STATISTICAL METHODS FOR EVALUATING EXPOSURE-HEALTH RELATIONSHIPS

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CIVIL AND ENVIRONMENTAL ENGINEERING

Nur H. Orak

B.S., Environmental Engineering, Istanbul Technical University, Turkey

M.S., Environmental Sciences and Engineering, Istanbul Technical University, Turkey

Carnegie Mellon University

Pittsburgh, PA

August 2016

This dissertation is dedicated to my beloved mother Vasfiye, my father Abdülhadi, and my dear siblings Nilay Merve and Mehmet Safa for their unconditional love, endless support, and lifetime encouragement.

ACKNOWLEDGEMENTS

I would like to acknowledge thankfully the financial support provided by the Republic of Turkey Ministry of National Education Fellowship Program and Carnegie Mellon University's Dean's Fellowship Program.

I would like to express my sincere appreciation to my adviser, Prof. Mitchell J. Small, who became a mentor for me over the last four years and guided me to become a lifelong researcher. He encouraged me to learn new things, ask critical questions, and love what I do. He always supported my independence and enthusiasm in research. He will forever play a role in my future success. I hope to emulate his mentorship with my future students.

I am deeply grateful to my committee members, Prof. Matteo Pozzi, Prof. Kelvin Gregory, and Prof. Marek J. Druzdzel, for their guidance and critical feedback during my Ph.D. They made an exceptional interdisciplinary research committee that inspired me to ask challenging questions and find novel paths to solve them.

I would also like to thank a few special people without whose support I would not have been able to achieve this dissertation. The former Civil and Environmental Engineering (CEE) department head, Prof. James H. Garrett, for his encouragement and support; the assistant director of the Global Communication Center, Nisha Shanmugaraj, for her academic writing guidance and patience; my editor, Amanda Fratrik, for her outstanding work; all the CEE staff, especially Ms. Maxine Leffard, for their dedication, assistance, and friendship.

I feel blessed to have the support of my dearest friends and colleagues in Pittsburgh and Istanbul. I thank all of you for all of your patience and lifetime friendship. I will always miss our ladies' tea times and late-night studying sessions. Finally, I would like to dedicate this dissertation

to my family, who supported me even in the darkest times, from thousands of miles away in Turkey, without fully understanding my research. The wisdom of my mom, Vasfiye, and my father, Abdulhadi, thought me to serve humanity, work hard, and be patient all the time. And I need to thank them now in Turkish . . .

Sevgili ailem, bu hayatta inandığım işin peşinden koşmama, benim ben olmama izin verdiğiniz için, koşulsuz sevgi ve desteğiniz için çok teşekkür ederim. Araştırma cesaretini annemden, merakını babamdan, sevgisini ablamdan, ve azmini kardeşimden öğrenebileceğim bir aile verdiği için Yaradan'a sonsuz şükürlerimi sunuyorum.

THESIS ABSTRACT

Conventional experimental techniques are sometimes limited in their ability to assess the actual risk of chemical exposures. Therefore, there is a rising awareness of mathematical, computational, and statistical approaches to provide insight into the adverse effects of environmental contaminants. Richard Bach once wrote: "Any powerful idea is absolutely fascinating and absolutely useless until we choose to use it." Likewise, any data may be viewed as absolutely fascinating and absolutely useless until we choose to understand and use it. Recent advances in science and technology provide alternative paths to develop effective risk-assessment methods for environmental contaminants. Moreover, these methods are more efficient in terms of time and cost. Therefore, I develop three Chapters to show the importance of statistical methods in environmental-health risk assessment, and highlight the potency of data-driven knowledge and multidisciplinary research for the future of environmental science and engineering.

In Chapter 1, I review the potential risks of missing chemical data and concentration variability on mixture toxicity by developing 27 occurrence scenarios based on data from the literature. The @RISK software simulates random concentrations, assuming multivariate lognormal distributions for the mixture components. In Chapter 2, I demonstrate how a performance analysis can be implemented for a Bayesian Network (BN) representation of a dose-response relationship. I explore the effect of different sample sizes on predicting the strength of the relationship between true responses and true doses of environmental toxicants. In Chapter 3, I characterize the risk factors of a prenatal arsenic exposure network by using Bayesian Network (BN) modeling as a tool for health risk assessment.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1: IMPLICATIONS OF A STATISTICAL OCCURRENCE MODEL FOR MIXTURE TOXICITY ESTIMATION

This chapter, written by Nur H. Orak and co-authored by Mitchell J. Small, is being reviewed by the *Journal of Human and Ecological Risk Assessment*, 2016, ID: BHER-2016-0215

Award: First Place, World Environmental & Water Resources Congress 2016, Graduate Student Technical Paper Competition

ABSTRACT

Assessing chemical mixture toxicity presents serious challenges to human and ecosystem health scientists. To provide insight into the potential implications of chemical co-occurrence, as well as the effects of incomplete chemical identification and/or inclusion that could occur in a risk assessment, this study characterizes the risk implications of concentration variability and correlation among co-acting compounds, considering the effect of missing chemical data. The potential risks of missing chemical data and concentration variability on mixture toxicity are explored by developing a set of multiple occurrence scenarios for mixtures, including a range of low to high toxicity chemicals, exhibiting low to high concentration variability, varying degrees of inter-chemical correlation, and omission of different chemicals in the mixture toxicity calculation.

The calculations are performed for hypothetical mixtures of a group of ten synthetic antibiotics ("aquinolones") that have been tested on marine bacterium for the endpoint of long-term bioluminescence inhibition to fit dose-response relationships for each, with mixture toxicities computed and compared for the assumptions of independent joint action theory and concentration/dose addition theory. These methods yield different toxicity estimates, but a similar direction and magnitude of effects from concentration variability and compared omission. I recommend a pre-assessment of the effect of different chemical occurrence patterns on mixture toxicity computed using different models for chemical interaction. This will help prioritize the experiments needed to differentiate between these models when estimating mixture effects.

1. INTRODUCTION

Problems associated with chemical mixture occurrence and toxicity pose a number of challenges to human and ecological risk assessment and management. Currently, approximately 85,000 unique chemicals are registered under the Toxic Substances Control Act (TSCA) (USEPA 2015), yet only a small part of this group is well characterized, that is, their toxicities are defined for regulatory actions (Judson et al. 2008). Currently, laboratory methods used to characterize environmental risks for the majority of these chemicals can only measure toxicities above a certain concentration; additionally these methods are expensive and time consuming (Jin et al. 2014). Consequently, these experimental methods have been applied to only a very small number of chemicals or groups of chemicals (Ryker and Small 2008), suggesting that we are only seeing the tip of a vast iceberg, with the undefined chemicals constituting the invisible part and posing potential risks. Another challenge associated with human and ecological risk management is that chemicals in the real world work in mixtures, so understanding the combined effect (mixture toxicity) of a group of chemicals is also important for regulatory actions (Konemann and Pieters 1996; Altenburger et al. 2012). However, it is impossible to test and characterize all possible combinations of chemicals through experimental methods (Cassee et al. 1998; Hadrup 2014). In addition, both single chemicals and those in mixtures often occur below the detection limit, and thus are generally omitted from chemical assessment (Zwart and Posthuma 2005; Beyer et al. 2014; Altenburger et al. 2015). Finally, regulated limit values are generally specified for single compounds even though humans are exposed to mixtures of chemicals (Evans et al. 2015).

A growing awareness about chemical mixtures has shown the limitations of experimental methods on mixture toxicity, and a need for an integrated risk assessment approach. These include a lack of knowledge of which chemicals to include in a comprehensive mixture

assessment; the large number of combinations under which these different chemicals can occur; the time and resources required to perform assays (that measure cellular toxicity) for all or even some of these combinations; and a lack of knowledge of occurrence patterns of groups of chemicals in the environment and associated exposure media. Recently, regulatory actions have demonstrated an awareness that single chemical risk assessment may underestimate the actual toxicity of chemicals (Altenburger et al. 2012).

The growing field of computational toxicology is successful in estimating the unknown toxicities of single chemicals; however, this method has limitations for mixture toxicity prediction (Kim et al. 2012), in part because only highly restrictive assumptions such as those made in additivity theory, provide feasible and efficient results for mixture toxicity prediction. In particular, the inclusion of interaction terms (e.g., synergism and antagonism) among chemicals creates too many degrees of freedom for estimation from limited toxicity studies (Altenburger et al. 2009; Molgaard et al. 2012), especially those based on field studies in which chemical concentrations vary over time and space, and some are below detection limits for measurement. Prior to characterizing these effects on model identification and parameter estimation, it is necessary to understand their implications in a predictive model of mixture toxicity. This study aims to identify the effect of concentration (and exposure) variability and correlation on mixture toxicity, as well as the effects of missing chemical data. Data from a recent study of quinolones (Backhaus et al. 1999) are used to predict single and mixture toxicities, and an extensive set of statistical occurrence scenarios are generated to investigate the effects of concentration variability, correlation and missing data on toxicity estimates.

2. METHODS

The mixture model used in this study was built on the experimental work of Backhaus et al. (1999) from which, 10 (ten) different compounds were selected: conixacin, enoxacin, flumequine, lomefloxacin, nalidixic acid, norflozacin, ofloxacin, oxolinic acid, pipedimic acid, and piromidic acid. All these chemicals belong to an important group of synthetic antibiotics called quinolones and these chemicals have been tested on marine bacterium *Vibrio fischeri* (Backhaus et al. 1999). The study has analyzed the long-term bioluminescence inhibition of *V. fischeri* and tested each quinolone at different concentrations to determine individual and mixture toxicities. Also, they estimated mixture toxicities by concentration addition and independent action theories; there is a small difference between the effect concentrations of the two predictions. The main reasons for choosing this paper are the detailed documentation of compounds, their chemical structures, and conducting both experimental and modeling studies. Therefore, we can use the paper to establish our predicted individual toxicities and develop the mixture scenarios based on those results.

As a first step, I compute concentration-response relationships for individual chemicals and mixtures of these chemicals as identified in Section 2.1. To develop a base case of no variability in concentrations, median values were chosen for each chemical's concentration as determined in Section 2.1. Twenty-seven (27) variability scenarios were then generated (Section 2.2) to analyze the effect of concentration variability on toxicity. The effects of concentration correlation on toxicity and examined in Section 2.3.

In general, there are two theoretical additivity models for mixture toxicities: i) concentration/dose addition; and ii) independent joint action (IA) (Bliss 1939; Drescherand and Boedeker 1995; Altenburger et al. 2012). What is important here is that both theories assume

that there is no interaction between single substances in a mixture in the biological environment (Gregorio et al. 2013).

Independent joint action theory (IA) assumes the mixture toxicity of independently acting chemicals can be estimated by the product of single responses of substances (toxicities) denoted by ($E(c_i)$) (Bliss 1939; Ashford 1981; Kamo and Yokomizo 2015):

$$E(c_{mix}) = E(c_1 + \cdots + c_N) = 1 - \prod_{i=1}^{n}[1 - E(c_i)] \quad (1)$$

Alternatively, concentration/dose addition theory assumes that the effect of a mixture of these chemicals ($E(c_{mix})$) can be predicted by the sum of the equivalent concentrations (doses) (C_i) (Bliss 1939; Liu et al. 2015). This approach is usually limited to compounds with similar modes of action. To implement concentration addition (CA) a reference compound is first selected, often the most toxic of the mixture component compounds, which is ofloxacin in the application that follows. Then equivalent concentrations of the reference compound ($C_{ref,i}$) are computed for each compound, yielding the same toxicity. The total effect of the mixture $E_{ref}(C_{ref,eq})$ is then calculated with the dose-response equation for the reference component. For ofloxacin this is the Weibull dose-response function, so that Weibull regression model:

$$C_{ref,i} = E_{ref}^{-1}\{E_i[C_i]\}$$

$$C_{ref,eq} = \sum_{i=1}^{n} C_{ref,i}$$

$$E(c_{mix}) = 1 - \exp(-\exp\left(\theta_1 + \theta_2 log_{10}\left(C_{ref,eq}\right)\right)) \quad (2)$$

In this study, it is first assumed that mixture toxicity follows the independent joint action model in Equation 1, and is not influenced by interactions among the chemicals, providing a baseline characterization of the effects of chemical variability and omission on toxicity prediction for independently acting chemicals. Another reason of choosing IA model is analyzing each compounds' effect on mixture scenarios. For comparison, concentration addition theory is applied to a subset of the scenarios those with no correlation.

2.1 Concentration-response relationships

The individual toxicity concentration-response models (see Table 1) and the concentration-response relationship parameters (see Table 2) were obtained from Backhaus et al. (1999). The half-maximal concentrations (EC50) were calculated for each compound and added to Table 2 for toxicity comparison in other sections. According to the EC50 values, ofloxacin is the most, and pipedimic acid is the least toxic of the compounds. Oxolinic acid was chosen to represent the intermediate toxicity compounds for some of the comparisons that follow.

First, individual toxicities and the total mixture toxicity were calculated for selected concentrations (0.0001 µmole/L – 10 µmole/L) to generate concentration-response curves for each compound (see Figure 1).

Table 1. Concentration-response models used for calculating the concentration response relationships for the long-term bioluminescence inhibition of *V. fischeri* (Backhaus et al. 1999)

Regression model	**Formula**
Weibull	$\mathbf{E(c) = 1 - \exp(-\exp(\theta_1 + \theta_2 \log_{10}(c)))}$
Generalized Logit	$\boldsymbol{E(c) = \frac{1}{(1 + \exp(-\theta_1 - \theta_2 log_{10}(c)))^{\theta_3}}}$
Box-Cox Weibull	$\boldsymbol{E(c) = 1 - \exp(-\exp\left(\theta_1 + \theta_2 \frac{c^{\theta_3} - 1}{\theta_3}\right))}$

[a]E(c) denotes the effect of a concentration c, given the two (θ_1, θ_2) or three (θ_1, θ_2,θ_3) parameters for the relationships.

Table 2. Concentration-response relationship parameters of the ten compounds (adapted from Backhaus, Scholze, & Grimme, 1999)

	Mixture components			
	Fit	$\widehat{\boldsymbol{\theta}}_1$	$\widehat{\boldsymbol{\theta}}_2$	$\widehat{\boldsymbol{\theta}}_3$
Pipedimic acid	Box-Cox Weibull	-3.942	2.153	0.495
Nalidixic acid	Generalized Logit	-8.213	128.48	0.042
Cinoxacin	Generalized Logit	0.848	6.289	0.437
Piromidic acid	Weibull	1.018	3.687	
Enoxacin	Generalized Logit	16.557	26.812	0.133
Oxolinic acid	Generalized Logit	145.193	159.45	0.0297
Flumequine	Generalized Logit	100.582	103.12	0.0413
Norfloxacin	Weibull	7.497	6.77	-
Lomefloxacin	Generalized Logit	71.694	72.11	0.0475
Ofloxacin	Weibull	4.829	3.648	-

A relationship for the total mixture toxicity with each component (assumed to be at the given concentration) was also computed using Equation 1 and was plotted in Figure 1. The total mixture toxicity curve exceeds that of any individual component, although only by a moderate

amount compared to the most toxic compound, ofloxacin (the total mixture has ~2 times higher toxicity than ofloxacin alone, over much of the indicated concentration range).

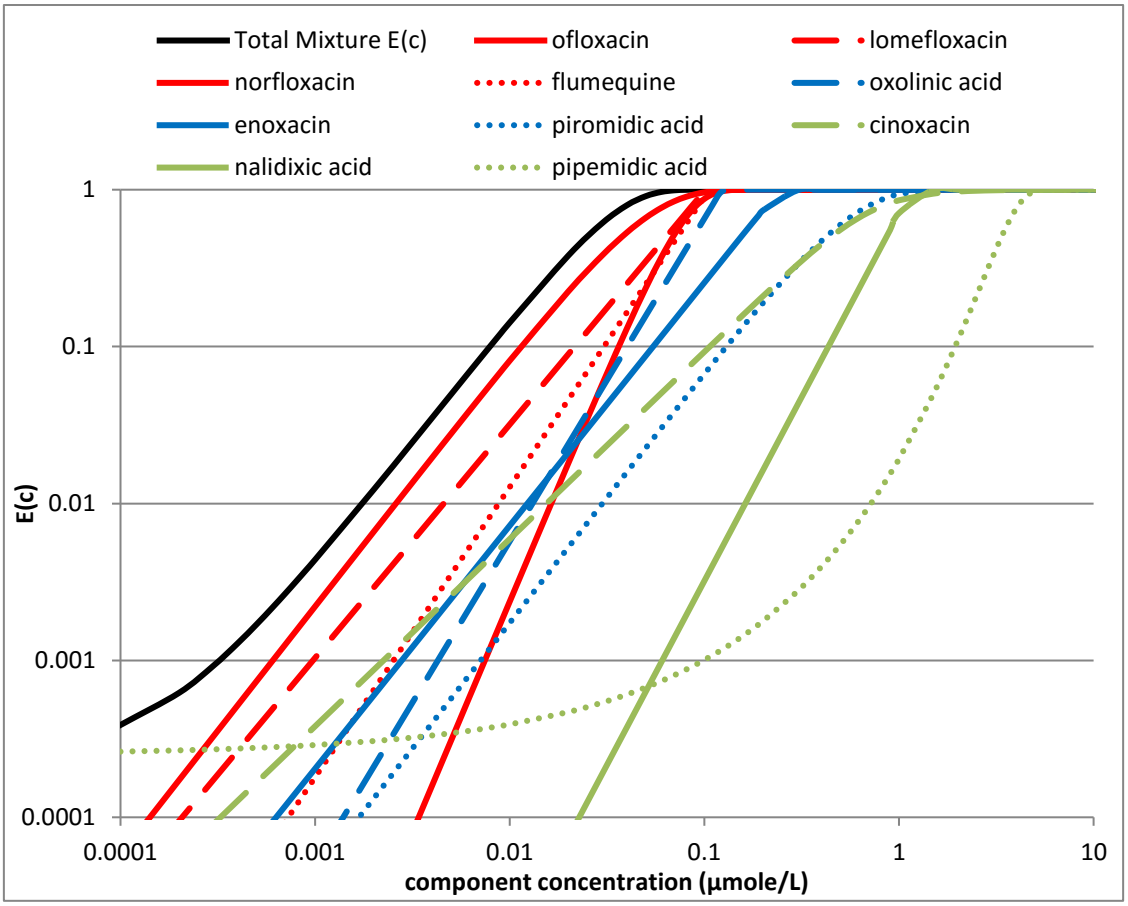


Figure 1. Concentration-response Relationships

2.2 Effect of Concentration Variability

Several factors affect the concentration variability of chemicals in a natural water system. Some are climatic features (e.g. temperature, precipitation, and seasonal changes such as photoperiod), anthropogenic events, and the physico-chemical character of a site. The first step of model evaluation was to identify the effect of concentration variability on mixture toxicity in conjunction with the concentration-response relationships. The multivariate lognormal distribution was chosen to describe the joint distribution of chemical concentrations at a hypothetical site.

A lognormal distribution defines the concentrations of compounds (c). Different median (C_{50}) and coefficient of variation (v) values were chosen to determine the lognormal parameters a and b, corresponding to the mean and standard deviation of ln(c), respectively. The values of a and b may also be determined from the median concentration (C_{50}) and the coefficient of variation of the concentration (υ) as:

$$a = \ln(C_{50})$$

$$b = \{\ln[\upsilon^2 + 1]\}^{\frac{1}{2}}$$

(3)

The median and coefficient of variation are used to describe the 27 scenarios in Table 3.

Simulated compound concentrations and calculated mixture toxicities for each scenario were generated for 1000 samples by the @RISK software and repeated for the twenty-seven (27) scenarios in Table 3. The simulation sample size (N) was selected to yield sufficient convergence after comparing different sample size results. Latin hypercube sampling (LHS) was used for the simulation. LHS sampling is a stratified sampling method designed to represent multidimensional distributions with a smaller sample size than can be achieved with random sampling (USEPA 1997).

The joint lognormal concentration distribution of the ten component species is thus specified as: C_i; i=1,10 ~ LN(a_i; i=1,10; b_i; i=1,10; [r_{ij}]) where the a_i are determined from the selected median values, the b_i are calculated with coefficient of variations, and [r_{ij}] is the correlation coefficient matrix between the ln concentration ofchemical i and j. Table 3 shows that each of the 27 cases assumes a common lognormal concentration distribution that applies to each of the 10 component species, ranging from the case of very low concentrations and no

variability (Case 1: median=0.0001 µmole/L, ν=0) to the case with very high concentrations and high variability (Case 27: median=5 µmole/L, ν=3).

Median (C_{50}) and coefficient of variation (ν) values apply to all ten component species for each case. For these cases species concentrations are assumed to be independent; $r_{ij}=0$.

