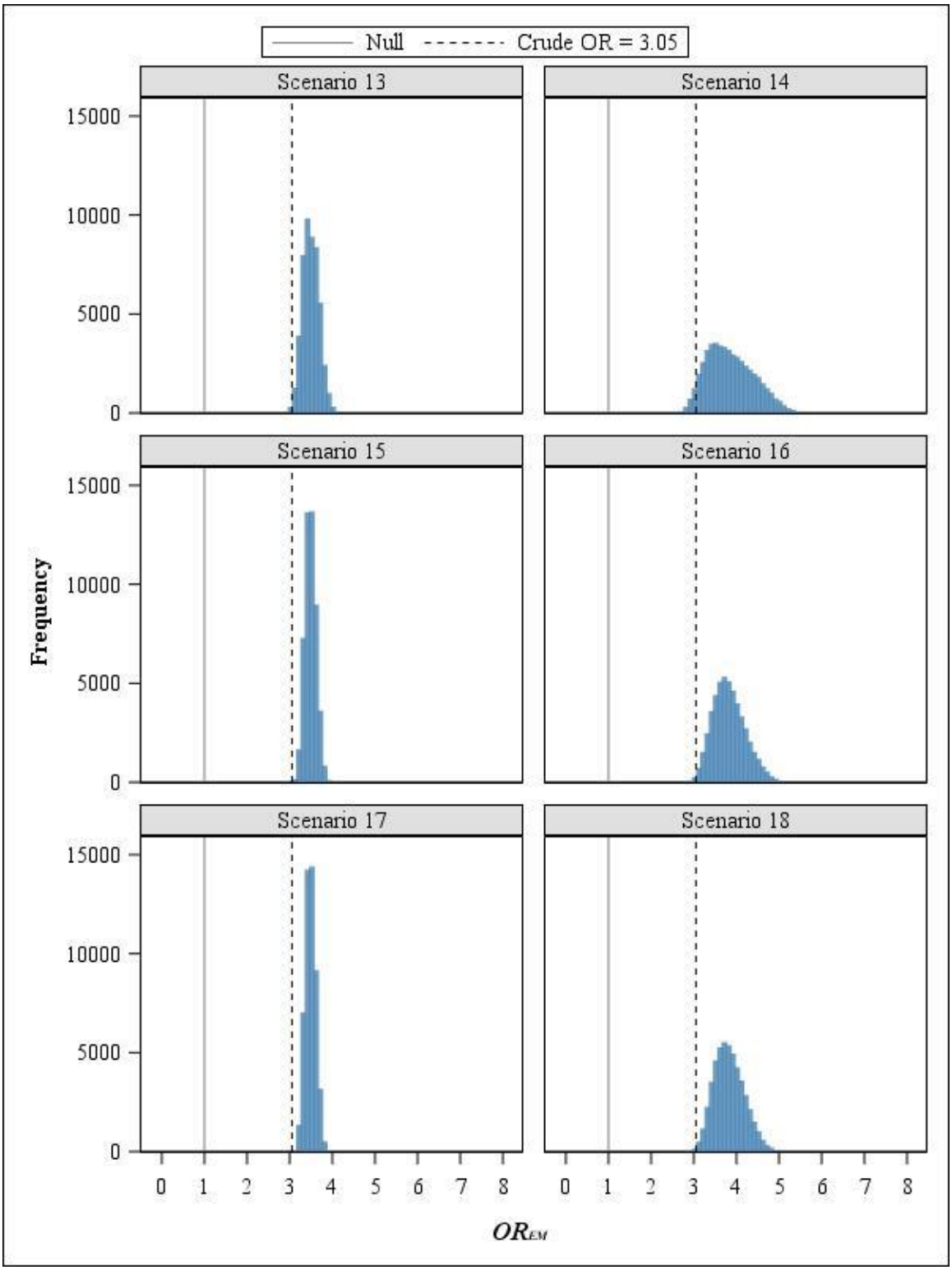


Figure 3C. Frequency distributions of OR_{EM} for differential B scenarios.

ranging from 0.25% to 0.75%), the geometric mean $\mathcal{E}_{EM}$ was considerably closer to one and the geometric mean OR_{EM} of 2.89 was smaller than the $OR_{observed}$.