Table 3. Summary of twenty-seven (27) scenarios

Scenario	**Median(C_{50}) µmole/L**	**ν**	**Scenario**	**Median(C_{50}) µmole/L**	**ν**	**Scenario**	**Median(C_{50}) µmole/L**	**ν**
1	0.0001	0	10	0.0001	1	19	0.0001	3
2	0.001	0	11	0.001	1	20	0.001	3
3	0.003	0	12	0.003	1	21	0.003	3
4	0.01	0	13	0.01	1	22	0.01	3
5	0.03	0	14	0.03	1	23	0.03	3
6	0.1	0	15	0.1	1	24	0.1	3
7	0.3	0	16	0.3	1	25	0.3	3
8	1.0	0	17	1.0	1	26	1.0	3
9	5.0	0	18	5.0	1	27	5.0	3

In this study, the mixture chemical interaction mechanism was first considered an independent joint action, that is, it is assumed that there is no interaction among the chemicals. The mixture toxicities were calculated by independent joint action theory using single toxicities of compounds ($E(c_i)$) as stated in Equation 1 (Backhaus et al. 1999).

2.3 Effect of Chemical Omission

Omission scenarios were generated to determine the effect on the estimated toxicity of missing (unmeasured) observations and the effect of chemicals that were measured but were below the detection limit and assumed to be not present ($c_i=0$).

Omission factors were calculated as stated in 4 to analyze the effect of omitting compounds in a chemical mixture.

$$OF_j = \text{omission factor for chemical j} = \left\{\frac{1-\prod_{i=1,i\neq j}^{n}[1-E(C_i)]}{1-\prod_{i=1}^{n}[1-E(C_i)]}\right\} \quad (4)$$

As indicated, OF_j is the ratio of the mixture toxicity computed with compound j omitted to the total mixture toxicity.

2.4 Effect of Concentration Correlation

As a final step of the model, the effects of correlated concentrations on toxicity were analyzed. A joint lognormal distribution is assumed with correlation coefficients (r_{ij}) of +0.4 and +0.9 between $\ln(C_i)$ and $\ln(C_j)$. Correlated concentrations were simulated for 5000 samples by the @RISK software for each of the 27 scenarios, to predict individual and mixture toxicities. The full set of scenarios simulated includes 27 cases with no correlation between species (see Table 3), 27 cases with moderate correlation between species (r_{ij}=0.4), and 27 cases with high correlation (r_{ij}=0.9).

3. RESULTS AND DISCUSSIONS

3.1 Effects of Variability and Correlation

As was pointed out in Section 2, single concentration-response relationships were modeled to be used for mixture toxicity simulation.

The effect of variability in all components on individual and mixture toxicity was tested by generating different sets of simulated concentrations using two alternative coefficients of variation (ν=1 and ν=3). Plots of toxicity versus median concentration for three of the ten compounds (representative of high, medium and low toxicity) are shown for graphic comparison: ofloxacin, oxolinic acid and pipemidic acid shown in Figures 2b, c, and d. The x-axis represents each single component's median concentration as previously shown in Table 3; the median concentration for the total mixture (Figure 2a) is thus ten times that of each single component concentration.

As shown in Figure 2, I found that higher variability in concentrations does cause higher effective (average) mixture toxicity when low to intermediate toxicity is associated with the median of the concentration distribution. The biggest enhancement of toxicity due to variability occurs at lower concentrations, since the presence of variability allows for occasional high concentrations and very high associated toxicity. However, if the coefficient of variation is very high, toxicity is slightly lower for high median concentration values because occasionally high and low concentration values occur in the system and the effect of very high concentrations is capped as $E[c_i]$ approaches 1.0 over much of the range of median concentrations considered. The total mixture $E[c_i]$ (Figure 2a) increases by a factor of 2.2 in changing from the case of no variability (ν=0 to ν=1) and by a factor of 11.5 in changing from ν=0 to ν=3. In this high variability case (ν=3), estimates of total mixture toxicity based on median concentrations will significantly underestimate the expected toxicity.

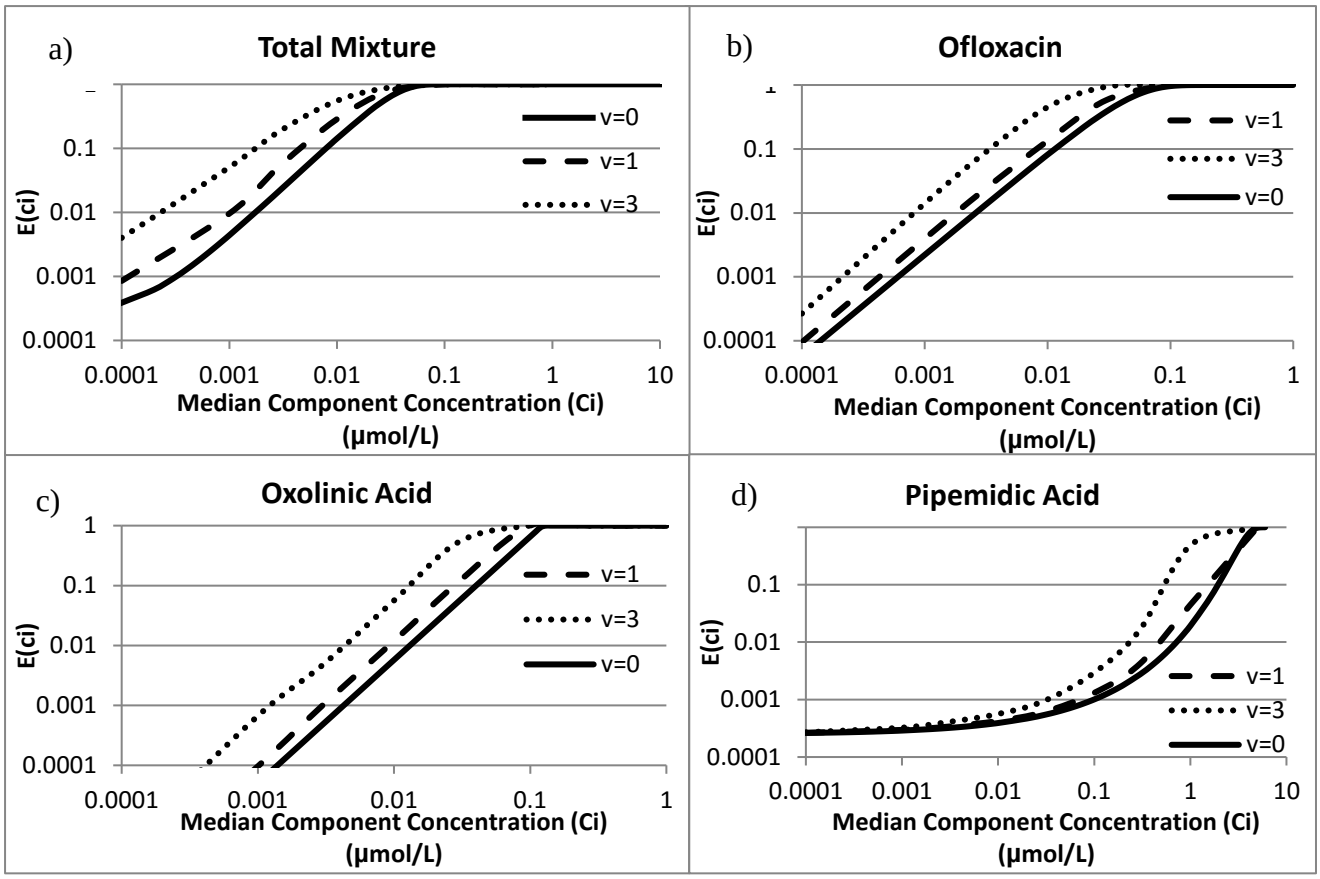


Figure 2. Effect of varying median concentration and coefficient of variation on individual chemical toxicity and total mixture toxicity, $E[C_i]$; i=1,10, r=0

As Figure 2a shows, the total mixture toxicity is very similar to the response in ofloxacin (2b), which means that the most toxic compound is the dominant component in the total mixture.

Figure 3 outlines the results of IA theory versus CA theory. IA results are slightly lower than the CA results and the occurrence of different levels of variability does not affect this trend. Note that the mixture toxicity prediction results for the CA model are somewhat higher than for the IA predictions. However, the increased toxicity effect due to concentration variability (moving from υ=0 to υ=3) is similar for the IA and CA models.

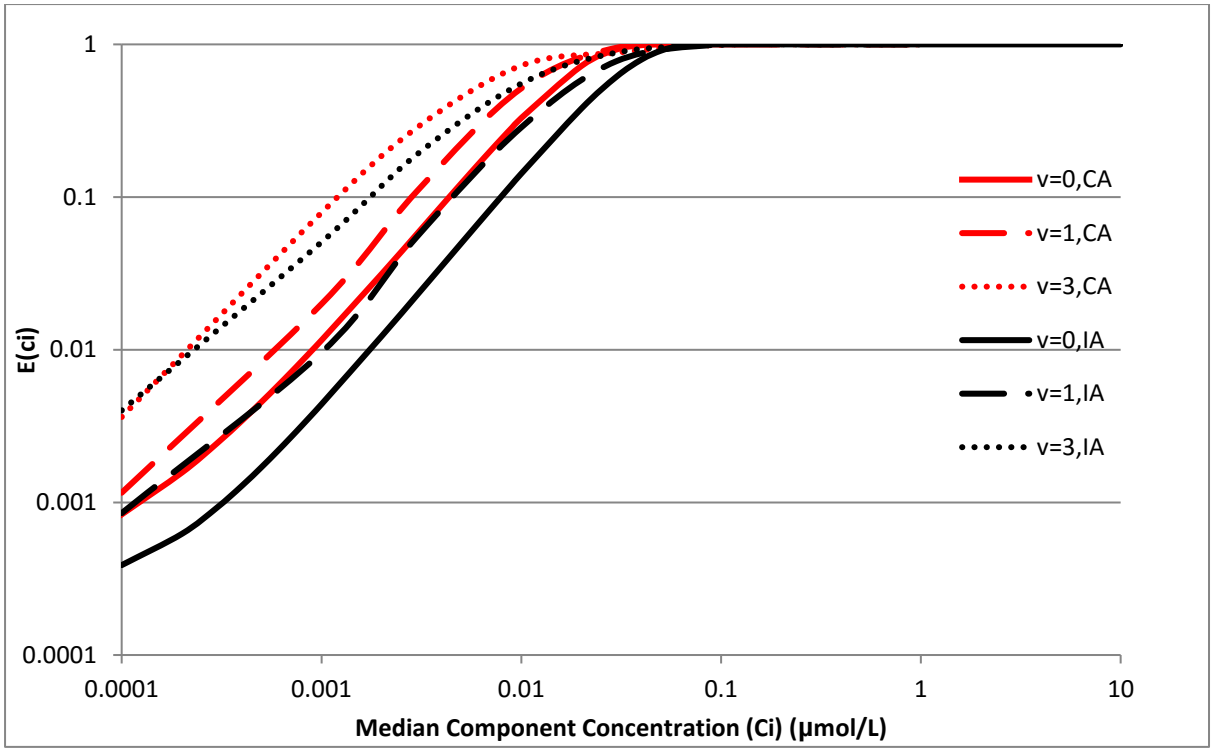


Figure 3. Independent action versus concentration addition, effect of varying median concentration and coefficient of variation on total mixture toxicity, E[C_i]; i=1,10, r=0

Figure 4 and 5 show the results of the simulations for estimating the effect of correlated concentration. Contrary to our original hypothesis, our findings show that positively correlated concentrations do not systematically lead to increased mixture toxicity. In the low variability scenario (ν=1), a change in correlation coefficients causes differences in toxicity at high concentrations, but correlated median concentrations decrease the mixture toxicity: $E(ci)_{\nu=1,r=0} > E(ci)_{\nu=1,r=+0.4} > E(ci)_{\nu=1,r=+0.9}$ (see Figure 4). The high coefficient of variation (ν=3) case follows a similar trend as the low coefficient of variation (ν=1) case, although causes higher mixture toxicity for the same median concentrations. However, a part of this toxicity occurred as a result of the higher coefficient of variation (Figure 5). Thus, in Figure 5, the gap between the trend for ν=0 and $\nu=3_{r=0}$ shows the effect of the coefficient of variation on toxicity as analyzed in Figure 2a. The gap between r=0, r=+0.4 and r=+0.9 delineates the effect of correlation coefficient in both Figures 4 and 5. This gap is bigger for higher correlation.

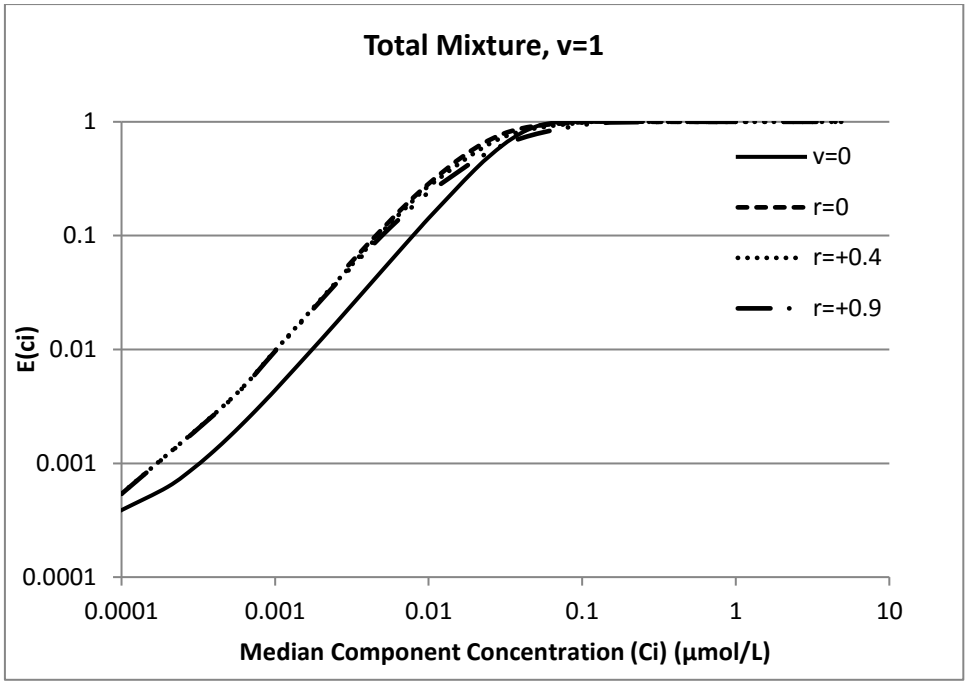


Figure 4. Effect of correlated concentrations on total mixture toxicity, ν=1, E[Ci]; i=1,10, IA model

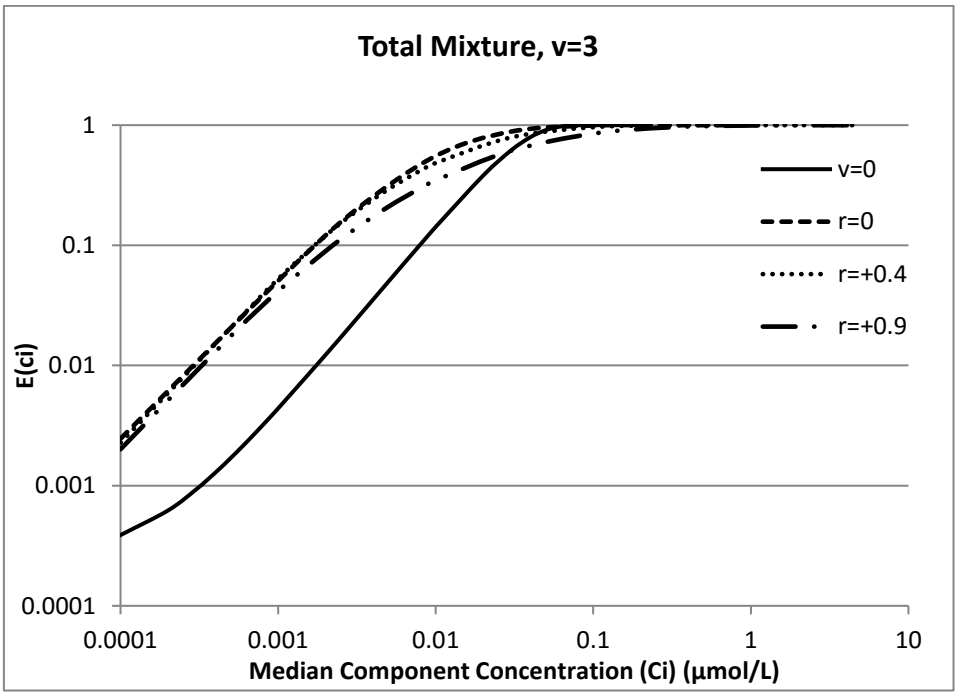


Figure 5. Effect of correlated concentrations on total mixture toxicity, ν=3, E[Ci]; i=1,10, IA model

As shown, the correlation coefficient has only a slight effect on mixture toxicity.

These results evoke a question: Is the risk of estimation of mixture toxicity higher for higher variability scenarios? To answer that question the effects of omission were analyzed.

3.2 Effects of Species Omission

First, omission factors (OF) are compared for each component at the fixed (no variability) concentrations of all components. Table 4 indicates that there is no significant difference between omission factors for C=0.001 μmole/L and C=0.01 μmole/L component concentrations. The OF results suggest that the increasing individual component concentrations (when ν=0) do not substantively change the impact of the omitted component's toxicity on the mixture toxicity. Second, the effect of variation on the omission scenarios is analyzed. Figure 6 outlines the estimated toxicity for the total mixture and three omitting scenarios under the occurrence of variability. As Figure 6 shows, omitting the least toxic compound in the system does not significantly affect the mixture toxicity. However, omitting the most toxic compound (ofloxacin) causes a notable decrease in the mixture toxicity. The ν=1 scenario yields a smaller reduction in the estimated toxicity (compared to the total mixture toxicity) for the compounds than do the ν=0 or ν=3 scenarios.

Table 4. Omission factors for ten components, ν=0, C=0.001 μmole/L and C=0.01 μmole/L

	Omission Factor (OF) for C=0.001 μmole/L, ν=0	**Omission Factor (OF) for C=0.01 μmole/L, ν=0**	**Omission Factor (OF) for C=0.001 μmole/L, ν=0**	**Omission Factor (OF) for C=0.01 μmole/L, ν=0**	**IA/CA C=0.001**	**IA/CA C=0.01**
Pipemidic acid	0.93	1	0.85	0.98	1.09	1.02
Nalidixic acid	1	1	1	1	1.00	1.00
Cinoxacin	0.91	0.96	0.82	0.91	1.10	1.06
Piromidic acid	0.99	0.99	0.95	0.96	1.04	1.03
Enoxacin	0.95	0.96	0.88	0.90	1.08	1.07
Oxolinic acid	0.99	0.97	0.95	0.91	1.04	1.06
Flumequine	0.96	0.92	0.89	0.85	1.08	1.08
Norfloxacin	1	0.99	0.99	0.95	1.01	1.04
Lomefloxacin	0.76	0.8	0.68	0.74	1.12	1.09
Ofloxacin	0.5	0.47	0.51	0.52	0.99	0.90

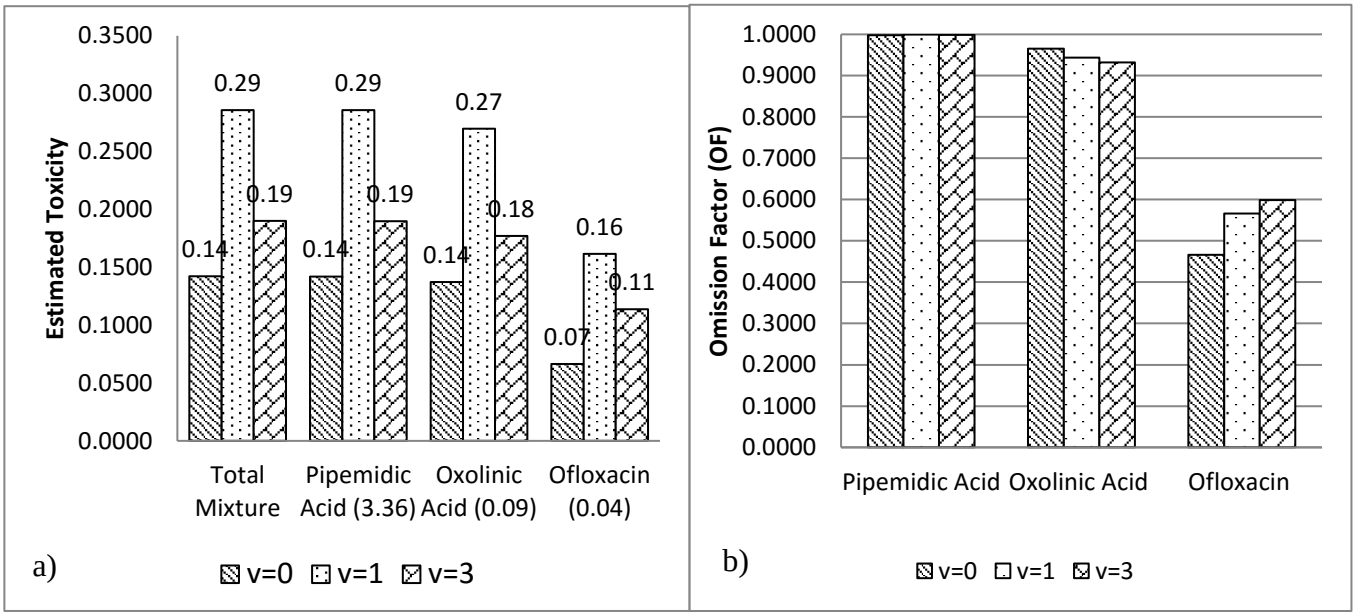


Figure 6. Omitting scenarios for different coefficient of variations (C=0.01 µmole/L), EC50 values shown in parenthesis

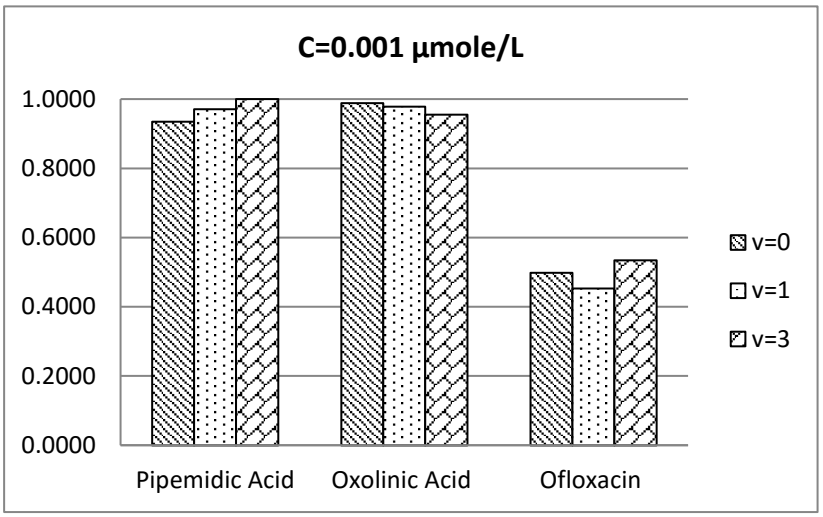


Figure 7. Omission factor (OF) for different omitting scenarios and for different coefficient of variations, v=0, v=1 and v=3, x-axis: omission factor

Figure 7 shows that variability has only a relatively small effect on the omission factor. Moreover, the OF does not change significantly with baseline concentration. Omitting pipemidic acid in the case of low median concentrations causes lower OF (i.e., 0.93) than in the case of high median concentration (i.e., 1.00), because pipemidic acid exhibits a different shape dose-response curve than the other compounds. A low concentration of pipemidic acid

causes higher toxicity than other chemicals but this relationship changes for high concentrations.

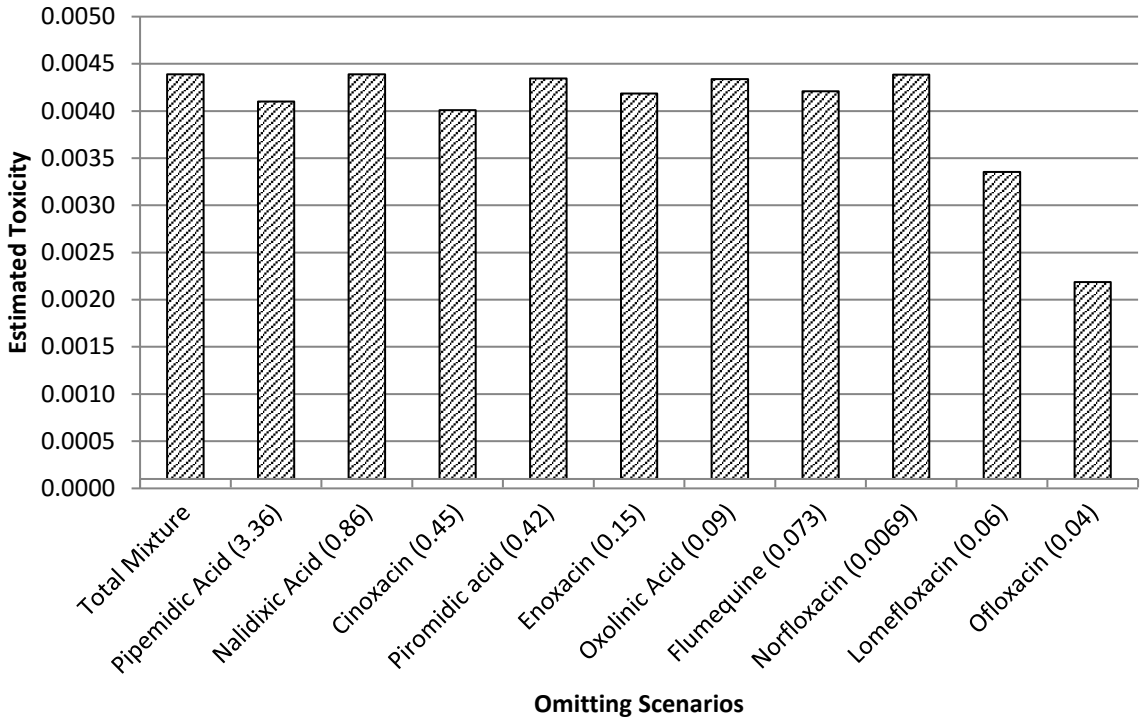
An important limitation of this study is that the models are based on a specific set of 10 chemicals, with results that may not be representative of other mixtures. However, we tried to generalize the outcomes by categorizing the compounds as the most, moderate, and least toxic. Also, for this study, mixture toxicity is dominated by the most toxic compound, ofloxacin, but this result could be different if no chemical dominates the mixture toxicity.

4. CONCLUSION

Current measurement and assessment techniques are limited in their ability to provide enough information about chemical mixtures' toxicities. This predictive occurrence model study indicates that the occurrence of variability in the system and correlated concentrations might have an important effect on mixture toxicities under specific conditions. There are several natural and anthropogenic reasons for concentration variability, so these results can be useful to analyze real case scenarios. In addition, omitting compounds (especially those with toxicities) can cause large underestimation of the mixture toxicities although not in all cases. These findings show that mixture occurrence and toxicity estimation should be explored concurrently to prioritize exposure sampling and mixture toxicity studies, and to conduct more accurate risk assessments for mixtures.

5. APPENDIX A

EXTRA CALCULATIONS AND FIGURES



Effect of Single Chemical Omission, EC_{50} shown in parenthesis, C=0.001 µmole/L, ν=0

Omit two compounds; $OF_{i,j}$

$$OF_{j,k} = omission\ factor\ for\ chemical\ j\ and\ k = \left\{\frac{1-\prod_{i=1,i\neq j,k}^{n}[1-E(C_i)]}{1-\prod_{i=1}^{n}[1-E(C_i)]}\right\}$$

	Omission Factor ($OR_{i,j}$) for C=0.001 µmole/L
Ofloxacin & Lomefloxacin (high toxicity range)	0.25
Oxolinic Acid & Enoxacin (middle toxicity range)	0.92
Pipemidic Acid & Cinoxacin (low toxicity range)	0.96
Ofloxacin & Pipemidic Acid (high and low toxicity)	0.46

$OF_{ofloxacin} * OF_{lomefloxacin} = 0.47*0.80 = 0.376 \neq OF_{ofloxacin,lomefloxacin}$

$OF_{oxolinic\ acid} * OF_{enoxacin} = 0.97*0.96 = 0.931$ close to $OF_{oxolinic\ acid,enoxacin}$

$OF_{pipemidic\ acid} * OF_{cinoxacin} = 1.00*0.96 = 0.96 = OF_{pipemidic\ acid,\ cinoxacin}$

$OF_{ofloxacin} * OF_{pipemidic\ acid} = 0.47*1.00 = 0.47$ close to $OF_{ofloxacin,\ pipemidic\ acid}$

Case	Assumed median, r=+0.4	Simulated median, r=+0.4	Assumed median, r=+0.9	Simulated median, r=+0.9	c.o.v.
1	0.0001	0.0001	0.0001	0.0001	1
2	0.0005	0.0005	0.0005	0.0005	1
3	0.001	0.001	0.001	0.001	1
4	0.002	0.002	0.002	0.002	1
5	0.003	0.003	0.003	0.003	1
6	0.005	0.005	0.005	0.005	1
7	0.008	0.008	0.008	0.008	1
8	0.01	0.01	0.01	0.01	1
9	0.03	0.03	0.03	0.03	1
10	0.1	0.1	0.1	0.1	1
11	0.3	0.3	0.3	0.3	1
12	1	1	1	1.0	1
13	5	5	5	5.0	1
14	0.0001	0.0001	0.0001	0.0001	3
15	0.0005	0.0005	0.0005	0.0005	3
16	0.001	0.001	0.001	0.001	3
17	0.002	0.002	0.002	0.002	3
18	0.003	0.003	0.003	0.003	3
19	0.005	0.005	0.005	0.005	3
20	0.008	0.008	0.008	0.008	3
21	0.01	0.01	0.01	0.01	3
22	0.03	0.03	0.03	0.03	3
23	0.1	0.1	0.1	0.1	3
24	0.3	0.3	0.3	0.3	3
25	1	1.0	1	1.0	3
26	5	5.0	5	5.0	3

Assumed correlation matrix, r=+0.4

	Ofl.	Lomef.	Norflo.	Flume.	Oxol.	Enox.	Pirom.	Cinox.	Nalid.	Pipem.
Ofloxacin	1									
Lomefloxacin	0.4	1								
Norfloxacin	0.4	0.4	1							
Flumequine	0.4	0.4	0.4	1						
Oxolinic Acid	0.4	0.4	0.4	0.4	1					
Enoxacin	0.4	0.4	0.4	0.4	0.4	1				
Piromidic Acid	0.4	0.4	0.4	0.4	0.4	0.4	1			
Cinoxacin	0.4	0.4	0.4	0.4	0.4	0.4	0.4	1		
Nalidixic Acid	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	1	
Pipemidic Acid	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	1

Simulated correlation matrix, r=+0.4 (example case 1)

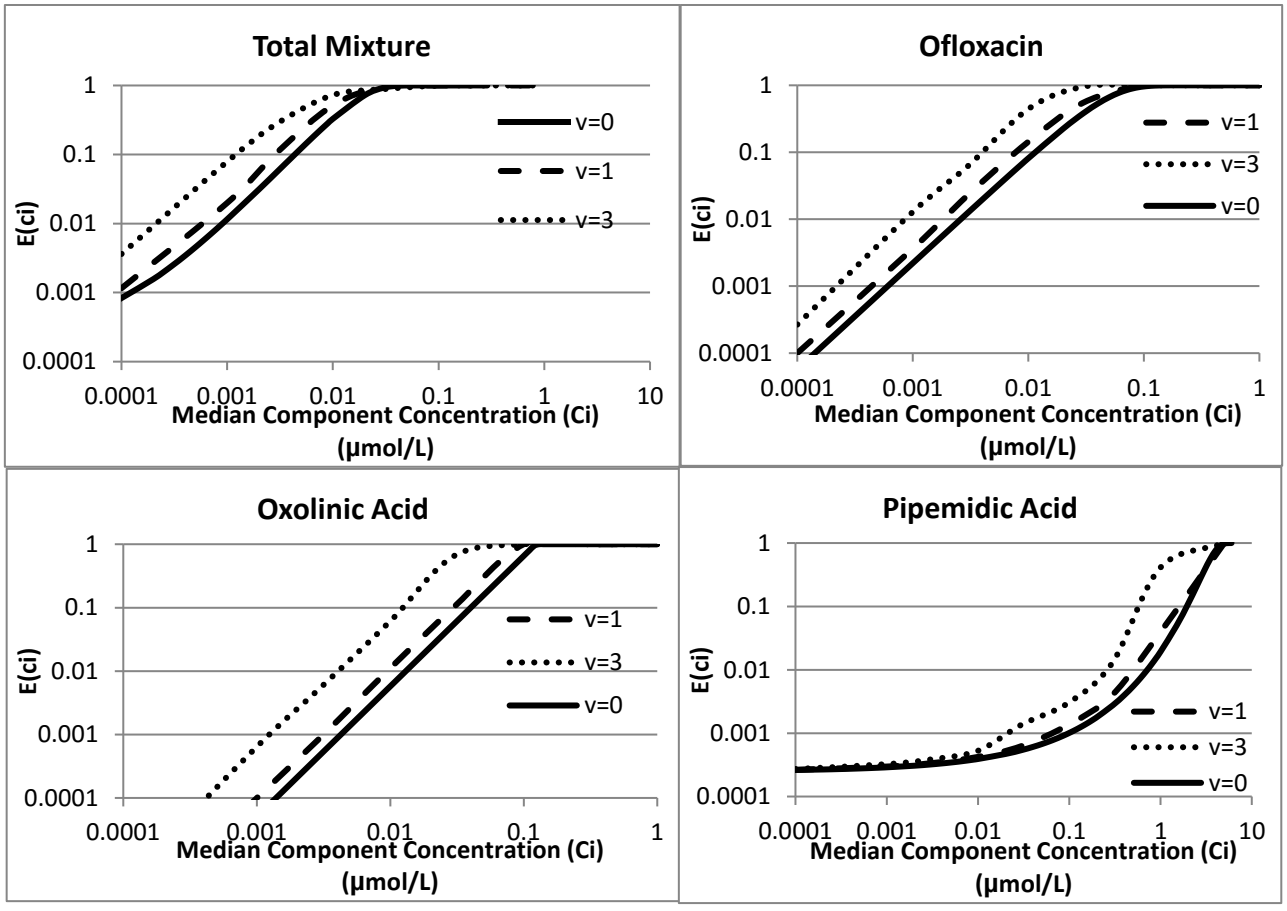
	Ofl.	Lomef.	Norflo.	Flume.	Oxol.	Enox.	Pirom.	Cinox.	Nalid.	Pipem.
Ofloxacin	1.00									
Lomefloxacin	0.39	1.00								
Norfloxacin	0.39	0.40	1.00							
Flumequine	0.40	0.40	0.40	1.00						
Oxolinic Acid	0.39	0.39	0.40	0.40	1.00					
Enoxacin	0.41	0.39	0.38	0.39	0.39	1.00				
Piromidic Acid	0.40	0.40	0.41	0.41	0.39	0.39	1.00			
Cinoxacin	0.41	0.41	0.39	0.40	0.38	0.39	0.39	1.00		
Nalidixic Acid	0.41	0.39	0.38	0.40	0.37	0.39	0.40	0.39	1.00	
Pipemidic Acid	0.40	0.41	0.38	0.38	0.39	0.39	0.39	0.39	0.38	1.00

Assumed correlation matrix, r=+0.9

	Ofl.	Lomef.	Norflo.	Flume.	Oxol.	Enox.	Pirom.	Cinox.	Nalid.	Pipem.
Ofloxacin	1									
Lomefloxacin	0.9	1								
Norfloxacin	0.9	0.9	1							
Flumequine	0.9	0.9	0.9	1						
Oxolinic Acid	0.9	0.9	0.9	0.9	1					
Enoxacin	0.9	0.9	0.9	0.9	0.9	1				
Piromidic Acid	0.9	0.9	0.9	0.9	0.9	0.9	1			
Cinoxacin	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1		
Nalidixic Acid	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1	
Pipemidic Acid	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1

Simulated correlation matrix, r=+0.9 (example case 1)

	Ofl.	Lomef.	Norflo.	Flume.	Oxol.	Enox.	Pirom.	Cinox.	Nalid.	Pipem.
Ofloxacin	1.00									
Lomefloxacin	0.90	1.00								
Norfloxacin	0.90	0.90	1.00							
Flumequine	0.90	0.90	0.90	1.00						
Oxolinic Acid	0.90	0.90	0.90	0.90	1.00					
Enoxacin	0.90	0.90	0.90	0.90	0.90	1.00				
Piromidic Acid	0.90	0.90	0.90	0.90	0.90	0.90	1.00			
Cinoxacin	0.90	0.90	0.90	0.90	0.90	0.90	0.90	1.00		
Nalidixic Acid	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	1.00	
Pipemidic Acid	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	1.00



Effect of varying median concentration and coefficient of variation on individual chemical toxicity and total mixture toxicity, E[Ci]; i=1,10, r=0, Concentration Additivity Theorem

6. REFERENCES

Altenburger, R, Ait-Aissa, S, Antczak, P, *et al.* 2015. Future water quality monitoring--adapting tools to deal with mixtures of pollutants in water resource management. The Science of the total environment. 512-513: 540-51.

Altenburger, R, Nendza, M, and Schuurmann, G. 2009. Mixture Toxicity and Its Modeling by Quantitative Structure-Activity Relationships. Environmental Toxicology and Chemistry. 22.

Altenburger, R, Scholz, S, Schmitt-Jansen, M, *et al.* 2012. Mixture toxicity revisited from a toxicogenomic perspective. Environ Sci Technol. 46: 2508-22.

Ashford, JR. 1981. General Models for the Joint Action of Mixtures of Drugs. Biometrics. 37: 457-74.

Backhaus, T, Scholze, M, and Grimme, LH. 1999. The single substance and mixture toxicity of quinolones to the bioluminescent bacterium Vibrio fischer. Aquatic Toxicology.

Beyer, J, Petersen, K, Song, Y, *et al.* 2014. Environmental risk assessment of combined effects in aquatic ecotoxicology: a discussion paper. Marine environmental research. 96: 81-91.

Bliss, CI. 1939. The Toxicity of Poisons Applied Jointly Annals of Applied Biology. 26: 585-615.

Cassee, FR, Groten, JP, Bladeren, PJv, *et al.* 1998. Toxicological Evaluation and Risk Assessment of Chemical Mixtures. Critical Reviews in Toxicology. 73-101.

Drescherand, K and Boedeker, W. 1995. Assessment of the Combined Effects of Substances: The Relationship between Concentration Addition and Independent Action. Biometrics. 51: 716-30.

Evans, RM, Martin, OV, Faust, M, *et al.* 2015. Should the scope of human mixture risk assessment span legislative/regulatory silos for chemicals? The Science of the total environment.

Gregorio, V, Chevre, N, and Junghans, M. 2013. Critical issues in using the common mixture toxicity models concentration addition or response addition on species sensitivity distributions: a theoretical approach. Environmental toxicology and chemistry / SETAC. 32: 2387-95.

Hadrup, N. 2014. Evidence from pharmacology and pathophysiology suggests that chemicals with dissimilar mechanisms of action could be of bigger concern in the toxicological risk assessment of chemical mixtures than chemicals with a similar mechanism of action. Regulatory Toxicology and Pharmacology. 69: 281-3.

Jin, X, Jin, M, and Sheng, L. 2014. Three dimensional quantitative structure-toxicity relationship modeling and prediction of acute toxicity for organic contaminants to algae. Computers in biology and medicine. 51: 205-13.

Judson, R, Richard, A, Dix, D, *et al.* 2008. ACToR--Aggregated Computational Toxicology Resource. Toxicology and applied pharmacology. 233: 7-13.

Kamo, M and Yokomizo, H. 2015. Explanation of non-additive effects in mixtures of similar mode of action chemicals. Toxicology. 335: 20-6.

Kim, J, Kim, S, and Schaumann, GE. 2012. Reliable predictive computational toxicology methods for mixture toxicity: toward the development of innovative integrated models for environmental risk assessment. Rev Environ Sci Biotechnol.

Konemann, WH and Pieters, MN. 1996. Confusion of Concepts in Mixture Toxicology. Food and Chemical Toxicology. 34: 1025-31.

Liu, Y, Vijver, MG, Qiu, H, *et al.* 2015. Statistically significant deviations from additivity: What do they mean in assessing toxicity of mixtures? Ecotoxicology and environmental safety. 122: 37-44.

Molgaard, B, Hussein, T, Corander, J, *et al.* 2012. Forecasting size-fractionated particle number concentrations in the urban atmosphere. Atmos Environ. 46: 155-63.

Ryker, SJ and Small, MJ. 2008. Combining occurrence and toxicity information to identify priorities for drinking-water mixture research. Risk analysis : an official publication of the Society for Risk Analysis. 28: 653-66.

USEPA. (U.S. Environmental Protection Agency). 1997. Guiding Principles for Monte Carlo Analysis.

USEPA, 2015, the Toxic Substances Control Act (TSCA): U.S. Environmental Protection Agency 2015, http://www2.epa.gov/tsca-inventory/about-tsca-chemical-substance-inventory.