To examine which parameters most influenced the direction of the error in scenarios 7, 9, and 11, we modified the proportions specified in scenario 11 for each classification pair individually (*i.e.*, e_{21} and e_{24}, e_{42} and e_{45}, etc.). Changing the classification proportions $e_{21/24}$, $e_{12/15}$, and $e_{42/45}$ produced the smallest difference (geometric mean OR_{EM} range: 3.02 – 3.07) from the unadjusted odds ratio, while independently increasing $e_{13/16}$, $e_{31/34}$, and $e_{33/36}$ had the greatest influence on the direction of the error. More precisely, varying $e_{13/16}$ produced a geometric mean OR_{EM} greater than $OR_{observed}$, but increasing the parameters $e_{31/34}$ or $e_{33/36}$ yielded geometric mean OR_{EM}s of 2.86 and 2.85, respectively, suggesting these parameters are driving the results observed for scenarios 7, 9, and 11.

Discussion

We conducted a probabilistic uncertainty analysis to quantify how one source of systematic error – exposure misclassification – may affect the exposure-disease response in a historical cohort of occupational workers. We chose to investigate several scenarios because very little published information existed for this cohort with which to validate exposure classification, no information about the shapes of the parameter distributions was available, and we had no reason to assume exposure misclassification was completely independent of disease status.

More importantly, we provided details about conducting such an analysis when exposure is a polytomous variable. Furthermore, we did not assume misclassification occurred only between adjacent exposure categories.

Similar to Rosenbaum *et al*. [93], our example demonstrates that even small amounts of exposure misclassification can noticeably influence an estimate of effect. In our study, the magnitude of impact fluctuated based on whether misclassification was differential or non-differential and the degree of misclassification. Equally noteworthy is the importance of the size of cell counts. Our findings suggest that moving even a few individuals from those cells with the smallest counts can considerably affect the direction and magnitude of the error. While the analyses described here provide some broad understanding of the nature and extent of the effects that differential and non-differential exposure misclassification may have on a measure of effect for the highest category in a multi-level exposure study, we suggest using caution when generalizing these results to other historical cohorts because the direction and degree of bias we observed may be specific to the data structure of the cohort used in this example.

Detailed exposure validation data was not available for the cohort of New Zealand workers studied by McBride and colleagues [22], and it wasn't possible to use data from other cohorts of TCDD-exposed chemical production workers since they generally had much higher exposures than those at the New Zealand TCP plant. Using the exposure information that was described in the published

literature for the NZ workers, we were able to form some initial assumptions as to what the parameter distributions looked like and then modify these scenarios in order to explore the sensitivity of $OR_{observed}$ to the effects of different levels and types of exposure misclassification. However, there is certainly the potential for higher proportions of exposure misclassification than we proposed in our analyses. The extraction of study data for probabilistic uncertainty analysis and specification of probability distributions for error-term parameters can, at times, be dependent on somewhat unrefined methods, with some assumptions merely being nothing more than one's "best guess." Undoubtedly, probabilistic bias analysis may be an uncertain means of quantifying error from exposure misclassification when no information or data about the level of misclassification can be obtained.

The uncertainty analyses described here do not constitute a complete probabilistic bias analysis as only one source of study error was assessed. In addition, our examples explicitly assume that there was no disease misclassfication, no selection bias, and no residual confounding, which is not likely to be accurate. Accordingly, it would be inappropriate to make any causal judgment regarding the exposure-disease relationship for TCDD and IHD mortality in this group of workers based on the results presented here. Nonetheless, our findings demonstrate that probabilistic uncertainty analysis of historical cohort mortality studies can be an effective tool for understanding trends in study error stemming from exposure misclassification in these types of studies

and confirms the importance of quantifying sources of systematic error prior to inferring whether an exposure-disease association is causal. Combining this probabilistic uncertainty model with other models we developed to adjust for sources of disease misclassification [39] is the next step in constructing a comprehensive model which researchers can use to quantitatively evaluate study bias in mortality studies. We encourage public health researchers to conduct such assessments of those historical cohort mortality studies with defined multi-level exposure categories, and hope further development of this method will produce an epidemiologic tool that is helpful in guiding human health risk assessments and public health policy decisions.