Zwart, DD and Posthuma, L. 2005. Complex Mixture Toxicity fir Single and Multiple Species: Proposed Methodologies. Environmental Toxicology and Chemistry. 24: 2665-76.

CHAPTER 2: NETWORK-BASED FRAMEWORK FOR DOSE-RESPONSE STUDY DESIGN AND INTERPRETATION

Chapter 2, written by Nur H. Orak and co-authored by Mitchell J. Small, and Marek J. Druzdzel, will be submitted to the Journal of Environmental Health Perspectives.

ABSTRACT

Conventional environmental-health risk-assessment methods are limited in their analyses of the actual risks of contaminant exposure. These methods are also incapable of interpreting the different sizes of datasets, which could lead to a better understanding of uncertainties (Wilson 2001a; USNRC 2013a). Therefore, I aim to demonstrate how a performance analysis can be implemented for a Bayesian Network (BN) representation of a dose-response relationship. I explore the effect of different sample sizes on predicting the strength of the relationship between true responses and true doses of environmental toxicants.

The results can guide the use of dose-response studies in regulatory decision-making by determining if data analysis is valid in certain cases, according to the strength of interactions between variables in a network and the sample size. The results will promote the use of model-based dose-response assessment for the dose-selection process and could help to inform the determination of future regulations and guidelines for toxic substances.

1. INTRODUCTION

Dose-response assessment is one of the most critical steps of the environmental risk-assessment process (see Figure 8). It gives information about the adverse health effects of different exposure levels. However, the assessment has uncertainty and variability problems due to experimental error (e.g., an imperfectly controlled environment, human factors, etc.), animal-to-human uncertainties, and other uncertainties (Dong et al. 2015). One primary cause of these uncertainties is that the relation between the actual dose level of a toxicant and the actual response is very difficult to estimate by direct measurements (Brown 1978; Gustafson 2004). Generally, experiments are done with high-dose compound exposure in laboratory animals, and these results are used to predict the potential adverse health endpoint(s) in humans, assuming that similar effects would be expected. However, the levels of chemical exposure in real life are generally much lower than tested levels (Andersen and Krewski 2009; Dong et al. 2015). Decisions about setting maximum contaminant limits are biased by these measured responses. Therefore, this study starts with one of the best known uncertainty problems in experimental studies: How should prior knowledge be used to learn about the strength of the relationship between true exposure and true response? And how do measurement errors in exposure and response affect experimental design for toxicological and epidemiological studies, in particular, the sample sizes needed to determine whether a significant dose (or exposure)-response relationship is present?

We know that a better understanding of prior knowledge can increase the likelihood of successful experimental designs for future trials and for clinical use. In order to achieve this goal, I propose a Bayesian Network (BN) model-based approach to analyze the probabilistic relationship between true exposure and true response. BNs provide a simple yet holistic approach to the use of both

quantitative and qualitative knowledge, with the distinct advantage of combining all available information (Pollino and Henderson 2010).

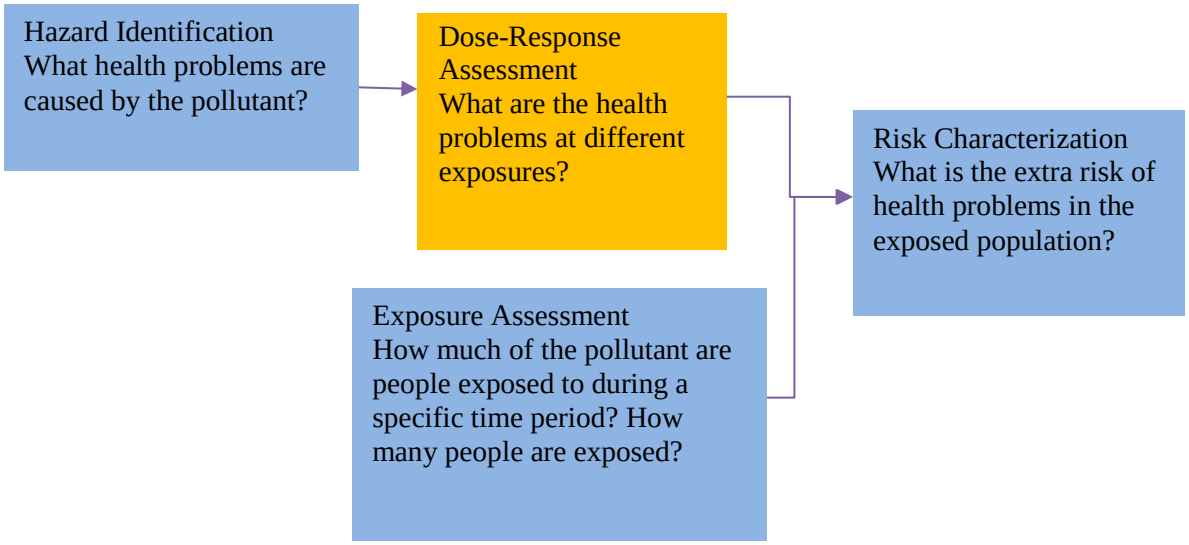


Figure 8. Components of the risk-assessment process (Source: https://epa.gov/)

Measurement error in statistical science is a well-studied topic in the literature (Rhomberg et al. 2011). However, effects of measurement error on the strength of concentration-response relationships in toxicological studies have been limited. BNs can help to understand the effects of measurement errors on the dose-response relationship network. There are three effects of measurement error in covariates: (1) it causes bias in parameter estimation, (2) it leads to a loss of power for the prediction of a relationship, and (3) it makes graphical-model analysis difficult (Carroll et al. 2006). Sonderegger et al. (2009) investigated the effects of unmeasured temporal variation, and they suggest temporal variation in contaminant concentrations causes important bias in the dose-response relationship.

In the next section we discuss our model, giving background on BNs and our estimation of model parameters. In section 3, I present our results. Section 4 discusses the possible applications of our results and adds a brief conclusion.

2. MODEL PROCESS

Using BNs as a risk-assessment tool allows us to investigate and quantify the causal relationships between several interacting variables and outcomes because there is a theoretical relation between causality and probability (Taroni et al. 2006; Mittal and Kassim 2007). Therefore, I aim to predict the strength of relationship between True Exposure (*TE*) and True Response (*TR*) based on different levels of exposure and different sample sizes.

BNs capture cause-and-effect relationships through influence diagrams, so understanding and designing the diagrams is critical. Figure 9 shows the influence diagram of a theoretical dose-relationship assessment. This simplified influence diagram considers several limitations under different nodes. *Accuracy of exposure measurement* includes experimental limitations, human error factors, and the limitations of animal models. *Accuracy of response measurement* includes similar limitations in measurements of responses. *True exposure* and *true response* are the actual exposure and response levels in the real world; regulations and limitations, however, are based on the *measured exposure* and *measured response*.

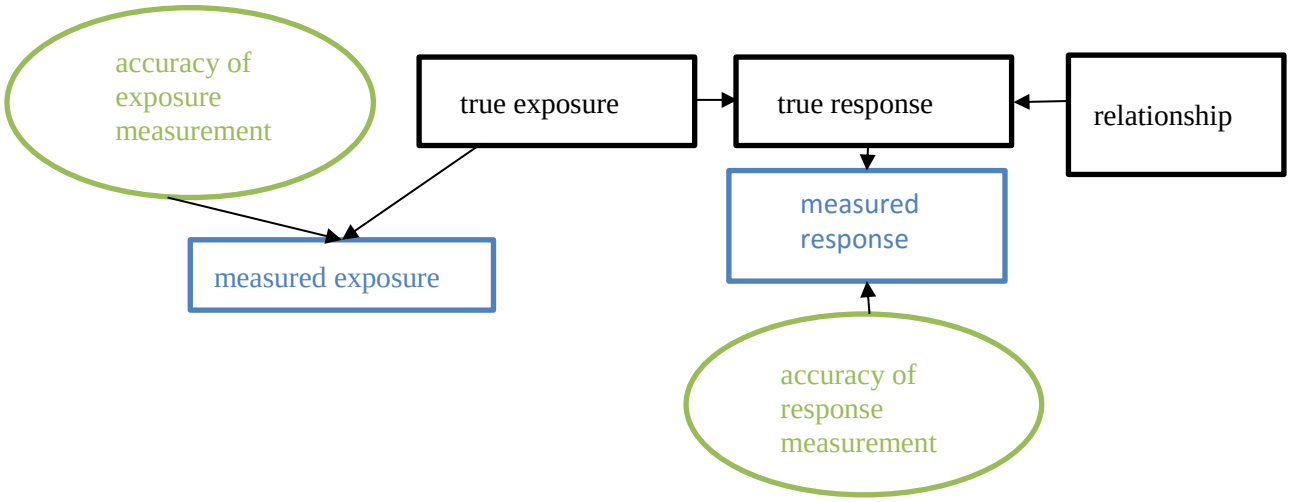


Figure 9. Influence diagram for a dose-response assessment

2.1 Background on Bayesian Networks

Bayesian Networks were developed in the late 1980s to visualize probabilistic dependency models via directed acyclic graphs (DAG) in order to understand the joint probabilistic relationships between variables (Pearl 1988; Newton 2009). BNs are strong modeling tools and are relatively simple compared to other modeling approaches (Pollino and Henderson 2010). The characterization of linkages between variables is typically probabilistic, rather than deterministic, so that BNs allow use of both quantitative and qualitative information (Newton 2009). Bayesian models are particularly appropriate for environmental systems because uncertainty is inherent, and BN's have been used widely for ecological applications (McCann et al. 2006). Similar potential exists in the field of human health risk assessment (Kraisangka et al. 2016).

BNs have been used to analyze problems, and to plan, monitor, and evaluate diverse cases of varying size and complexity in several different disciplines (Weber et al. 2012; Beaudequin et al. 2015; Yang et al. 2016). They also have been used for environmental-related prediction, evaluation, diagnosis, and classification in the literature (Liu et al. 2011; Weber et al. 2012). Specifically, a few studies have investigated the relationship between true exposure and true response through BNs (Marella and Vicard 2013). There are also a few examples of BN applications in health-risk assessment (Woodworth 2004; Mittal and Kassim 2007). A few studies investigated interactions among cancer risk components caused by environmental exposure by using a probability tree approach (Sielken and Valdez-Flores 1999; Small 2008). The authors focus on dose-response predictions as a part of fundamental assumptions of the cancer risk network.

BNs apply Bayes' theorem (also known as Bayes' rule or Bayes' law), which was first derived by Thomas Bayes and published in 1764 (Murphy 2012). According to Bayes' theorem, a

prior probability provides information about the likelihood of a parameter, and the posterior probability is calculated based on the conditional probability of that likelihood (Su et al. 2013). This feature of the theorem differentiates Bayesian statistical models from ordinary non-Bayesian statistical models because the Bayesian approach is a mixture of ordinary linear models and a joint distribution over the measured variables, and it may incorporate subjective prior beliefs (Spirtes et al. 1993). Bayes' rule (Eq. 5) continuously updates the belief probability of each node in the network (Murphy 2012; Tang et al. 2016).

$$p(X = x|Y = y) = \frac{p(X = x, Y = y)}{p(Y = y)} = \frac{p(X = x)p(Y = y|X = x)}{\sum_{x'} p(X = x')p(Y = y|X = x')} \quad (5)$$

BNs bring a holistic approach to understanding the important pathways in networks, which are not easily expressed by mathematical equations, by integrating qualitative expert knowledge, equations, probabilistic modeling, and empirical data (Pearl 1988; Gat-Viks et al. 2006; Tighe et al. 2013). This classification approach is used when the response is categorical, as in our case (Denison et al. 2002).

I developed a BN (Figure 10) based on the preliminary influence diagram (Figure 9) by using the Graphical Network Interface (GeNIe) software package (bayesfusion.com 2016a). I chose this software because of its flexible data-generation feature and its user-friendly interface. The accuracy of exposure-measurement and response-measurement levels are represented by *AcEM* and *AcRM*, respectively. These accuracy levels can be affected by experimental limitations, animal-to-human differences, or human factors. The measured (observed) values of exposure and response are termed *ME* and *MR*, respectively. The true exposure (*TE*) and true response (*TR*) values are the actual exposure and response levels. The node *R* represents the complex relationship between *TE* and *TR*. For instance, if *R* is strong, then the degree of causal influence of *TE* on *TR* is high and the correlation between *TE* and *TR* approaches 1; in other words, an increasing strength

of relationship indicates an increased health risk associated with exposure. The state *none* assumes there is no potential linkage between true exposure and true response. For instance, the response to an environmental exposure can be high even for low measurement accuracy level if R is strong. On the other hand, if R is none, increasing exposure levels do not change the risk of the targeted health effect. The node *ER Match* takes into account the potential nine combinations of *ME* and *MR* outcomes. When the measured exposure and measured risk are the same (i.e., states ll, mm, or hh) this lends support to the belief that a relationship exists between the true exposure and the true risk, especially when the measurement errors are low. When the states do not match, this lends support to the belief that the relationship is not strong, and possibly there is no relationship at all (or the relationship is masked by measurement error).

2.2 Bayesian Network Model Parameters Estimation

I assume that there is no prior information about the distributions of nodes in the network. Therefore, I use the uniform prior probability distribution over each variable, i.e., I assume that each state in a node with three outcomes has a 33% probability of occurrence, except the relationship (R) node. The *R* node prior probability is designed to investigate any potential relationship in addition to the strength of relationship. There is a 50% probability of an existing relationship (medium or strong) or of no relationship (see Figure 10).

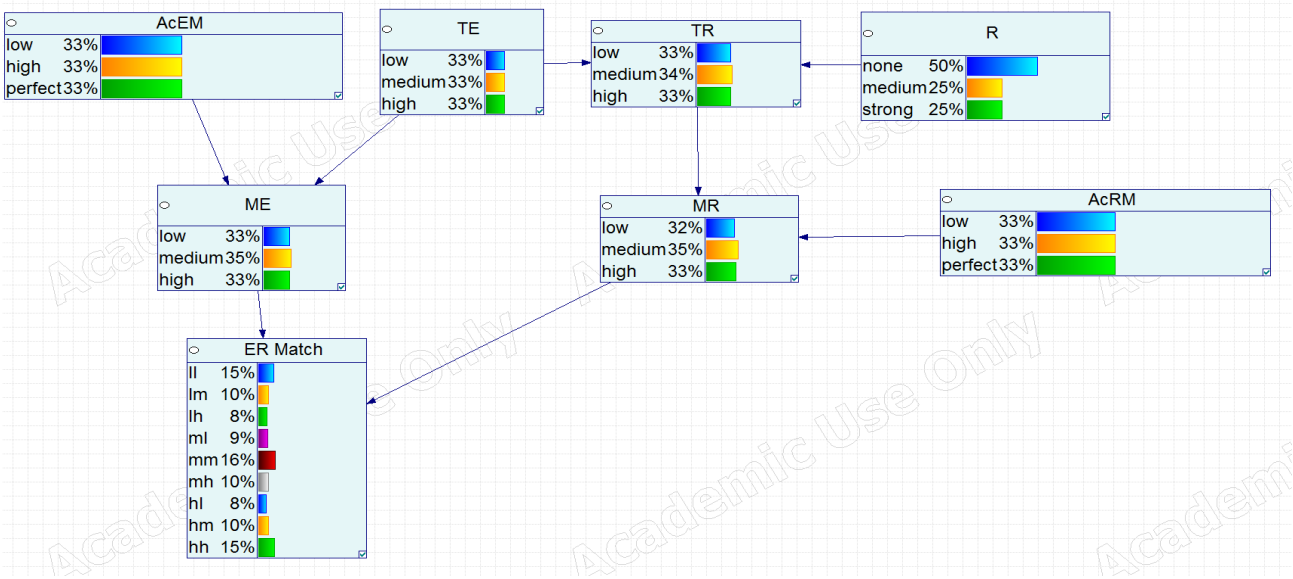


Figure 10. BN model for dose-response assessment with prior probabilities

For the prior network shown in Fig. 10, the conditional probability tables (CPTs) are estimated for *ME*, *MR*, and *TR* to represent the effect of different strength of relationship (none, medium, strong) and different accuracy levels of measurement and response (low, high, perfect) on TE and TR. The CPT of *ME* (Table 5) outlines the probabilities for *ME* based on *AcEM* and *TE*. In case of low accuracy, the probability of accurately predicting *ME* or *MR* is 50%, this value increases to 80% and 100% for high and perfect accuracy levels, respectively.

Table 5. Conditional probability distributions for measured exposure, *ME*

accuracy of ex...	low			high			perfect		
true exposure	low	medium	high	low	medium	high	low	medium	high
low	0.5	0.25	0.2	0.8	0.1	0.05	1	0	0
medium	0.3	0.5	0.3	0.15	0.8	0.15	0	1	0
high	0.2	0.25	0.5	0.05	0.1	0.8	0	0	1

In the case of no relationship (*none*), the probability distribution over all states of the *TR* variable is uniform. Increasing the level of relationship increases the probability of predicting *TR*; 60% for *medium* and 90% for *strong*.

Table 6. Conditional probability distributions for true response, *TR*

relationship	none			medium			strong		
true exposure	low	medium	high	low	medium	high	low	medium	high
low	0.333	0.333	0.333	0.6	0.2	0.15	0.9	0.05	0.03
medium	0.333	0.333	0.333	0.25	0.6	0.25	0.07	0.9	0.07
high	0.334	0.334	0.334	0.15	0.2	0.6	0.03	0.05	0.9

The default belief updating algorithm in GeNIe is the clustering algorithm, the fastest-known exact algorithm for Bayesian networks. The clustering algorithm was originally proposed by Lauritzen and Spiegelhalter (1988) and improved by several researchers (Jensen et al. 1990; Dawid 1992).

2.3 Data Simulation and Analysis

I simulate several random cases for 9 potential combinations of strength of relationships, accuracy level, and sample size summarized in Table 7. GeNIe allows the user to generate random cases that are representative of the network, and it also allows the user to generate these cases with different states selected for some of the network nodes. Each case represents a hypothetical individual in a group of N that was exposed to a potential amount of toxicant in an environment. A "true" population is first simulated with an assumed strength of relationship (none, medium, or strong) and specified levels of exposure and effect measurement error (low, medium or high for each). Given multiple sets of random cases with each (true) specification, I use each of the case sets to update a new "blank" copy of the network (that is, one with the prior specifications) and infer the posterior probability that the strength of relationship (informed by the case set) is none, medium, or strong. If the inferred probabilities align with the true strength of relationship used to generate the cases, then I conclude that the simulated study has the power to properly infer the strength of relationship. This power depends on the accuracy of the measurements and the sample size (N), i.e., the number of random cases in each case set. As N increases, the power for proper

inference likewise increases. In order to demonstrate the comparative results for different sample sizes, I simulated several N values: 20, 50, 100, and 1000.

The simulation analysis involves the following steps:

1- Assign a true state for *R, AcEM,* and *AcRM* (e.g., scenario Figure 11),

2- Generate a dataset D of size N for the selected scenario, and repeat for 10 trials,

3- Count the frequency for each state of *ER Match,* and calculate the average,

4- Calculate the posterior distribution for each state of *R* based on prior knowledge generated by the selected scenarios, and sequential network updates calculated for each case in the dataset D, and

5- Repeat steps 1–4 for different sample sizes (N).

Table 7. Nine scenarios for power evaluation

	Scenario	
Simulation No	**Relationship (R)**	**AcEM - AcRM**
1	None	Low-Low
2	None	High-High
3	None	Perfect-Perfect
4	Medium	Low-Low
5	Medium	High-High
6	Medium	Perfect-Perfect
7	Strong	Low-Low
8	Strong	High-High
9	Strong	Perfect-Perfect

To understand the potential value of new information and to predict a probability distribution for *R*, each case is used as evidence and assigned as the prior for the next case. The Bayes factor (BF) is used as an updating factor to predict the cumulative posterior probability for both the null and the alternative hypothesis. In other words, the BF is a weighted average of the likelihood ratio of the null hypothesis over the alternative hypothesis (Jarosz and Wiley 2014):

$$Bayes\ Factor(BF) = \frac{likelihood\ of\ data\ given\ H_0\ (Posterior\ Odds)}{likelihood\ of\ data\ given\ H_1\ (Prior\ Odds)} \quad (6)$$

An increasing BF indicates increasing evidence in support of the null hypothesis. Posterior probabilities of 10 trials for each scenario are calculated based on BFs. One important advantage of BF is that it is not affected by sample size or other factors, because it is a ratio of probabilities. This feature means that two BFs of equal value provide the same amount of evidence for the alternative hypothesis. Posterior odds of each state of *R* are calculated by multiplying the prior probabilities by the BF, then calculating posterior probabilities based on those outcomes. Then, cumulative posterior probabilities along the sample size are calculated by adding up the posterior probabilities for each case.

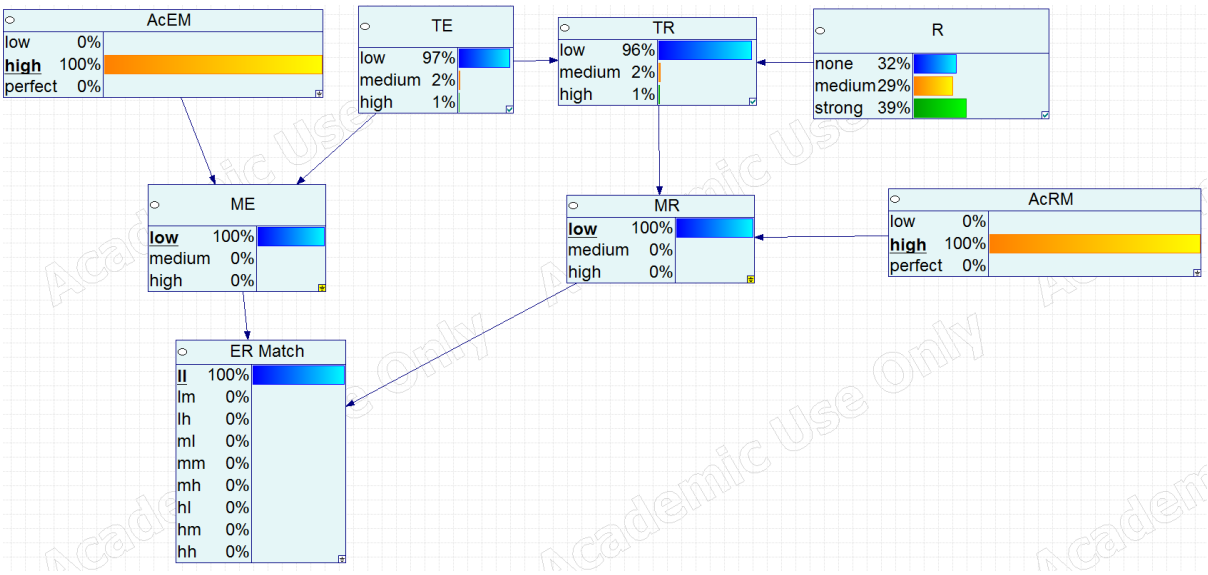


Figure 11. An example: updated BN model for *AcEM-AcRM*: low-low associated relationship (*R*) assessment and a single ll case

Different scenarios are compared with power analysis of the false negative rate and determination of false positive rates (FPRs) as defined below.

$$Average\ FPR = 1 - selectivity = P[medium] + P[strong]$$

(7)

$$Average\ False\ Negative\ Rate\ (FNR)(medium) = 1 - sensitivity = P[none]$$

$$Average\ FNR(strong) = 1 - sensitivity = P[none]$$

$$Average\ power = 1 - FNR$$

(8)

Average power provides important information about the performance of classifiers by plotting the 1-FNR vs. sample size. The accuracy level of prediction increases with increasing power.

3. RESULTS

I evaluate the efficiency of the model by how well it predicts the strength of relationship based on synthetic *ER Match* results. Three figures summarize the outcomes based on different cases generated under three *AcEM* and *AcRM* scenarios (low-low, high-high, perfect-perfect).

In each figure, the title line represents the predicted posterior probabilities of R; each column is for prediction of one class (none, medium or strong). The y-axis indicates the actual R that cases generated. A thick trend line shows the average value of 10 trials. The noise on the graph visualizes the variance as a result of 10 trials.

Figure 12 compares posterior probabilities of predicting *R* for the actual *R* cases (none, medium, and strong) under the low-low *AcEM-AcRM* scenario over 1000 cases. Increasing actual strength of relationship increases the probability of accurately predicting *R* classification. Also, increasing the sample size increases the accuracy level for the predicted *R* for all three scenarios. For example, for the actual case none, posterior P[none] goes up from 50% prior to 75% after N=400.

Figure 13 shows posterior probabilities of predicted *R* for the actual *R* class over 1000 cases. Posterior P(none) from the actual *none* rises dramatically from 50% to 100% after N=150. The previous scenario, *AcEM-AcRM*: low-low, could reach a maximum of 80% average posterior P(none); increasing the accuracy level significantly contributes to the efficiency of *R* prediction. The number of occasional high and low posterior probabilities increase compared to the low-low scenario.

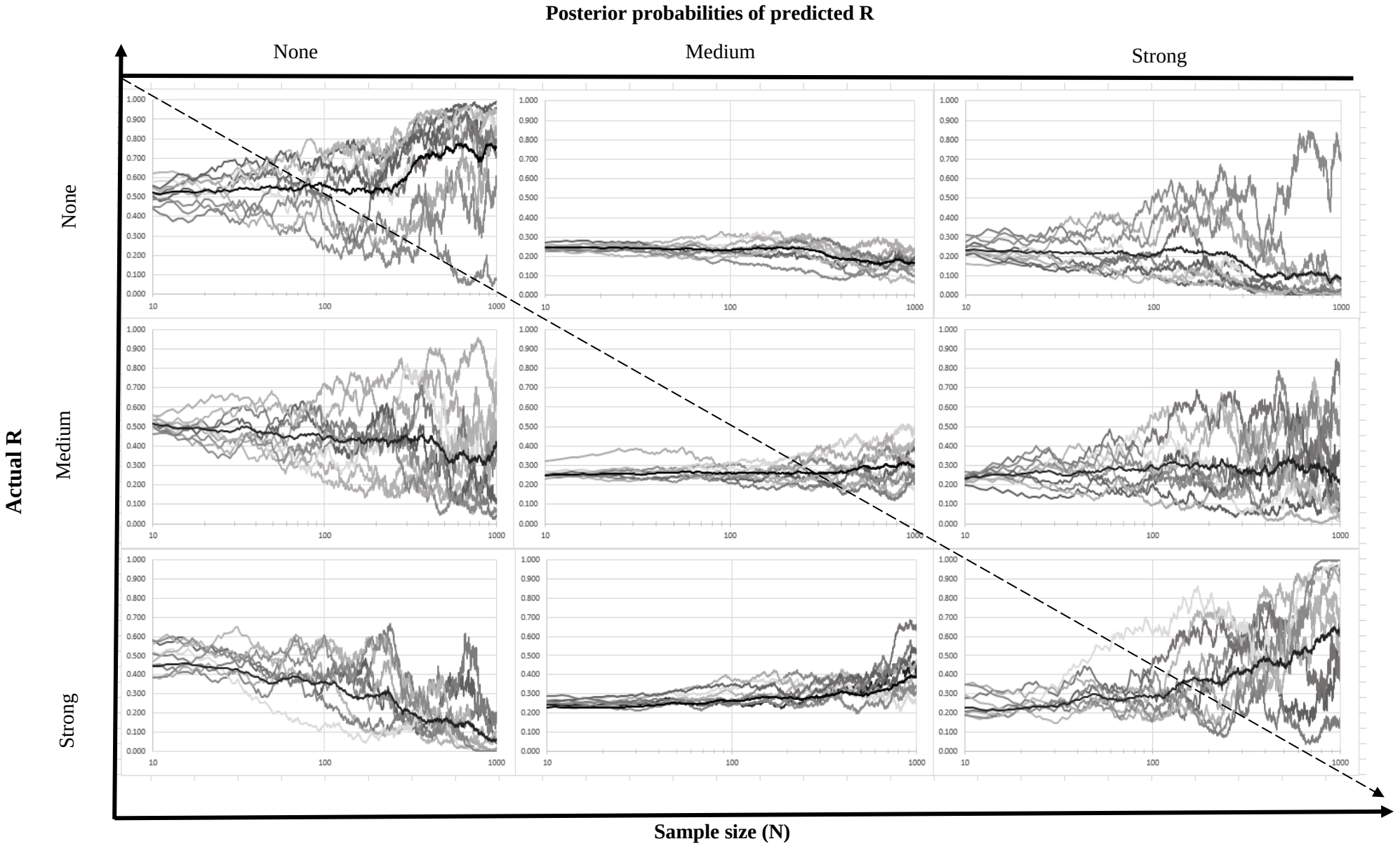


Figure 12. Posterior probabilities of different strength of relationship for the case of low-low accuracy level (title indicates the actual strength of relationship of dataset)

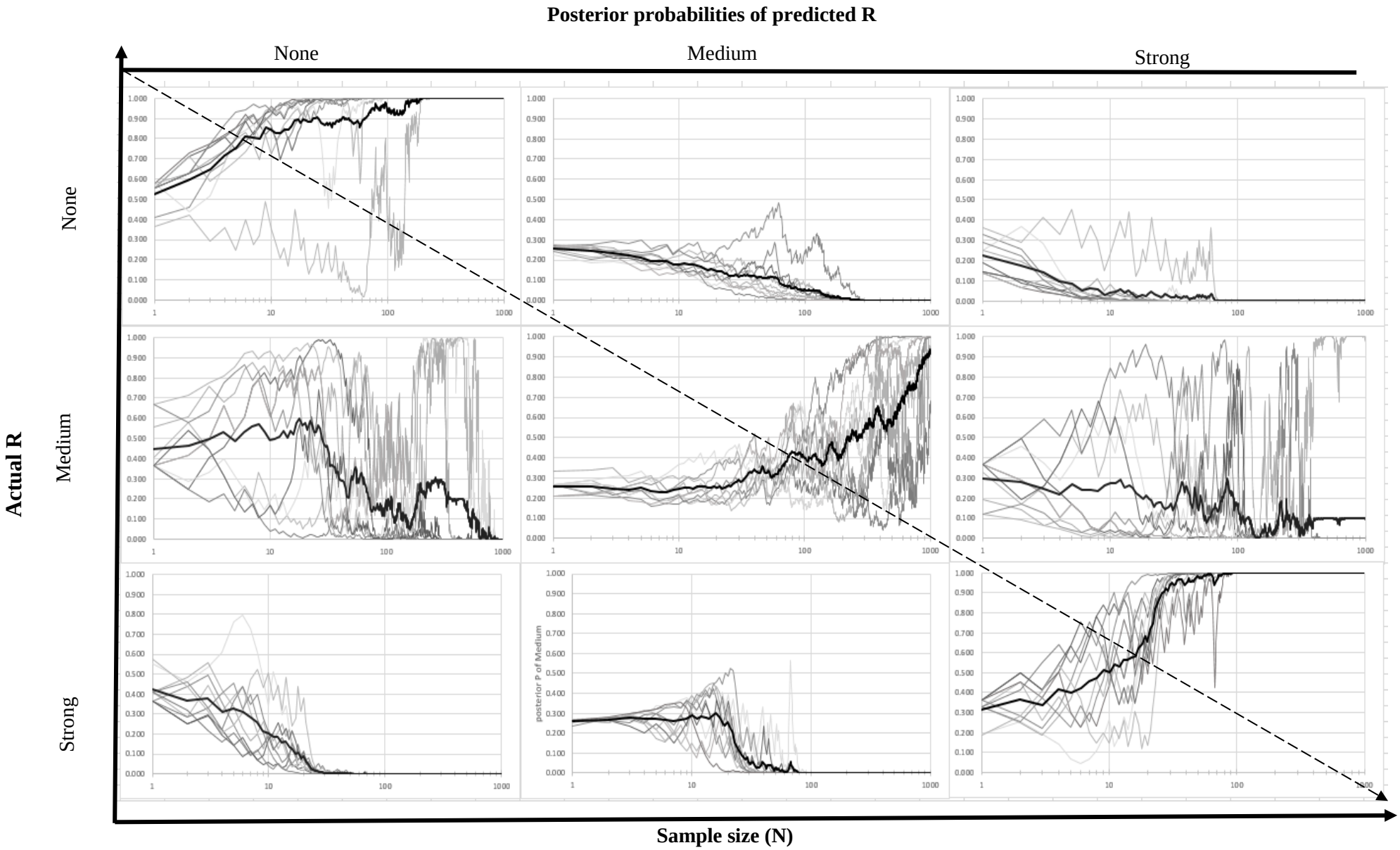


Figure 13. Posterior probabilities of different strength of relationship for the case of high-high accuracy level (title indicates the actual strength of relationship of dataset)

Figure 14 summarizes the outcomes of posterior probabilities of predicted *R* over different sample sizes under perfect-perfect accuracy levels. A perfect accuracy level means that there is a 100% relationship between *TR* and *MR* or *TE* and *ME*. The average trend line converges to 100% more quickly than in other accuracy scenarios. To learn more from different accuracy level scenarios, average power and FPR are calculated. Table 8 lists power for the actual *medium* for four different sample sizes. Prediction power is significantly lower for a very small sample size (N=20), and increasing accuracy levels do not help the effectiveness. For instance, power for the perfect accuracy level increases by only 8%. On the other hand, this gap increases dramatically for big (N=100) and very big sample sizes (N=1000).

Table 8. Average power (medium) vs. sample size (N) for three accuracy levels

Accuracy Level\Sample size	**20**	**50**	**100**	**1000**
Low	0.50	0.51	0.53	0.61
High	0.47	0.56	0.67	0.87
Perfect	0.58	0.79	0.89	0.98

Table 5 summarizes average power for predicting *strong* cases at different accuracy levels. Power is higher even for small sample sizes when compared to Table 9. Big sample size is enough to reach almost 100% prediction accuracy with high experimental accuracy.

Table 9. Average power (strong) vs. sample size (N) for three accuracy levels

Accuracy Level\Sample size	**20**	**50**	**100**	**1000**
Low	0.57	0.60	0.62	0.82
High	0.97	0.99	0.99	1.00
Perfect	1.00	1.00	1.00	1.00

Table 10 indicates FPRs vary with accuracy level and sample size. The average false positive ratio decreases with increasing accuracy levels and sample sizes.

Table 10. Average FPR vs. sample size (N) for three accuracy levels

Accuracy Level\Sample size	20	50	100	1000
Low	0.49	0.47	0.46	0.33
High	0.25	0.19	0.13	0.02
Perfect	0.19	0.11	0.06	0.01

Table 11 shows the sample size needed to (on average) infer with 90% posterior probability the correct strength (for the three true strengths of relationship) and the three accuracy levels. Increasing accuracy levels requires smaller sample sizes to predict the strength of true relationship. For instance, increasing accuracy level causes a dramatic decrease of sample size (1000+ to 6) for the case of *strong* relationship.

Table 11. The sample size needed to infer with 90% posterior probability of the correct strength

	True strengths of relationship		
Accuracy Level	**None**	**Medium**	**Strong**
Low	1000+	1000+	1000+
High	133	983	25
Perfect	32	205	6

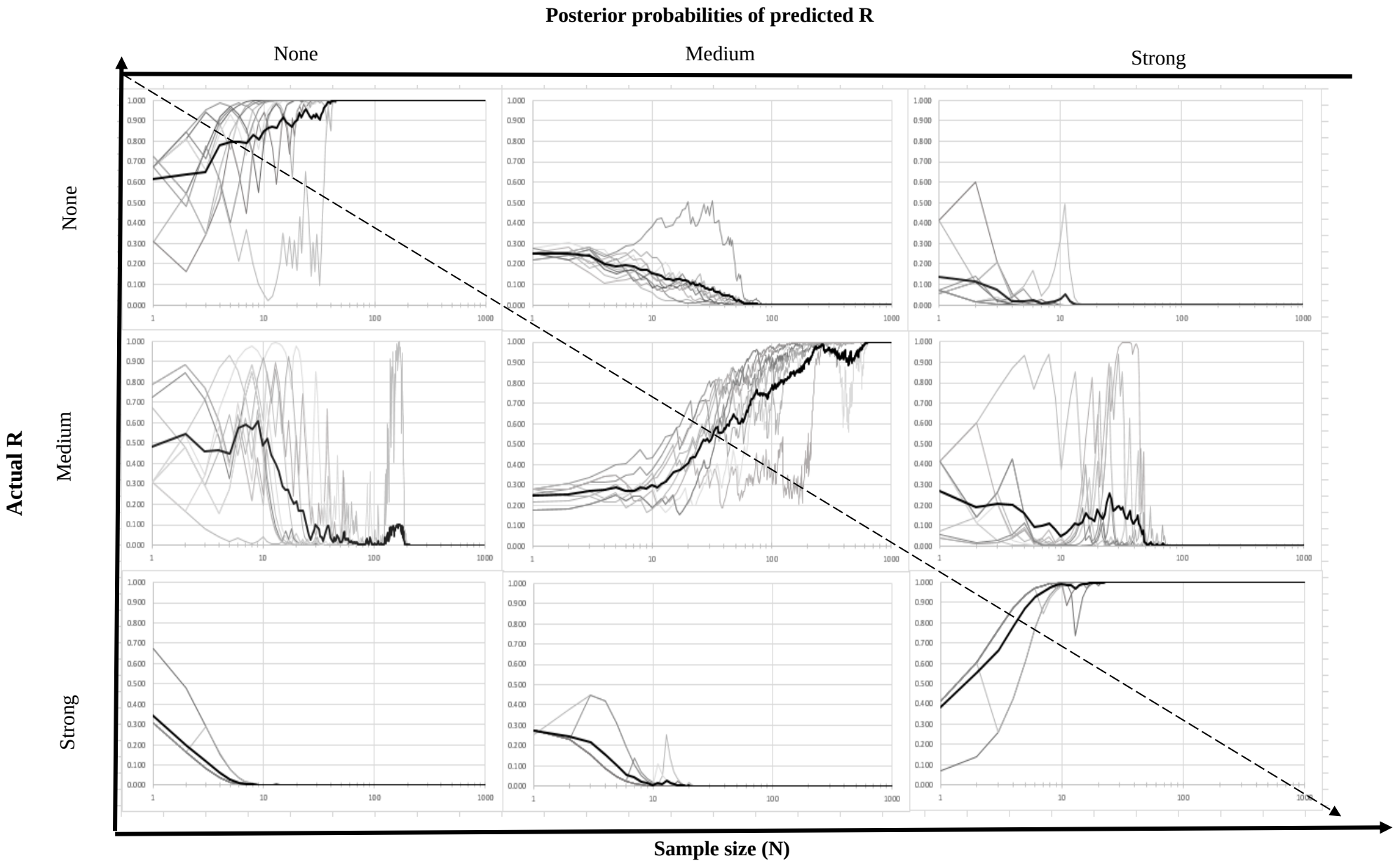


Figure 14. Figure 7. Posterior probabilities of different strength of relationship for the case of perfect-perfect accuracy level (title indicates the actual strength of relationship of dataset)

4. DISCUSSION AND CONCLUSION

Current environmental-health risk-assessment approaches are not effective in understanding actual dose-response relationships and the effects of measurement errors in different sample sizes. Directed graphs can provide a powerful approach for visualizing dependencies between variables in a network. In this study, we present a novel method to answer fundamental uncertainty questions in toxicological/epidemiological studies. I use BN as a tool to understand hidden biases due to unobserved covariates. Dose-response relationships are investigated among individual cases and within individuals.

Our findings show that increasing actual strength of relationship increases the accuracy level of predicting *relationship (R)* classification. Also, increasing the sample size increases the accuracy level for the predicted *R* for all scenarios. Moreover, increasing the experimental accuracy level significantly contributes to the efficiency of *R* prediction. These results can be applied to various contexts in toxicological and epidemiological studies.