CHAPTER V

CONCLUSIONS

This body of work initially began as a "simple" assessment of the combined impact of exposure and disease misclassification on a measure of effect for a historical cohort mortality study. Preliminary analyses demonstrated that 1) use of the odds ratio as a measure of effect results in disease misclassification due to loss to follow up and 2) when disease-specific mortality – rather than incidence or total mortality – is the outcome of interest, sensitivity and specificity proportions can inappropriately be applied to living study subjects. A third challenge of conducting the initial bias analysis resulted from exposure being measured as a multi-level variable. As the initial analysis evolved to adequately address each of these challenges, three different probabilistic bias-analysis models were developed to form a complete dissertation. The primary objectives of this research were to illustrate how to adjust an odds ratio for different sources of systematic study error using Monte Carlo simulation methods and to evaluate the effects of these specific sources of disease and misclassification error.

For ease of implementation, the development of all three methods followed a general framework: 1) calculate an observed measure of effect, 2) identify bias-analysis parameters, 3) identify sources of data needed to inform probability distributions for bias-analysis parameters, 4) specify probability distributions for

each bias-analysis parameter, 5) sample probability distributions using Monte Carlo methods to correct the appropriate cell counts, 6) calculate an adjusted measure of effect using corrected cell counts, and 7) summarize the results. Each of these steps was described in detail along with any assumptions necessary to implement the methods. For example, since there was no reason to believe that either exposure or disease misclassification was only non-differential, this research included both non-differential and differential simulation scenarios. Equally important, these methods illustrated the use of both internal and external sources of validation data for specifying probability distributions.

Conventional practice when interpreting epidemiological findings most often incorrectly concludes systematic study error is non-differential and either has little or no impact on the observed exposure-disease relationship or biases an effect toward the null. However, application of probabilistic bias analyses to adjust for different types of misclassification may well be a considerable improvement over making such conjectures. Results for each of the three methods developed here demonstrated that various sources of error can impact a measure of effect differently and that even small amounts of misclassification can noticeably affect study results, particularly when cell counts are small. In addition, it was shown that, while non-differential misclassification typically produces a measure of effect biased toward the null, bias does occur in both directions and is much more unpredictable when misclassification is differential.

Despite the benefits of using the methods described here to quantify the effects of exposure and disease misclassification, implementation can be difficult. Although access to the original study data is not necessary, it allows for fewer assumptions when constructing probability distributions for each bias-analysis parameter. The availability and quality of external validation data can also limit the usefulness of these methods. Lastly, the process of selecting the appropriate distributions and integrating all relevant, and sometimes conflicting, information into those distributions can itself be prone to error and personal bias.

This research has provided a strong foundation for incorporating probabilistic bias-analysis methods into standard epidemiological studies. Yet, there are several questions that need to be investigated further including whether modifying the degree of differentiality impacts the simulation results and whether there are any potential differences between using an odds ratio or incidence proportion ratio as the measure of effect. Additionally, it is unknown if the observed trends in direction and magnitude of systematic error are specific to a study's data structure or if they are indicative of broad findings that could be generalized to all historical cohort mortality studies. Finally, it will be necessary to develop a comprehensive model which researchers can use to quantitatively evaluate the combined effect of these sources of systematic bias and identify the mathematical function that best describes their relationship. The results of this research have already revealed that adjustment for disease misclassification due to

incorrect cause of death data must be done prior to adjustment for disease misclassification resulting from loss to follow-up. However, the order of adjustment for other study errors – such as disease misclassification and confounding – will need to be ascertained.

The degree of uncertainty in an estimate of effect can be useful knowledge as it allows others, particularly risk assessors and those involved in policy and funding decisions, to evaluate whether a given point estimate is considerably prone to systematic study error and to what end. Hopefully, the practice of conducting probabilistic bias analyses will become more routine with the continued development and refinement of methods to account for sources of these errors.

BIBLIOGRAPHY

1. Vena, J.E., et al., *Sources of bias in retrospective cohort mortality studies: a note on treatment of subjects lost to follow-up.* J Occup Med, 1987. **29**(3): p. 256-61.
2. Swaen, G.M. and J.M. Meijers, *Influence of design characteristics on the outcome of retrospective cohort studies.* Br J Ind Med, 1988. **45**(9): p. 624-9.
3. Rothman, K.J., S. Greenland, and T.L. Lash, *Modern epidemiology*. 3rd ed. 2008, Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins. x, 758 p.
4. Lash, T.L., et al., *Good practices for quantitative bias analysis.* Int J Epidemiol, 2014.
5. Maldonado, G., *Adjusting a relative-risk estimate for study imperfections.* J Epidemiol Community Health, 2008. **62**(7): p. 655-63.
6. Maldonado, G., *Infomal evaluation of bias may be inadequate.* Am J Epidemiol, 1998. **147**(11 Suppl): p. S82.
7. Maldonado, G. and S. Greenland, *Estimating causal effects.* Int J Epidemiol, 2002. **31**(2): p. 422-9.
8. Phillips, C. and G. Maldonado, *Using Monte Carlo methods to quantify multiple sources of error in studies.* Am J Epidemiol, 1999. **149**(11 Suppl): p. S17.
9. Fox, M.P., T.L. Lash, and S. Greenland, *A method to automate probabilistic sensitivity analyses of misclassified binary variables.* Int J Epidemiol, 2005. **34**(6): p. 1370-6.
10. Greenland, S., *Multiple-bias modelling for analysis of observational data.* Journal of the Royal Statistical Society Series A, 2005. **168**: p. 267-306.
11. Lash, T.L. and A.K. Fink, *Semi-automated sensitivity analysis to assess systematic errors in observational data.* Epidemiology, 2003. **14**(4): p. 451-8.
12. Phillips, C.V., *Quantifying and reporting uncertainty from systematic errors.* Epidemiology, 2003. **14**(4): p. 459-66.
13. Lash, T.L., *Heuristic thinking and inference from observational epidemiology.* Epidemiology, 2007. **18**(1): p. 67-72.

14. Greenland, S., *Interval estimation by simulation as an alternative to and extension of confidence intervals.* Int J Epidemiol, 2004. **33**(6): p. 1389-97.
15. Lash, T.L., M.P. Fox, and A.K. Fink, *Applying quantitative bias analysis to epidemiologic data*. Statistics for biology and health. 2009, Dordrecht ; New York: Springer. xii, 192 p.
16. Jurek, A.M. and S. Greenland, *Adjusting for multiple-misclassified variables in a study using birth certificates.* Ann Epidemiol, 2013. **23**(8): p. 515-20.
17. Jurek, A.M., T.L. Lash, and G. Maldonado, *Specifying exposure classification parameters for sensitivity analysis: family breast cancer history.* Clin Epidemiol, 2009. **1**: p. 109-17.
18. Jurek, A.M., G. Maldonado, and S. Greenland, *Adjusting for outcome misclassification: the importance of accounting for case-control sampling and other forms of outcome-related selection.* Ann Epidemiol, 2013. **23**(3): p. 129-35.
19. Jurek, A.M., et al., *Periconceptional maternal vitamin supplementation and childhood leukaemia: an uncertainty analysis.* J Epidemiol Community Health, 2009. **63**(2): p. 168-72.
20. Savitz, D.A. and R. Moure, *Treatment of subjects lost to follow-up: effect on oil refinery cancer risks.* J Occup Med, 1988. **30**(2): p. 89-91.
21. Checkoway, H., N. Pearce, and D. Kriebel, *Research methods in occupational epidemiology*. 2nd ed. Monographs in epidemiology and biostatistics. 2004, New York: Oxford University Press. xiv, 372 p.
22. McBride, D.I., et al., *Mortality in workers exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin at a trichlorophenol plant in New Zealand.* J Occup Environ Med, 2009. **51**(9): p. 1049-56.
23. New Zealand Ministry of Health, *Mortality and Demographic Data 2008*. 2011: Wellington, NZ.
24. Vose, D. and D. Vose, *Risk analysis : a quantitative guide*. 2nd ed. 2000, Chichester ; New York: Wiley. x, 418 p.
25. Oracle Corporation, *Crystal Ball*. 2014: Redwood Shores, CA.