5. APPENDIX B

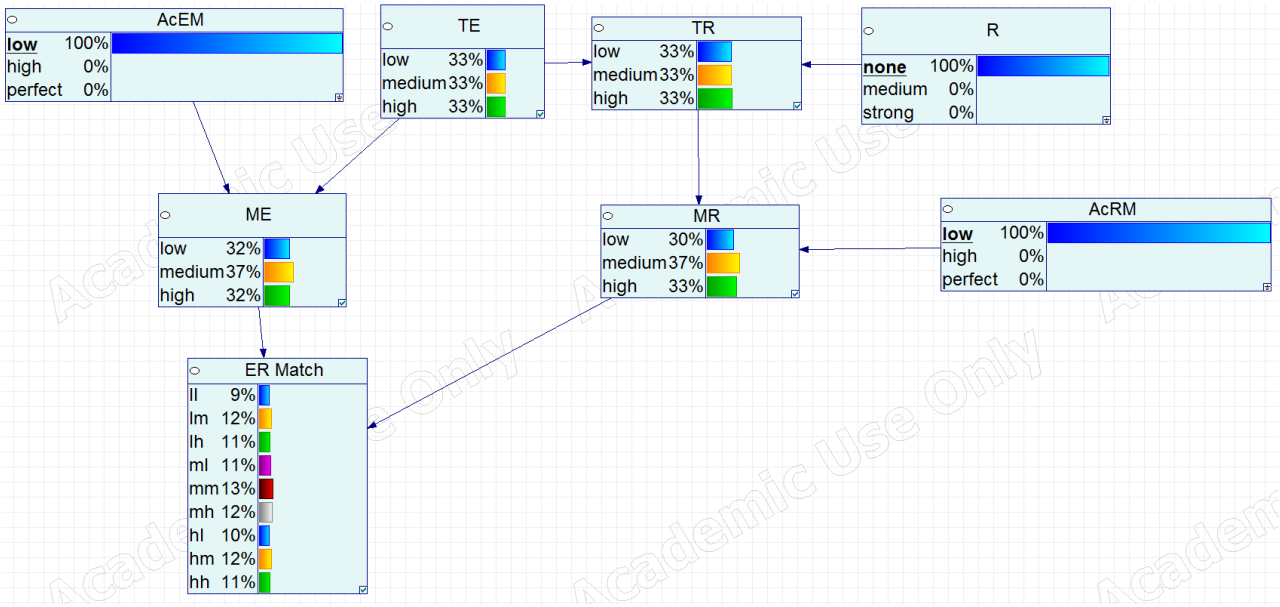


Figure 15. Data Simulation-Scenario AcEM-AcRM: low-low, R: none

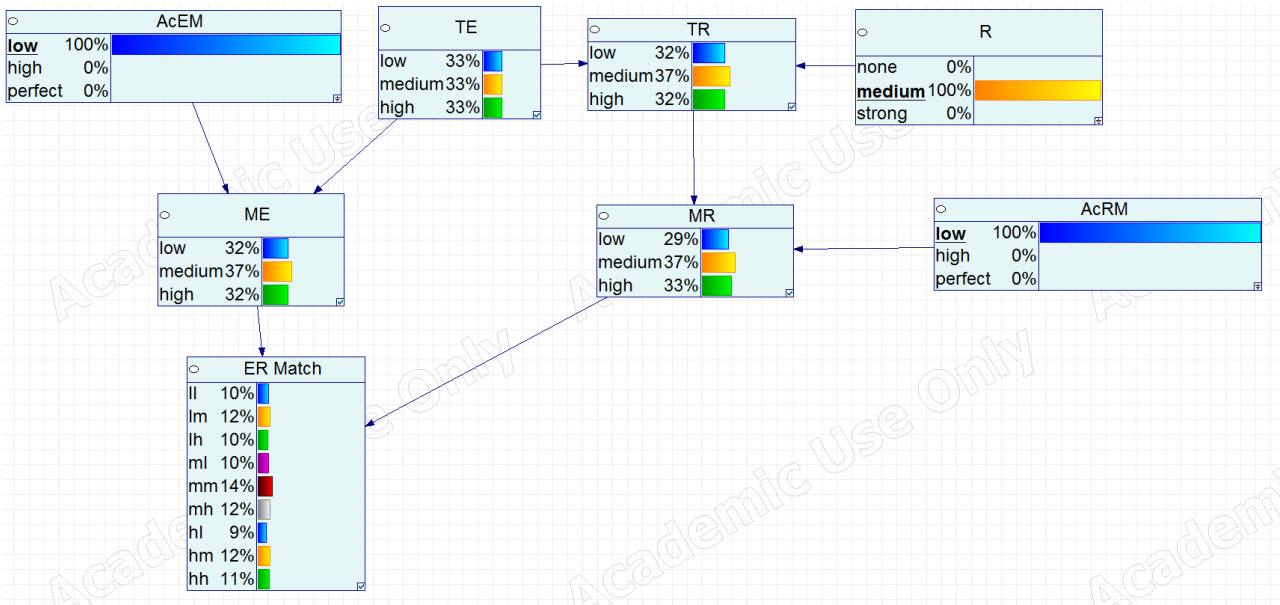


Figure 16. Data Simulation-Scenario AcEM-AcRM: low-low, R: medium

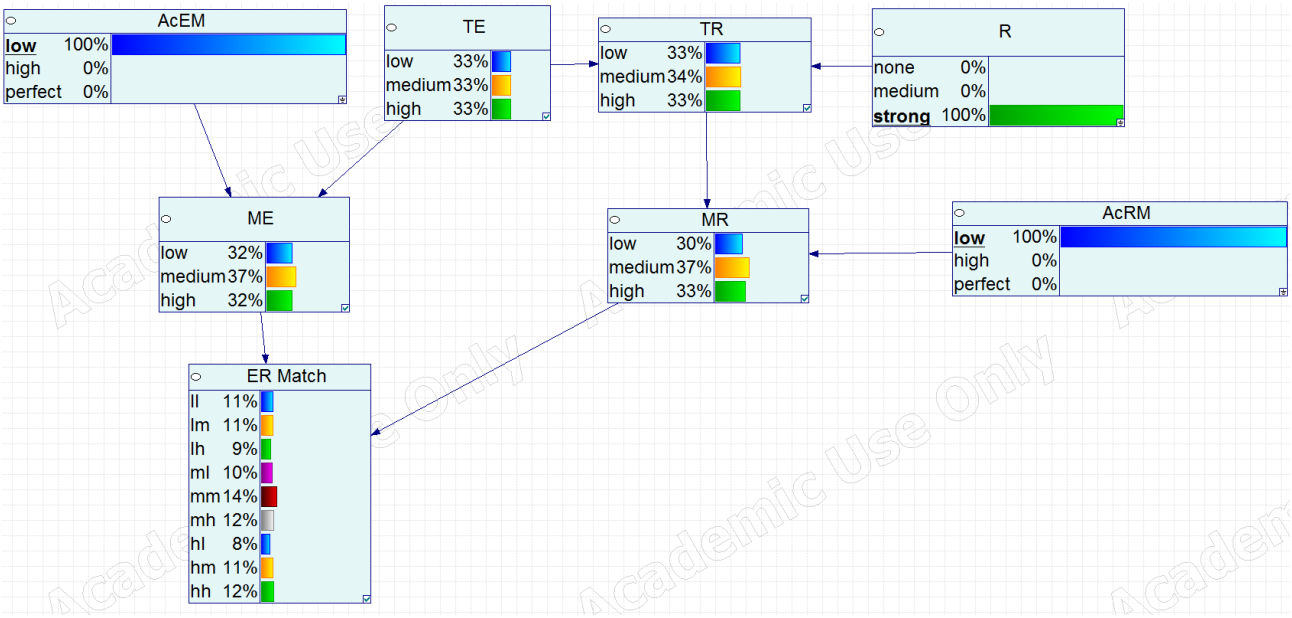


Figure 17. Data Simulation-Scenario AcEM-AcRM: low-low, R: strong

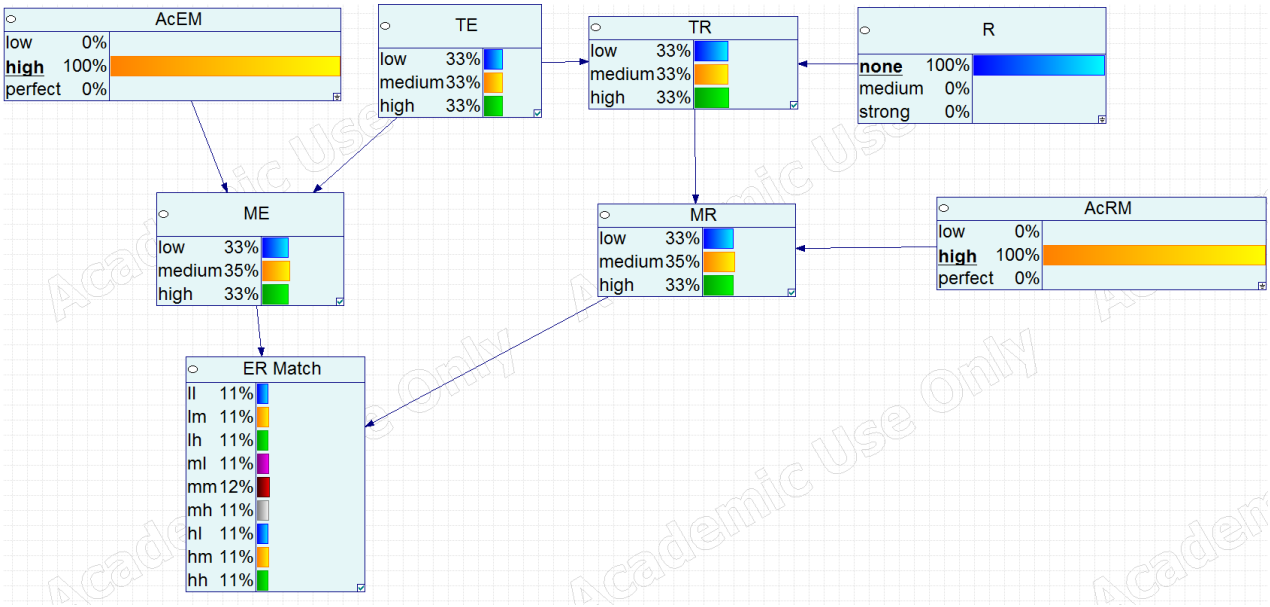


Figure 18. Data Simulation-Scenario AcEM-AcRM: high-high, R: none

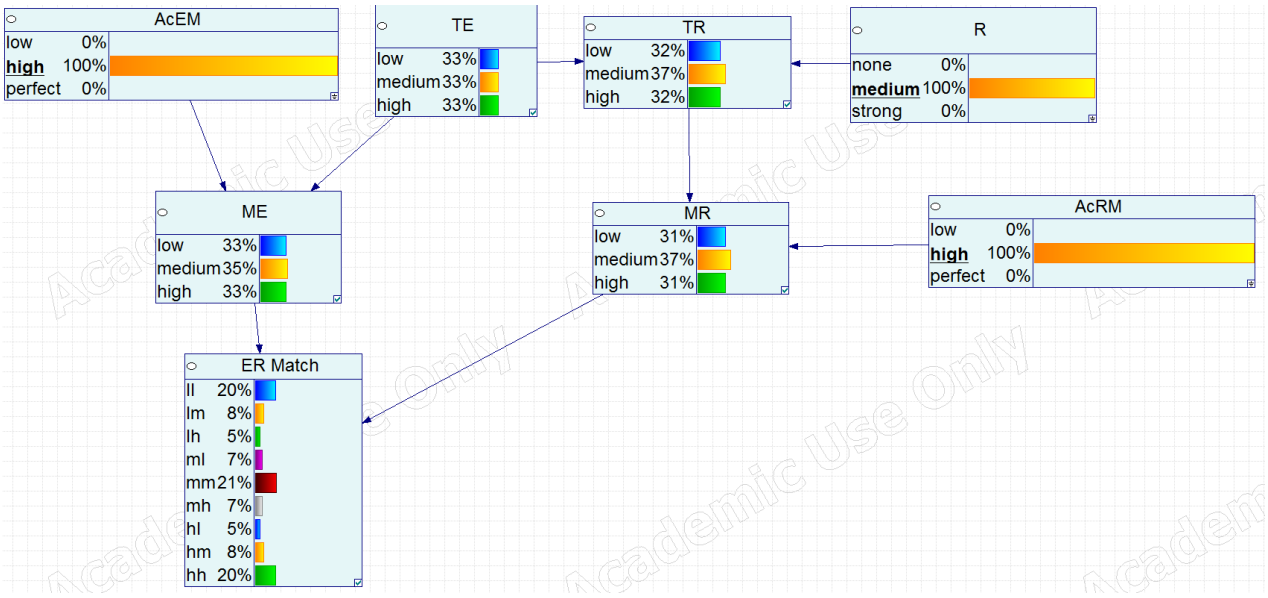


Figure 19. Data Simulation-Scenario AcEM-AcRM: high-high, R: medium

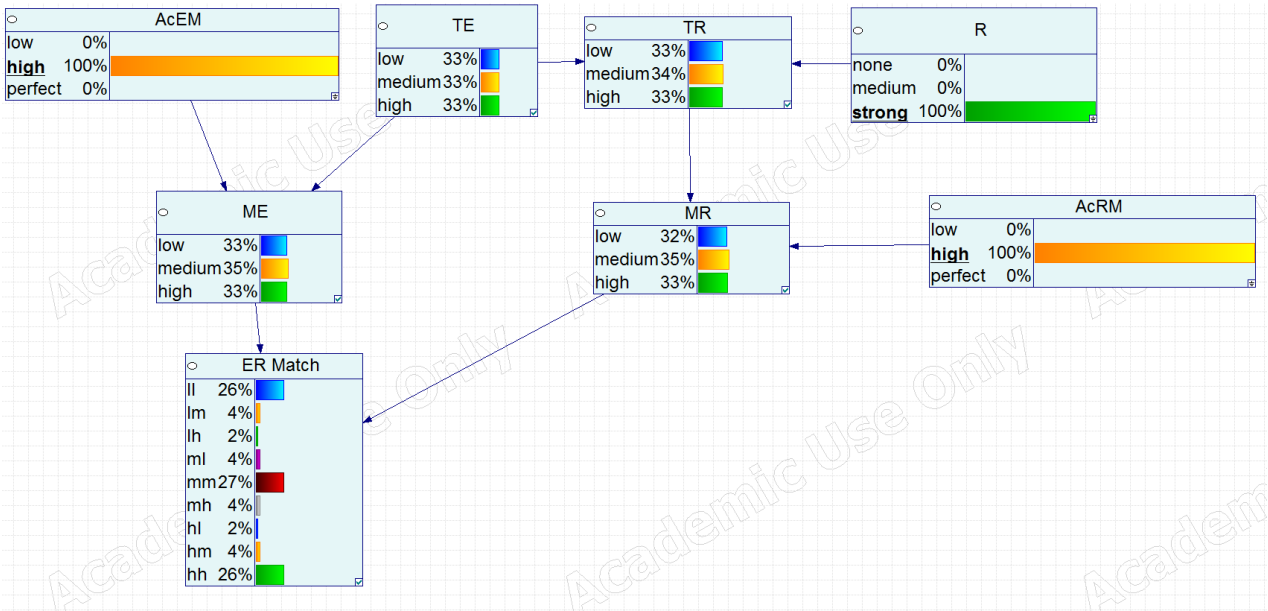


Figure 20. Data Simulation-Scenario AcEM-AcRM: high-high, R: strong

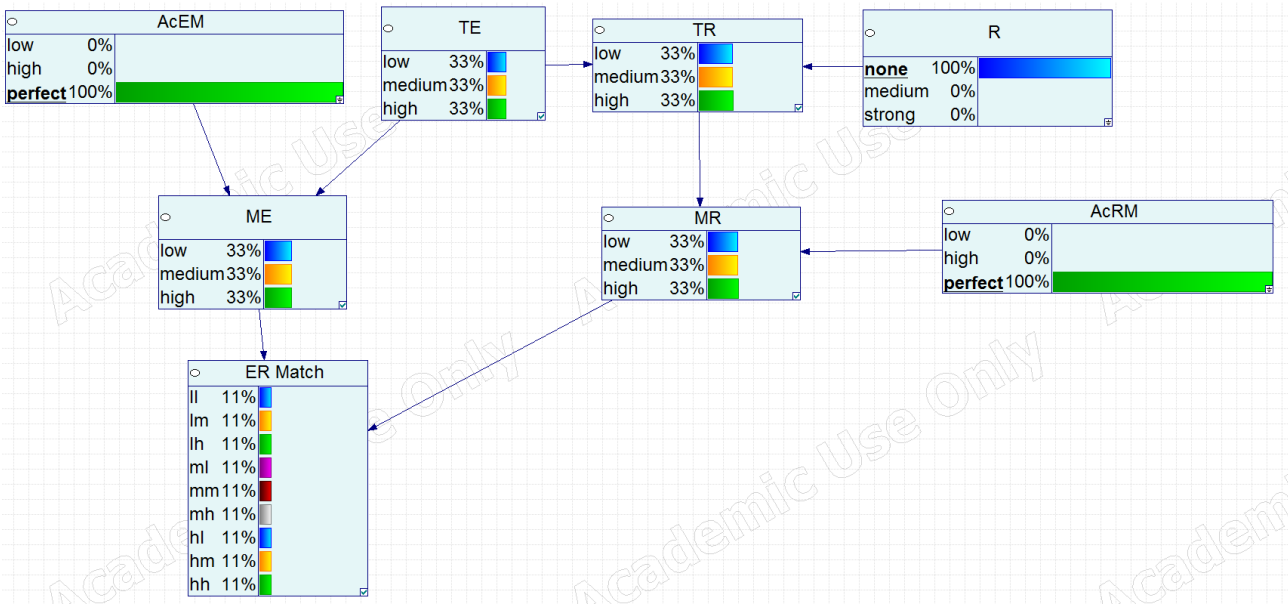


Figure 21. Data Simulation-Scenario AcEM-AcRM: perfect-perfect, R: none

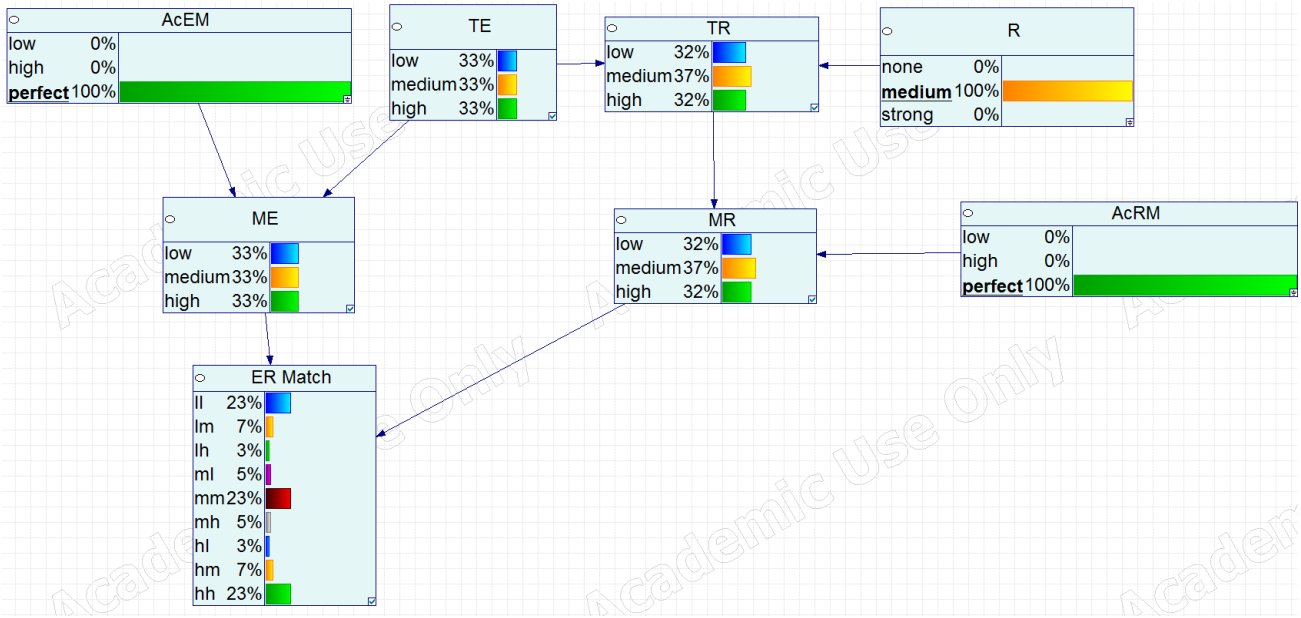


Figure 22. Data Simulation-Scenario AcEM-AcRM: perfect-perfect, R: medium

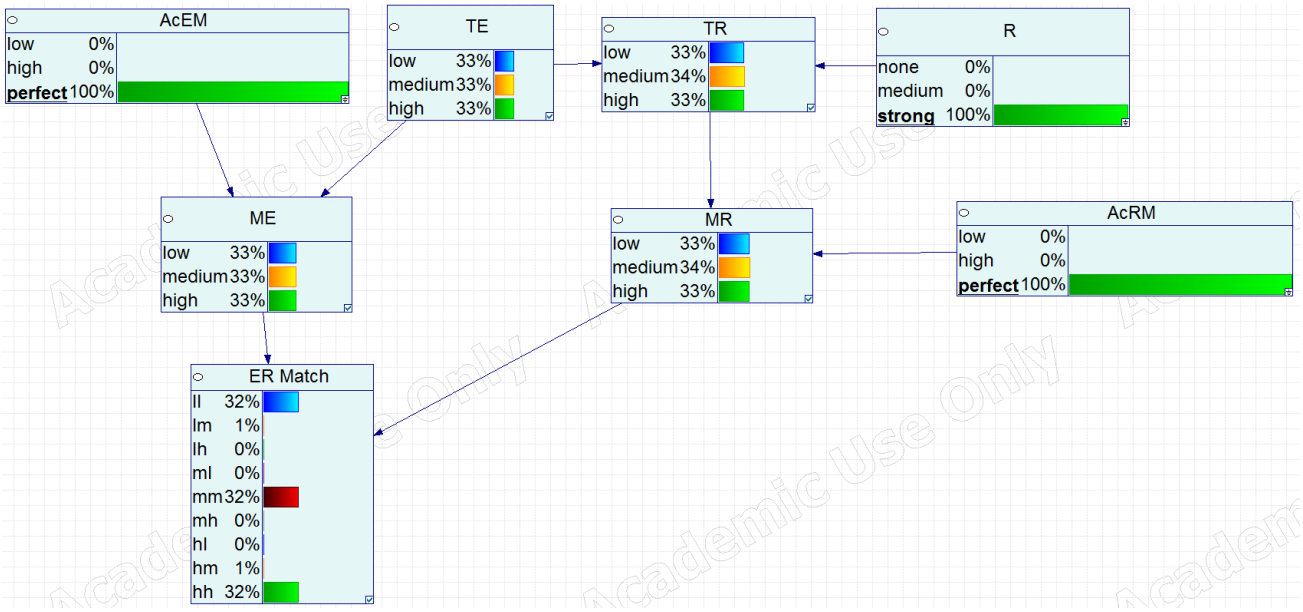


Figure 23. Data Simulation-Scenario AcEM-AcRM: perfect-perfect, R: strong

Table 12. Bayes Factor-AcEM-AcRM: low-low

	For None			For Medium			For Strong		
Data	**PostP**	**PostOdds**	**BF None**	**PostP**	**PostOdds**	**BF**	**PostP**	**PostOdds**	**BF**
ll	0.47	0.90	0.90	0.25	0.34	1.01	0.27	0.38	1.13
lm	0.50	1.01	1.01	0.25	0.34	1.01	0.25	0.32	0.97
lh	0.52	1.10	1.10	0.24	0.32	0.97	0.23	0.30	0.90
ml	0.51	1.05	1.05	0.25	0.33	0.98	0.24	0.32	0.96
mm	0.49	0.95	0.95	0.25	0.34	1.02	0.26	0.35	1.05
mh	0.50	1.01	1.01	0.25	0.33	0.99	0.25	0.33	0.99
hl	0.52	1.11	1.11	0.24	0.32	0.96	0.23	0.30	0.91
hm	0.50	1.01	1.01	0.25	0.34	1.01	0.25	0.32	0.97
hh	0.48	0.91	0.91	0.25	0.34	1.02	0.27	0.37	1.11