26. Jurek, A.M., et al., *Uncertainty analysis: an example of its application to estimating a survey proportion.* J Epidemiol Community Health, 2007. **61**(7): p. 650-4.
27. MacLehose, R.F. and P. Gustafson, *Is probabilistic bias analysis approximately Bayesian?* Epidemiology, 2012. **23**(1): p. 151-8.
28. Swift, B. and K. West, *Death certification: an audit of practice entering the 21st century.* J Clin Pathol, 2002. **55**(4): p. 275-9.
29. James, D.S. and A.D. Bull, *Information on death certificates: cause for concern?* J Clin Pathol, 1996. **49**(3): p. 213-6.
30. Weeramanthri, T.S., *Reporting deaths to the coroner. Death certification needs urgent overhaul.* BMJ, 1993. **306**(6891): p. 1539-40.
31. Sorlie, P.D. and E.B. Gold, *The effect of physician terminology preference on coronary heart disease mortality: an artifact uncovered by the 9th revision ICD.* Am J Public Health, 1987. **77**(2): p. 148-52.
32. Selikoff, I.J., *Use of death certificates in epidemiological studies, including occupational hazards: discordance with clinical and autopsy findings.* Am J Ind Med, 1992. **22**(4): p. 469-80.
33. Naghavi, M., et al., *Algorithms for enhancing public health utility of national causes-of-death data.* Popul Health Metr, 2010. **8**: p. 9.
34. Halanych, J.H., et al., *Agreement on cause of death between proxies, death certificates, and clinician adjudicators in the Reasons for Geographic and Racial Differences in Stroke (REGARDS) study.* Am J Epidemiol, 2011. **173**(11): p. 1319-26.
35. Edwards, J.K., et al., *Accounting for misclassified outcomes in binary regression models using multiple imputation with internal validation data.* Am J Epidemiol, 2013. **177**(9): p. 904-12.
36. Gilbert, R., et al., *Misclassification of outcome in case-control studies: Methods for sensitivity analysis.* Stat Methods Med Res, 2016. **25**(5): p. 2377-2393.
37. Lash, T.L., et al., *Methods to apply probabilistic bias analysis to summary estimates of association.* Pharmacoepidemiol Drug Saf, 2010. **19**(6): p. 638-44.

38. Zhang, Y. and K. Berhane, *Bayesian mixed hidden Markov models: a multi-level approach to modeling categorical outcomes with differential misclassification.* Stat Med, 2014. **33**(8): p. 1395-408.
39. Scott, L.L. and G. Maldonado, *Quantifying and Adjusting for Disease Misclassification Due to Loss to Follow-Up in Historical Cohort Mortality Studies.* Int J Environ Res Public Health, 2015. **12**(10): p. 12834-46.
40. Greenland, S. and D.G. Kleinbaum, *Correcting for misclassification in two-way tables and matched-pair studies.* Int J Epidemiol, 1983. **12**(1): p. 93-7.
41. The World Bank Group. *How Does the World Bank Classify Countries?* 2018 [cited 2018 July 1]; Available from: https://datahelpdesk.worldbank.org/knowledgebase/articles/378834-how-does-the-world-bank-classify-countries.
42. Jackson, R., et al., *Validation of coronary heart disease death certificate diagnoses.* N Z Med J, 1988. **101**(856 Pt 1): p. 658-60.
43. Agarwal, R., et al., *Overreporting of deaths from coronary heart disease in New York City hospitals, 2003.* Prev Chronic Dis, 2010. **7**(3): p. A47.
44. Coady, S.A., et al., *Validation of death certificate diagnosis for coronary heart disease: the Atherosclerosis Risk in Communities (ARIC) Study.* J Clin Epidemiol, 2001. **54**(1): p. 40-50.
45. Dobson, A.J., R.W. Gibberd, and S.R. Leeder, *Death certification and coding for ischemic heart disease in Australia.* Am J Epidemiol, 1983. **117**(4): p. 397-405.
46. Folsom, A.R., et al., *Out-of-hospital coronary death in an urban population--validation of death certificate diagnosis. The Minnesota Heart Survey.* Am J Epidemiol, 1987. **125**(6): p. 1012-8.
47. Goraya, T.Y., et al., *Validation of death certificate diagnosis of out-of-hospital coronary heart disease deaths in Olmsted County, Minnesota.* Mayo Clin Proc, 2000. **75**(7): p. 681-7.
48. Harriss, L.R., et al., *Accuracy of national mortality codes in identifying adjudicated cardiovascular deaths.* Aust N Z J Public Health, 2011. **35**(5): p. 466-76.