Table 13. Bayes Factor-AcEM-AcRM: high-high

	For None			For Medium			For Strong		
Data	**PostP**	**PostOdds**	**BF**	**PostP**	**PostOdds**	**BF**	**PostP**	**PostOdds**	**BF**
ll	0.37	0.58	0.58	0.27	0.37	1.10	0.37	0.58	1.73
lm	0.55	1.24	1.24	0.25	0.34	1.02	0.19	0.24	0.71
lh	0.67	2.01	2.01	0.21	0.26	0.79	0.12	0.14	0.42
ml	0.58	1.36	1.36	0.24	0.31	0.93	0.19	0.23	0.69
mm	0.41	0.69	0.69	0.27	0.36	1.09	0.33	0.49	1.46
mh	0.58	1.36	1.36	0.24	0.31	0.93	0.19	0.23	0.69
hl	0.67	2.01	2.01	0.21	0.27	0.80	0.12	0.14	0.42
hm	0.55	1.25	1.25	0.25	0.34	1.02	0.19	0.24	0.71
hh	0.37	0.58	0.58	0.27	0.37	1.10	0.36	0.57	1.72

Table 14. Bayes Factor-AcEM-AcRM: perfect-perfect

	For None			For Medium			For Strong		
Data	**PostP**	**PostOdds**	**BF**	**PostP**	**PostOdds**	**BF**	**PostP**	**PostOdds**	**BF**
ll	0.31	0.44	0.44	0.28	0.38	1.15	0.42	0.71	2.13
lm	0.68	2.08	2.08	0.25	0.34	1.02	0.07	0.08	0.23
lh	0.79	3.71	3.71	0.18	0.21	0.64	0.04	0.04	0.11
ml	0.73	2.66	2.66	0.22	0.28	0.84	0.05	0.06	0.17
mm	0.31	0.44	0.44	0.28	0.38	1.15	0.42	0.71	2.13
mh	0.73	2.67	2.67	0.22	0.28	0.84	0.05	0.06	0.17
hl	0.79	3.70	3.70	0.18	0.22	0.65	0.04	0.04	0.11
hm	0.68	2.08	2.08	0.25	0.34	1.02	0.07	0.08	0.23
hh	0.31	0.45	0.45	0.28	0.38	1.15	0.42	0.71	2.13

6. REFERENCES

Andersen, ME and Krewski, D. 2009. Toxicity testing in the 21st century: bringing the vision to life. Toxicological sciences : an official journal of the Society of Toxicology. 107: 324-30.

bayesfusion.com, 2016a, GeNIe 2016.

Beaudequin, D, Harden, F, Roiko, A, *et al.* 2015. Beyond QMRA: Modelling microbial health risk as a complex system using Bayesian networks. Environ Int. 80: 8-18.

Brown, CC. 1978. The Statistical Analysis of Dose-Effect Relationships. Princeiples of Ecotoxicology.

Carroll, RJ, Ruppert, D, Stefanski, LA, *et al.* 2006. Measurement Error in Nonlinear Models: A Modern Perspective. Chapman & Hall/CRC. pp

Dawid, PA. 1992. Applications of a general propagation algorithm for probabilistic expert systems. Statistics and Computing. 2: 25-36.

Denison, DGT, Holmes, CC, Mallick, BK, *et al.* 2002. Bayesian Methods for Nonlinear Classification and Regression. John Wiley & Sons, Ltd., England. pp

Dong, Z, Liu, Y, Duan, L, *et al.* 2015. Uncertainties in human health risk assessment of environmental contaminants: A review and perspective. Environ Int. 85: 120-32.

Gat-Viks, I, Tanay, A, Raijman, A, *et al.* 2006. A Probabilistic Methodology for Integrating Knowledge and Experiments on Biological Networks. Journal of Computational Biology. 13: 165-81.

Gustafson, P. 2004. Measurement Error and Misclassification in Statistics and Epidemiology. Chapman&Hall/CRC. pp

Jarosz, AF and Wiley, J. 2014. What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. The Journal of Problem Solving. 7.

Jensen, FV, Lauritzen, SL, and Olsen, KG. 1990. Bayesian updating in recursive graphical models by local computations. Computational Staisticals Quarterly. 4: 269-82.

Kraisangka, J, Druzdzel, MJ, and Benza, R. 2016. A Risk Calculator for the Pulmonary Arterial Hypertension Based on a Bayesian Network. 13th Annual Bayesian Modeling Applications Workshop (BMAW-2016), New York City, NY.

Liu, KF-R, Lu, C-F, Chen, C-W, *et al.* 2011. Applying Bayesian belief networks to health risk assessment. Stochastic Environmental Research and Risk Assessment. 26: 451-65.

Marella, D and Vicard, P. 2013. Object-Oriented Bayesian Networks for Modeling the Respondent Measurement Error. Communications in Statistics - Theory and Methods. 42: 3463-77.

McCann, RK, Marcot, BG, and Ellis, R. 2006. Bayesian belief networks: applications in ecology and natural resource management. Canadian Journal of Forest Research. 36: 3053-62.

Mittal, A and Kassim, A. 2007. Bayesian Network Technologies: Applications and Graphical Models. IGI Publishing. pp

Murphy, KP. 2012. Machine Learning A Probabilistic Perspective. The MIT Press, Cambridge, Massachusetts, London, England. pp

Newton, AC. 2009. Bayesian Belief Networks in Environmental Modeling: A Review of Recent Progress. Environmental Modelling 13-50.

Pearl, J. 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, Los Angeles. pp

Pollino, CA and Henderson, C. 2010. Bayesian networks: A guide for their application in natural resource management and policy. Australian Government.

Rhomberg, LR, Chandaliaa, JK, Longa, CM, *et al.* 2011. Measurement error in environmental epidemiology and the shape of exposure-response curves. Critical reviews in toxicology. 41.

Sielken, RL and Valdez-Flores, C. 1999. Probabilistic Risk Assessments Use of Trees and Distributions to Reflect Uncertainty and Variability and To Overcome The Limitations of Default Assumptions. Environment international. 25: 755-72.

Small, MJ. 2008. Methods for assessing uncertainty in fundamental assumptions and associated models for cancer risk assessment. Risk analysis : an official publication of the Society for Risk Analysis. 28: 1289-308.

Sonderegger, DL, Wang, H, Huang, Y, *et al.* 2009. Effects of measurement error on the strength of concentration-response relationships in aquatic toxicology. Ecotoxicology. 18: 824–8.

Spirtes, P, Glymour, C, and Scheines, R. 1993. Causation, Prediction, and Search. Carnegie Mellon University, Department of Philosophy. pp

Su, C, Andrew, A, Karagas, MR, *et al.* 2013. Using Bayesian networks to discover relations between genes, environment, and disease. BioData Mining. 6.

Tang, C, Yi, Y, Yang, Z, *et al.* 2016. Risk analysis of emergent water pollution accidents based on a Bayesian Network. Journal of environmental management. 165: 199-205.

Taroni, F, Aitken, C, Garbolino, P, *et al.* 2006. Bayesian Networks and Probabilistic Inference in Forensic Science. John Wiley & Sons, Ltd. pp

Tighe, M, Pollino, CA, and Wilson, SC. 2013. Bayesian Networks as a screening tool for exposure assessment. Journal of environmental management. 123: 68-76.

USNRC. 2013a. Critical Aspects of EPA's IRIS Assessment of Inorganic Arsenic: Interim Report. National Research Council, Washington D.C.

Weber, P, Medina-Oliva, G, Simon, C, *et al.* 2012. Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. Engineering Applications of Artificial Intelligence. 25: 671-82.

Wilson, JD. 2001a. Advanced Methods for Dose–Response Assessment: Bayesian Approaches—Final Report. Resources for the Future.

Woodworth, GG. 2004. Biostatistics: A Bayesian Introduction. John Wiley & Sons, Ltd. , New Jersey. pp

Yang, C, Ji, J, Liu, J, *et al.* 2016. Structural learning of Bayesian networks by bacterial foraging optimization. International Journal of Approximate Reasoning. 69: 147-67.

CHAPTER 3: A PROBABILISTIC METHODOLOGY FOR RISK ASSESSMENT OF ARSENIC EXPOSURE AND ADVERSE REPRODUCTIVE OUTCOMES

Chapter 3, written by Nur H. Orak and co-authored by Mitchell J. Small, will be submitted to the Journal of Environmental Science and Technology.

ABSTRACT

Arsenic contamination of drinking water affects more than 137 million people and has been linked to several adverse health effects including cognitive, cardiovascular, and metabolic disorders (Caldwell et al. 2015). Prior research has shown that arsenic poisoning can be attributed to arsenic concentration, human metabolism and arsenic species (arsenicals). However, conventional health-risk assessment methods are not very effective in analyzing the actual risk of arsenic exposure. These methods are also not capable of interpreting large amounts of data, the interpretation of which could lead to a better understanding of uncertainties. (Wilson 2001b; USNRC 2013b). The complex, integrated systems, where physical-biological-human systems interact, require a multidisciplinary approach. Therefore, we aim to explore the prenatal, inorganic arsenic-exposure network, the strength of interactions, and the potential causal relationships by combining Hypothesis Based Weight of Evidence (HBWoE) and Bayesian Network (BN) modeling. Our analysis demonstrates how BN can be used to quantify HBWoE evaluation.

First, I give some background information about the HBWoE framework, and inorganic arsenic (iAs) exposure, metabolism of iAs. Section 2 explains the model process steps, and sections 3-4 outline the results and potential implications. Because I provide detailed information about BNs in Chapter 2, I do not give detailed information in this chapter.

1. INTRODUCTION

The traditional toxicological approach, “dose-response” graphs, is limited in its ability to unveil the relationship between potential risk factors of arsenic exposure and adverse human health outcomes, which is critically important to understanding the risk at low arsenic exposure levels. The U.S. National Research Council (USNRC 2013b) published a report that recommends data-driven approaches over default practices for assessing multiple effects of inorganic arsenic. Therefore, to provide insight on the potential interactions of different variables in the arsenic-exposure network, this study characterizes the risk factors by combining an HBWoE framework with BN modeling as a tool for health risk assessment. The goal of this chapter is to design a BN to incorporate different types of information on arsenic exposure in drinking water and its effects on pregnant women.

1.1 Background on the Hypothesis-Based Weight of Evidence (HBWoE) Methodology

It is challenging in a systematic literature review to collect and evaluate the weight of evidence in each study. Hypothesis-based weight of evidence (HBWoE) provides an objective, operational, and transparent weight-of-evidence concept (Rhomberg 2013; Bailey et al. 2016). The main goal is integrating all the sources of evidence to support causality between variables. The secondary goal is providing procedures for evaluating evidence for given outcomes, Figure 24 shows the procedure of the HBWoE. We complete our litrature review and data collection by following the steps of Figure 14.

.

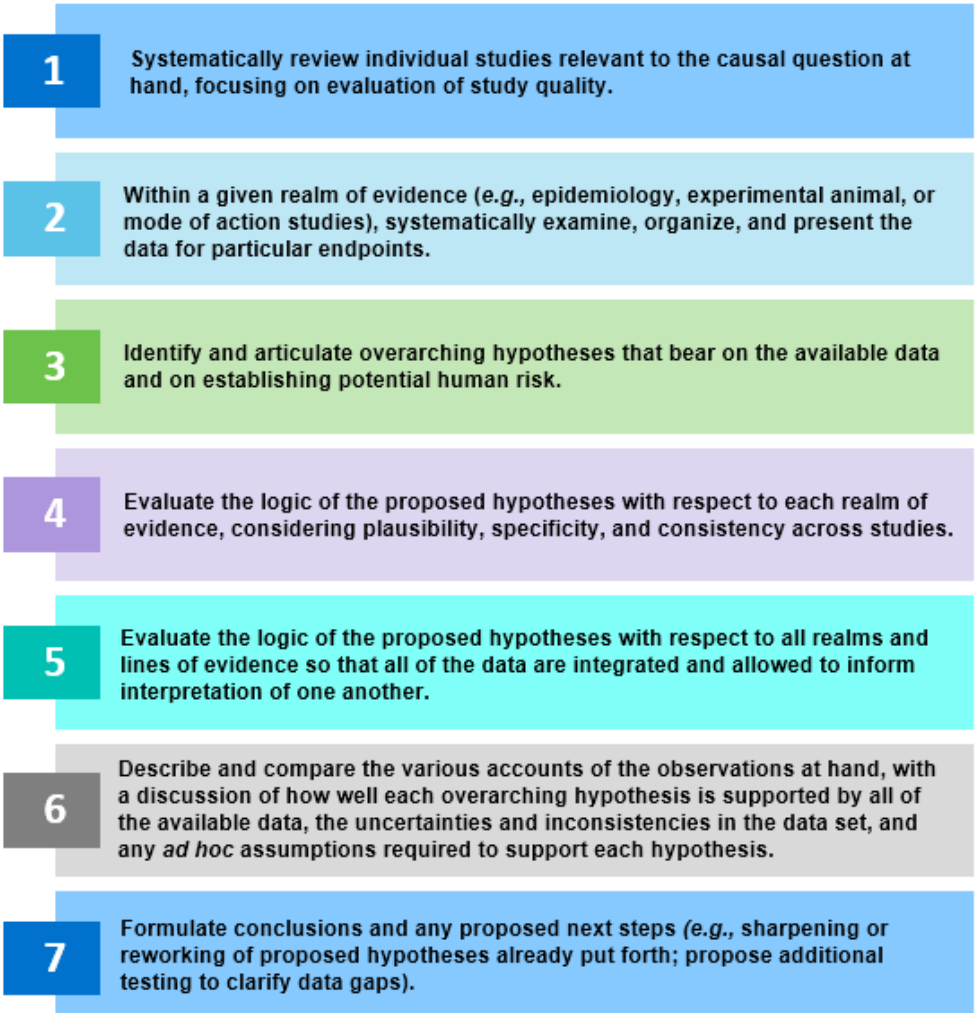


Figure 24. Seven steps of the Hypothesis-Based Weight-of-Evidence (HBWoE) approach (Rhomberg and Bailey 2015)

1.2 Literature Review on Inorganic Arsenic (iAs) Exposure and Metabolism of iAs

Natural water contamination by arsenic exposure is a global threat to public health (Stanton et al. 2015). Arsenic is the 20th most abundant element in the earth's crust, the 14th in seawater and the 12th in the human body. In other words, arsenic also occurs naturally (Mandal and Suzuki 2002). The most common forms of inorganic arsenic, arsenate (As^{5+}) and arsenite (As^{3+}), are more toxic than organic arsenic (Qi et al. 2014). High levels of arsenic exposure cause adverse effects on human health, such as dermal, respiratory, pulmonary, cardiovascular, gastrointestinal, hematological, hepatic, renal, neurological, immunologic, genotoxic, mutagenetic, and carcinogenic effects (WHO 2011; USNRC 2013b). In addition, prenatal

inorganic arsenic exposure is potentially linked with infant development and survival (Gardner et al. 2011; Rager et al. 2014). Therefore, for drinking water, the maximum permissible arsenic concentration is 10 µg/L by the U.S. Environmental Protection Agency (EPA), and the recommended value is 10 µg/L by the World Health Organization (WHO) (Stanton et al. 2015).

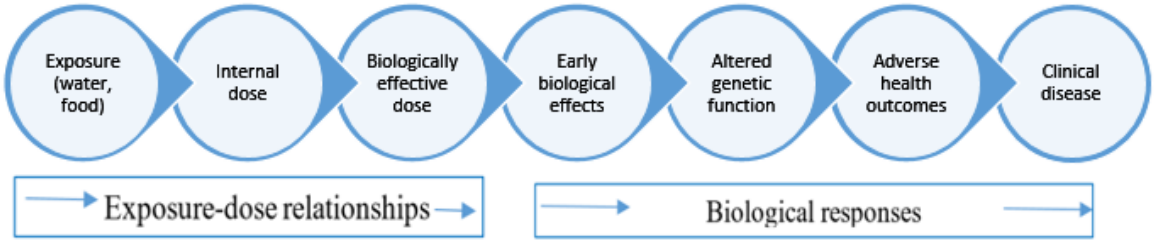


Figure 25. Pathway of arsenic metabolism

The literature documents the relationship between arsenic exposure in drinking water and adverse pregnancy outcomes. High concentrations of arsenic can cross the placental barrier and cause negative reproductive and developmental effects, such as spontaneous abortion, preterm birth, stillbirth, and decreased birth weight (Ahmad et al. 2001; Bailey and Fry 2015; Punshon et al. 2015). The adverse health outcomes are caused by several factors (variables), such as arsenic concentration in drinking water and characteristics of women, such as genetic components (sex, age, and body weight) and lifestyle (smoking habits, alcohol consumption, etc.) (NRC 2001; Laine et al. 2015). In addition, genetic factors can also affect the metabolism of arsenic to form different arsenic metabolites (arsenicals). Yet even though the interactions between these variables and the underlying reason of arsenic partitioning are extensively studied in the literature, the current dose-response methods do not explain these relationships well.

The metabolism of inorganic arsenic has been well studied in the literature. The metabolism of inorganic arsenic is complex, and the process forms several different arsenicals, that have different toxicities (USNRC 2013b). Prior research on arsenic detoxification has shown that an effective method is methylation (Thomas et al. 2001; Vahter 2002; Wanibuchi et al. 2004).