49. Iribarren, C., et al., *Validation of death certificate diagnosis of out-of-hospital sudden cardiac death.* Am J Cardiol, 1998. **82**(1): p. 50-3.
50. Lloyd-Jones, D.M., et al., *Accuracy of death certificates for coding coronary heart disease as the cause of death.* Ann Intern Med, 1998. **129**(12): p. 1020-6.
51. Mahonen, M., et al., *The validity of the routine mortality statistics on coronary heart disease in Finland: comparison with the FINMONICA MI register data for the years 1983-1992. Finnish multinational MONItoring of trends and determinants in CArdiovascular disease.* J Clin Epidemiol, 1999. **52**(2): p. 157-66.
52. Martin, C.A., M.S. Hobbs, and B.K. Armstrong, *Estimation of myocardial infarction mortality from routinely collected data in Western Australia.* J Chronic Dis, 1987. **40**(7): p. 661-9.
53. McIlwaine, W.J., et al., *Certification of death from ischaemic heart disease in Belfast.* Int J Epidemiol, 1985. **14**(4): p. 560-5.
54. Saito, I., *Review of death certificate diagnosis of coronary heart disease and heart failure in Japan.* Nihon Koshu Eisei Zasshi, 2004. **51**(11): p. 909-16.
55. Sexton, P.T., K. Jamrozik, and J.M. Walsh, *Death certification and coding for ischaemic heart disease in Tasmania.* Aust N Z J Med, 1992. **22**(2): p. 114-8.
56. Brenner, H., D.A. Savitz, and O. Gefeller, *The effects of joint misclassification of exposure and disease on epidemiologic measures of association.* J Clin Epidemiol, 1993. **46**(10): p. 1195-202.
57. World Health Organization (WHO), *Proposal for the Multi-national Monitoring of Trends and Determinants in Cardiovascular Disease and Provisional Protocol*, Cardiovascular Diseases Unit, Editor. 1981: Geneva, Switzerland.
58. Reid CM, Ryan P, and Wing LMH, *The 2nd Australian National Blood Pressure Study (ANBP2),* in *Clinical Trials in Hypertension*, H.R. Black, Editor. 2001, Marcel Dekker: New York. p. 587-604.
59. Reid, C.M., et al., *Feasibility of conducting cardiovascular outcome research in Australian general practice: results from the ANBP2 pilot study. Australian National Blood Pressure Study.* Clin Exp Pharmacol Physiol, 1997. **24**(5): p. 370-3.

60. The LIPID Study Group, *Design features and baseline characteristics of the LIPID (Long-Term Intervention with Pravastatin in Ischemic Disease) Study: a randomized trial in patients with previous acute myocardial infarction and/or unstable angina pectoris.* Am J Cardiol, 1995. **76**(7): p. 474-9.
61. The Long-Term Intervention with Pravastatin in Ischaemic Disease Study Group (The LIPID Study Group), *Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels.* N Engl J Med, 1998. **339**(19): p. 1349-57.
62. Gillum, R.F., et al., *International diagnostic criteria for acute myocardial infarction and acute stroke.* Am Heart J, 1984. **108**(1): p. 150-8.
63. Lee, M.H., N.O. Borhani, and L.H. Kuller, *Validation of reported myocardial infarction mortality in blacks and whites. A report from the Community Cardiovascular Surveillance Program.* Ann Epidemiol, 1990. **1**(1): p. 1-12.
64. New Zealand Health Information Service (NZHIS), *A guide to certifying causes of death : a guide for doctors and coroners on the provision of information on deaths to the New Zealand Health Information Service*. 2001, Wellington, N.Z.: New Zealand Health Information Service. 64 p.
65. Brown, S.H. and M. Frankovich, *How accurate are New Zealand death certificates?* N Z Med J, 1998. **111**(1072): p. 321-2.
66. Reid, A.W., *New Zealand death certificates.* N Z Med J, 1999. **112**(1082): p. 61.
67. Jurek, A.M., S. Greenland, and G. Maldonado, *How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null?* Int J Epidemiol, 2008. **37**(2): p. 382-5.
68. Murray, C.J., S.C. Kulkarni, and M. Ezzati, *Understanding the coronary heart disease versus total cardiovascular mortality paradox: a method to enhance the comparability of cardiovascular death statistics in the United States.* Circulation, 2006. **113**(17): p. 2071-81.
69. Blakely, T. and New Zealand. Ministry of Health., *The New Zealand census-mortality study : socioeconomic inequalities and adult mortality, 1991-94*. 2002, Wellington, N.Z.: Ministry of Health. 258.

70. Fett, M.J., *Measuring the accuracy of vital status data in cohort studies.* Am J Public Health, 1985. **75**(12): p. 1385-8.
71. Swart, A., et al., *Examining the quality of name code record linkage: what is the impact on death and cancer risk estimates? A validation study.* Aust N Z J Public Health, 2015. **39**(2): p. 141-7.
72. da Silveira, D.P. and E. Artmann, *Accuracy of probabilistic record linkage applied to health databases: systematic review.* Rev Saude Publica, 2009. **43**(5): p. 875-82.
73. Newman, T.B. and A.N. Brown, *Use of commercial record linkage software and vital statistics to identify patient deaths.* J Am Med Inform Assoc, 1997. **4**(3): p. 233-7.
74. Blair, A. and P.A. Stewart, *Do quantitative exposure assessments improve risk estimates in occupational studies of cancer?* Am J Ind Med, 1992. **21**(1): p. 53-63.
75. Checkoway, H. and E.A. Eisen, *Developments in occupational cohort studies.* Epidemiol Rev, 1998. **20**(1): p. 100-11.
76. Checkoway, H., N. Pearce, and D. Kriebel, *Selecting appropriate study designs to address specific research questions in occupational epidemiology.* Occup Environ Med, 2007. **64**(9): p. 633-8.
77. Parodi, S., et al., *Comparison bias and dilution effect in occupational cohort studies.* Int J Occup Environ Health, 2007. **13**(2): p. 143-52.
78. Dickinson, H.O., et al., *The sex ratio of children in relation to paternal preconceptional radiation dose: a study in Cumbria, northern England.* J Epidemiol Community Health, 1996. **50**(6): p. 645-52.
79. Greenberg, R.S., et al., *A meta-analysis of cohort studies describing mortality and cancer incidence among chemical workers in the United States and western Europe.* Epidemiology, 2001. **12**(6): p. 727-40.
80. Pearce, N., H. Checkoway, and D. Kriebel, *Bias in occupational epidemiology studies.* Occup Environ Med, 2007. **64**(8): p. 562-8.
81. Rosenman, K.D., *Cardiovascular disease and work place exposures.* Arch Environ Health, 1984. **39**(3): p. 218-24.

82. Fox, A.J. and P.F. Collier, *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): p. 225-30.
83. Aylward, L.L., et al., *Concentration-dependent TCDD elimination kinetics in humans: toxicokinetic modeling for moderately to highly exposed adults from Seveso, Italy, and Vienna, Austria, and impact on dose estimates for the NIOSH cohort.* J Expo Anal Environ Epidemiol, 2005. **15**(1): p. 51-65.
84. Kjuus, H., *Estimation of Etiologic Fractions: The Importance of the Reference Group,* in *Impact of Work Load and Work Exposures on Disease Incidence in the Nordic Countries: Proceedings and Papers from the Workshop for Nordic Researchers in Copenhagen, 14-15 March 1996*, O. Olsen, Editor. 1997, Nordic Council of Ministers: Copenhagen.
85. Shore, R.E., et al., *Use of human data in quantitative risk assessment of carcinogens: impact on epidemiologic practice and the regulatory process.* Regul Toxicol Pharmacol, 1992. **15**(2 Pt 1): p. 180-221.
86. Hemon, D., *[Epidemiology of occupational risks: methodological problems]*. Rev Epidemiol Sante Publique, 1986. **34**(4-5): p. 230-6.
87. Marsh, G.M. and R.J. Caplan, *Evaluating Health Effects of Exposure at Hazardous Waste Sites: A Review of the State-of-the-Art, with Recommendations for Future Research*, in *Health Effects from Hazardous Waste Sites*, J.B. Andelman and D.W. Underhill, Editors. 1987, Lewis Publishers: Chelsea.
88. Aylward, L.L., et al., *TCDD exposure estimation for workers at a New Zealand 2,4,5-T manufacturing facility based on serum sampling data.* J Expo Sci Environ Epidemiol, 2010. **20**(5): p. 417-26.
89. Collins, J.J., et al., *Serum concentrations of chlorinated dibenzo-p-dioxins and dibenzofurans among former New Zealand trichlorophenol workers.* Chemosphere, 2009. **76**(11): p. 1550-6.
90. Bates, M.N., et al., *Persistent organochlorines in the serum of the non-occupationally exposed New Zealand population.* Chemosphere, 2004. **54**(10): p. 1431-43.

91. Marshall, J.R., et al., *On the distortion of risk estimates in multiple exposure level case-control studies.* Am J Epidemiol, 1981. **113**(4): p. 464-73.
92. Birkett, N.J., *Effect of nondifferential misclassification on estimates of odds ratios with multiple levels of exposure.* Am J Epidemiol, 1992. **136**(3): p. 356-62.
93. Rosenbaum, W.L., T.D. Sterling, and J.J. Weinkam, *Correcting standardized rate ratios for imprecise classification of a polychotomous exposure variable with limited data.* Am J Epidemiol, 1995. **142**(4): p. 442-5.

APPENDIX A

Table A1. ICD-9 and ICD-10 codes for most diseases of the circulatory system.

Description of ICD-9 Code	ICD-9 Code	ICD-10 Code	Description of ICD-10 Code
Essential hypertension	401	I10	Essential (primary) hypertension
Hypertensive heart disease	402	I11	Hypertensive heart disease
Hypertensive chronic kidney disease	403	I12	Hypertensive chronic kidney disease
Hypertensive heart and chronic kidney disease	404	I13	Hypertensive heart and chronic kidney disease
Secondary hypertension	405	I15	Secondary hypertension
---	---	I16	Hypertensive crisis
Acute myocardial infarction	410	I21	Acute myocardial infarction
Other acute and subacute forms of ischemic heart disease	411	I24	Other acute ischemic heart diseases
Old myocardial infarction	412	I25	Chronic ischemic heart disease
Angina pectoris	413	I20	Angina pectoris
Other forms of chronic ischemic heart disease	414	I25	Chronic ischemic heart disease
---	---	I22	Subsequent ST elevation and non-ST elevation myocardial infarction
Certain sequelae of myocardial infarction not elsewhere classified	429.79	I23	Certain current complications following ST elevation and non-ST elevation myocardial infarction (within the 28 day period)
Acute pulmonary heart disease	415	I26	Pulmonary embolism
Chronic pulmonary heart disease	416	I27	Other pulmonary heart disease
Other diseases of pulmonary circulation	417	I28	Other diseases of pulmonary vessels
Acute pericarditis	420	I30	Acute pericarditis
		I32	Pericarditis in diseases classified elsewhere
Acute and subacute endocarditis	421	I33	Acute and subacute endocarditis
Acute myocarditis	422	I40	Acute myocarditis
		I41	Myocarditis in diseases classified elsewhere
Other diseases of the pericardium	423	I31	Other diseases of pericardium

Other diseases of the endocardium	424	I34	Nonrheumatic mitral valve disorders
		I35	Nonrheumatic aortic valve disorders
		I36	Nonrheumatic tricuspid valve disorders
		I37	Nonrheumatic pulmonary valve disorders
		I38	Endocarditis, valve unspecified
		I39	Endocarditis and heart valve disorders in diseases classified elsewhere
Cardiomyopathy	425	I42	Cardiomyopathy
		I43	Cardiomyopathy in diseases classified elsewhere
Conduction disorders	426	I44	Atrioventricular and left bundle-branch block
		I45	Other conduction disorders
Cardiac dysrhythmias	427	I46	Cardiac arrest
		I47	Paroxysmal tachycardia
		I48	Atrial fibrillation and flutter
		I49	Other cardiac arrhythmias
Heart failure	428	I50	Heart failure
Ill-defined descriptions and complications of heart disease	429	I51	Complications and ill-defined descriptions of heart disease
---	---	I52	Other heart disorders in diseases classified elsewhere

ICD, international classification of diseases
Shaded codes are for those considered to constitute ischemic heart disease