Inorganic arsenic compounds can be methylated to monomethylarsonic acid (MMA), dimethylarsinic acid (DMA) and trimethylarsine oxide (TMAsO) (Mandal and Suzuki 2002; Qi et al. 2014). These highly methylated species have fewer toxic effects compared to less methylated compounds (Wanibuchi et al. 2004; Laine et al. 2015). Several studies show that the partitioning of inorganic arsenic metabolites in the human body is 20-30% inorganic arsenic, 10-20% MMA, and 60-80% DMA (Gardner et al. 2011) (see Figure 13). Women are

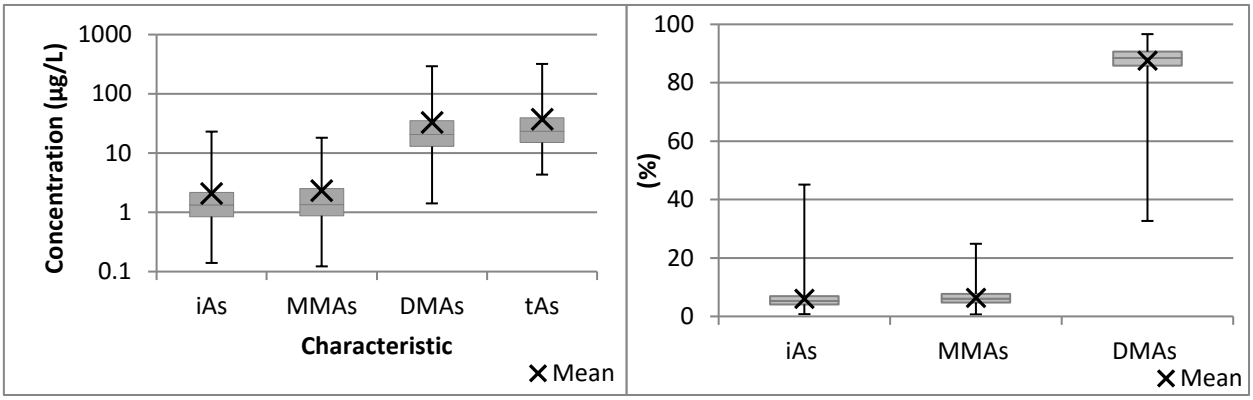


Figure 26. Levels of urinary arsenicals of the participants of the BEAR study, adapted from Laine et al. (2015)

more efficient in converting inorganic arsenic to DMA than men; this efficiency may be higher during pregnancy (Vahter et al. 2006). There are several biomarkers of arsenic exposure in the human body; the most common ones are blood arsenic and urinary arsenic (Jarup 2003).

2. MODEL PROCESS

The model development starts with setting goals for the prenatal arsenic exposure BN model (A-BN) (see Figure 27), which explore the causal relationships in the inorganic arsenic (iAs) exposure network and the effects of drinking water iAs exposure (DW-iAs) on low birth weight. For the second and third steps, I use knowledge from HBWoE to design the conceptual model and, identify and parameterize the model variables. The last step aims to evaluate the model.

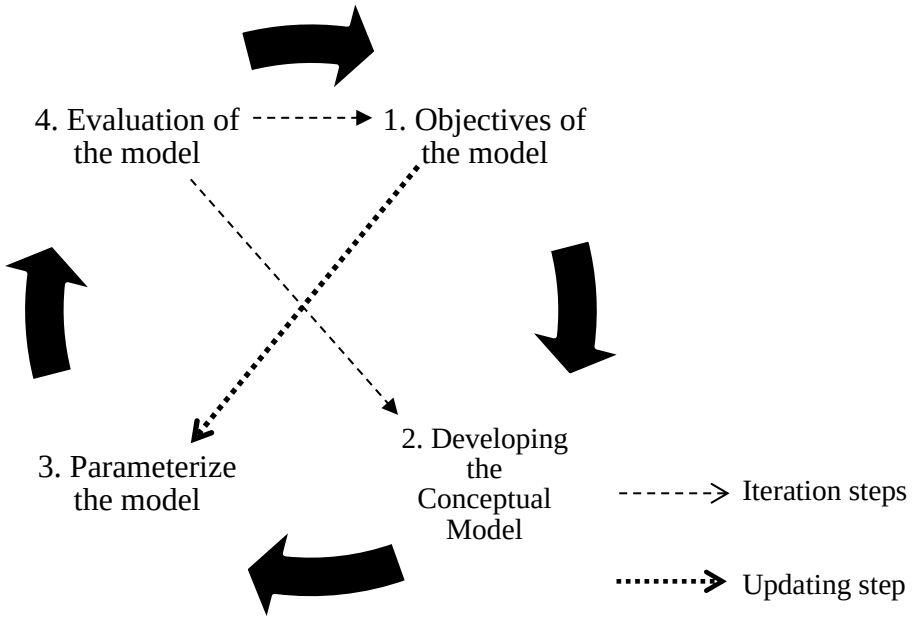


Figure 27. Steps to develop the Arsenic-Bayesian Network (A-BN)

2.1 Objectives

The first aim of the A-BN model is to explore the relative importance of each variable in the network. Using BN as a risk-assessment tool allows us to investigate and quantify the causal relationships between several interacting variables and outcomes because there is a

theoretical relation between causality and probability. Therefore, I aim to improve the understanding and prioritization of risk factors in the arsenic network by using BN.

2.2 Developing the Conceptual Model and Parameterization

We develop a conceptual model after defining the objectives, which has two important steps: first, identifying the network variables and, second, developing an influence diagram (model structure) using those variables. BNs capture cause-and-effect relationships through influence diagrams. Therefore, understanding and designing the influence diagram is critical. There are several different approaches; I develop an influence diagram based on evidence in the literature by following the first five steps of HBWoE. I collect data on DW-iAs, arsenicals in the human body, and demographic characteristics from different studies. I develop several BNs with GeNIe Software (bayesfusion.com 2016b) based on the preliminary influence diagram, and I predict conditional probability tables (CPTs) based on the outcomes of HBWoE. I compare different BNs to reach the most accurate prenatal arsenic-exposure network model.

This step requires several iterations to understand the significance of each variable and simplify the influence diagram. I organize the outcomes for the selected final four literature studies in Table 16. This table is the main information source for simplifying the arsenic network and finding the best discretization for each variable. In other words, I develop an evidence chart to prioritize the variables and understand the direction of a potential relation between two variables (y causes x to be more likely to occur). If this relationship is clearly known, then I keep it in the diagram.

When the influence diagram is ready, I update the nodes with BN algorithms, which are categorized in GeNIe as Exact algorithms or Stochastic algorithms. Exact algorithms include the clustering algorithm and the polytree algorithm. Stochastic sampling algorithms

have 7 types: Probabilistic Logic, Likelihood, Self-Importance, Heuristic Importance, Backward, AIS, and EPIS Samplings. The estimated posterior importance sampling (EPIS) algorithm is almost always the best sampling algorithm available, so I will focus on EPIS. This algorithm computes the posterior probability over all nodes with a loopy belief propagation, and later it uses *Importance Sampling* to refine the estimate (Druzdzel 2003).

I select a hypothesis for the first demonstration: does arsenic exposure affect low birth weight? I restructure the influence diagram to answer this question by comparing four literature studies.

2.3 Evaluation and Test

After I complete the development of A-BN, I need to see how much the results differ for the different conditional probability tables (CPTs) that are estimated based on Table 16. I calculate the false positive rate (FPR) and the power (TPR=1-FNR) of the evidence provided by the model.

Table 15. Table of Error Types

		Test Result	
		Negative	**Positive**
True State of World	**Negative**	True Negative Rate (selectivity)	False Positive Rate
	Positive	False Negative Rate	True Positive Rate (sensitivity)

3. RESULTS

First, I develop a preliminary A-BBN to identify all the important variables of the arsenic network (Figure 28): DW-iAs (µg/L), total arsenic (tAs) (µg/L), DMA (p_DMA) (%), MMA (p_MMA) (%), iAs (p_iAs) (%), gestational age (ges_age) (weeks), baby weight (baby_weight) (g), smoking status (smoker), alcohol consumption (alcohol), and mother age (age).

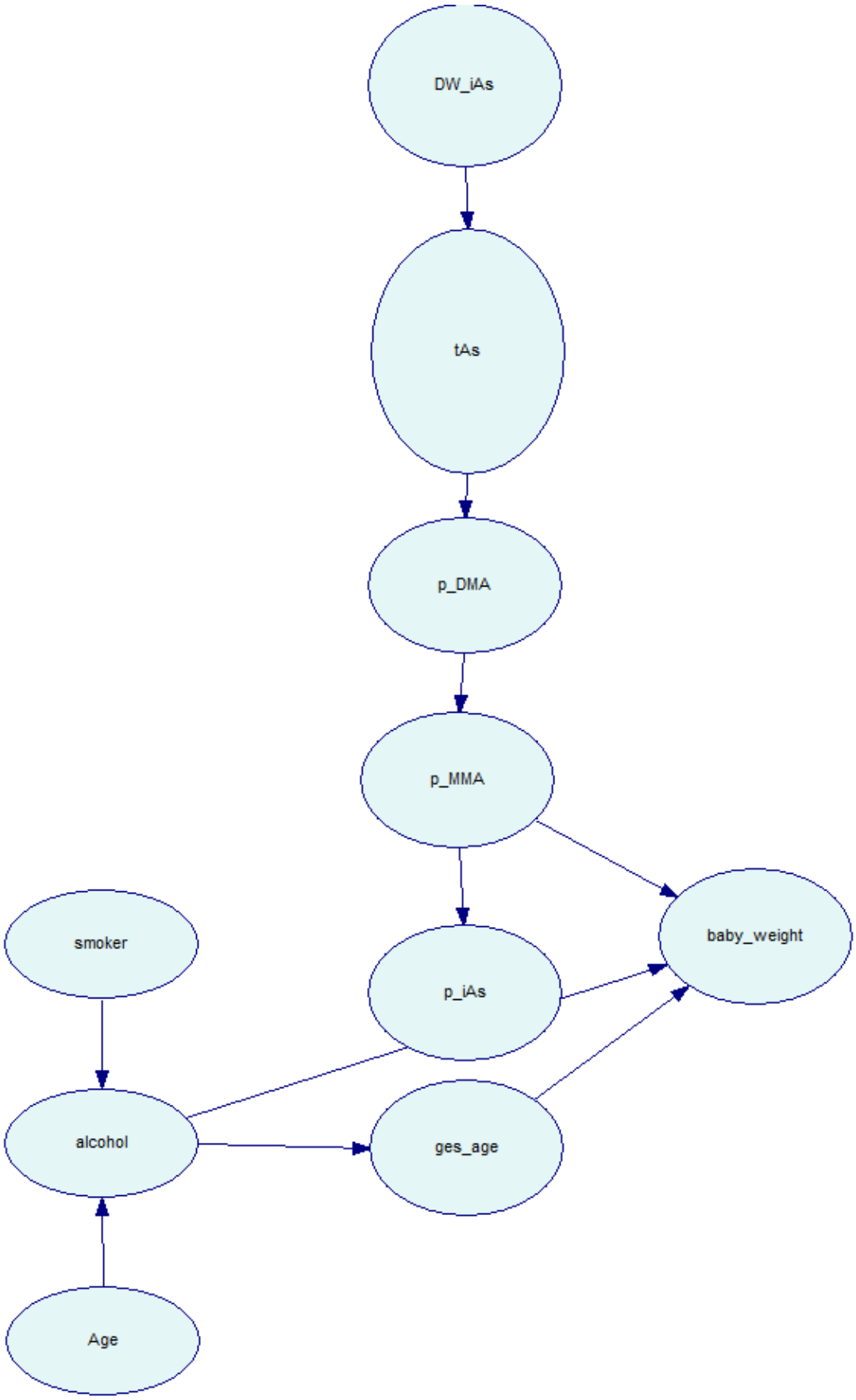


Figure 28. Influence diagram of the arsenic-exposure network (A-BN)

Second, I examine the literature for low birth weight risk and I ask several hypotheses based questions to understand the most important variables and the mechanism of risk. I implement the A-BN model inherited from the preliminary influence diagram on four literature studies (Table 16). I simplify the influence diagram to compare the four different studies because complex diagrams increase the uncertainty of the final outcome. Figure 29 shows the revised A-BN that I use to compare four literature studies: Gelmann et al. (2013), Hopenhayn et al. (2003), Punshon et al. (2015), and Laine et al. (2015).

Table 16. Comparison of four literature studies for Arsenic-Bayesian Network

Variables*	**Punshon et al.**	**Hopenhayn et al.**	**Laine et al.**	**Gelmann et al.**
Year	2015	2003	2015	2013
Sample size (N)	766	424	200	1870
Cohort region	U.S.	Chile	Mexico	Romania
Exposure (water µg/L)	0.38	40	24.6	>10
Urinary As (µg/L)	3.62	54.3	37.5	
DMA (%)	80.8	-	87.6	67
MMA (%)	9.1	-	6.4	18
iA (%)	10.1	-	6.1	15
Age	31.3	29.8	24	26.6
Smoking status (%)	6	53	7	11.1
Alcohol status (%)	-	15	20.5	0
Infant sex F (%)	50	51	96	21.9
BMI (kg/m2)	25.3	32.3	34.7	24.0
Birthweight (g)	3455	3396	3339	2405
Low birth weight (%)	4	-57 g	2	
Gestational age (weeks)		39.2	39	
MCL for country (µg/L)	10	10	25	10

*Detailed information for each study is documented in Appendix C.

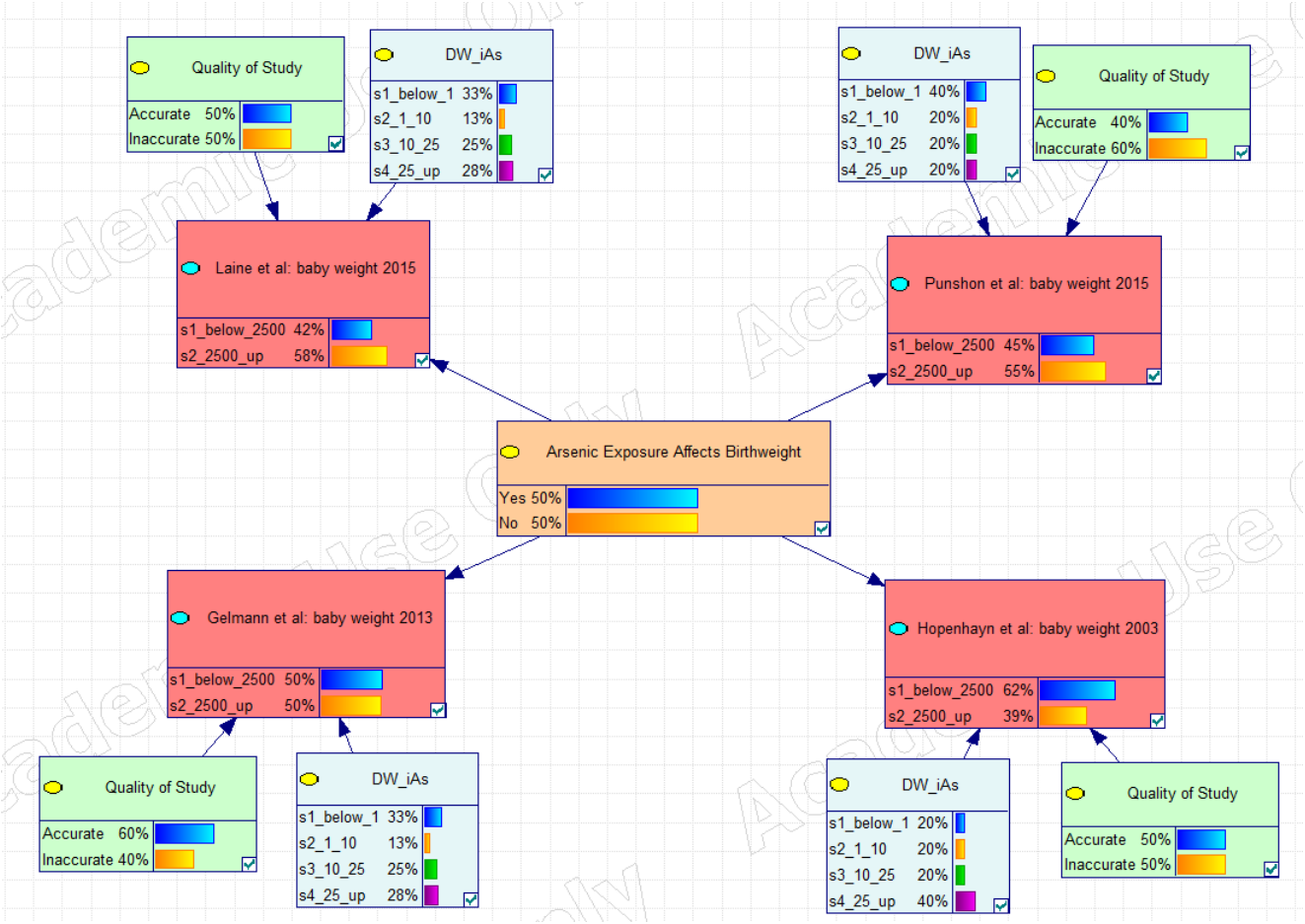


Figure 29. Preliminary revised Updated Arsenic Bayesian Network (A-BN) I use to compare four literature studies

The *Arsenic Exposure Affects Birth weight* node represents the hypothesis that exposure influences different outcomes for each study, which has a 50% prior distribution between "yes" and "no" outcomes. *Quality of study* considers all variables that affect the experimental results (e.g., sample size, limitations, etc.). I predict CPTs based on the provided information from each study (see Appendix C).

Figure 30 shows the updated network where different combinations of the states of the four study nodes surrounding the middle *Arsenic Exposure Affects Birthweight* are selected (based on the respective study outcomes), yielding an update of the middle box. A-BN outcome suggests there is no existing relationship between arsenic exposure and low birth weight (FPR=32%). On the other hand, Figure 31 shows that increasing arsenic concentration in

drinking water significantly increases the probability of an existing correlation between arsenic exposure and birth weight (FNR=6%).

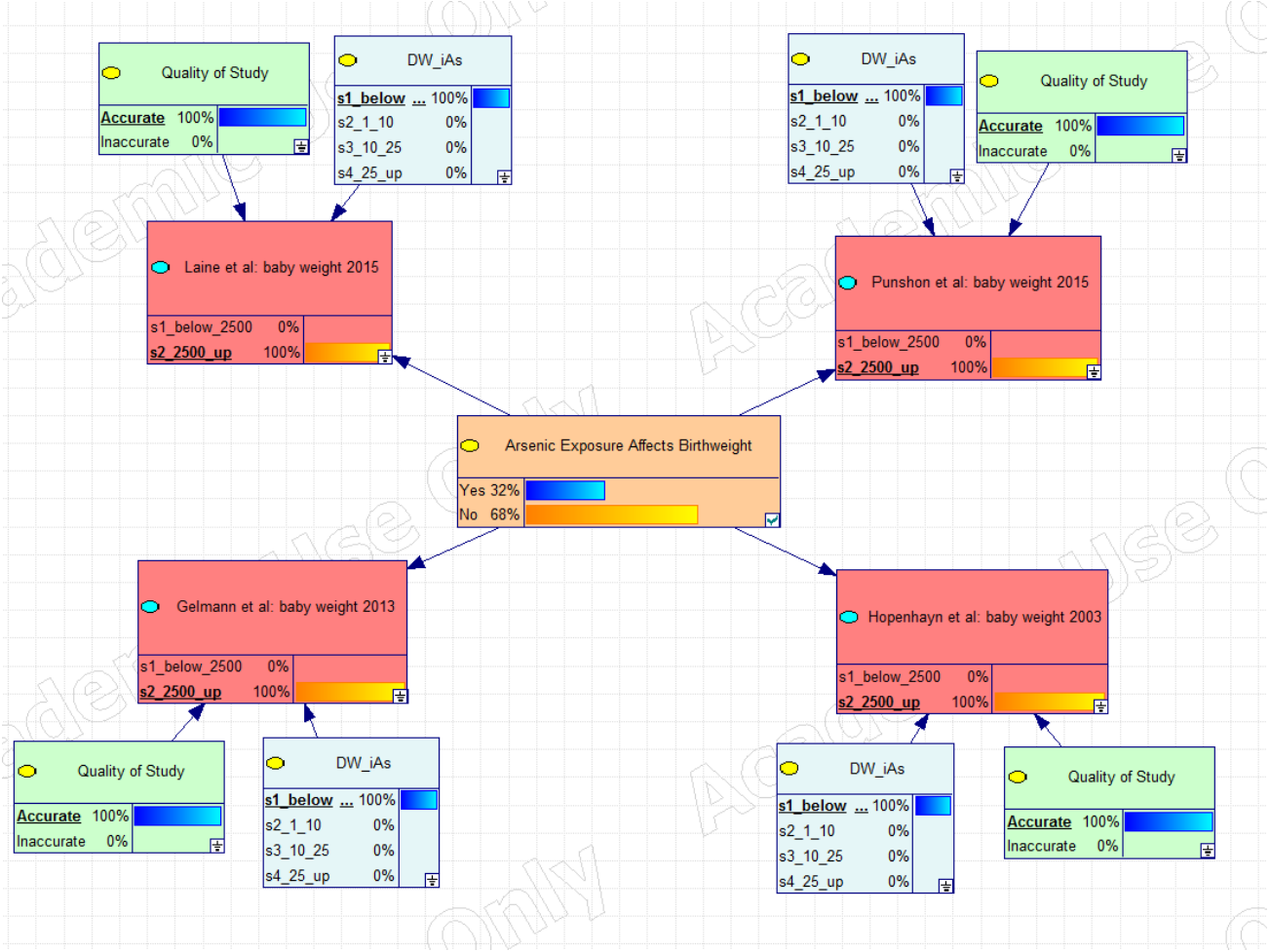


Figure 30. Updated Arsenic Bayesian Network (A-BN), DW-iAs < 1µg/L

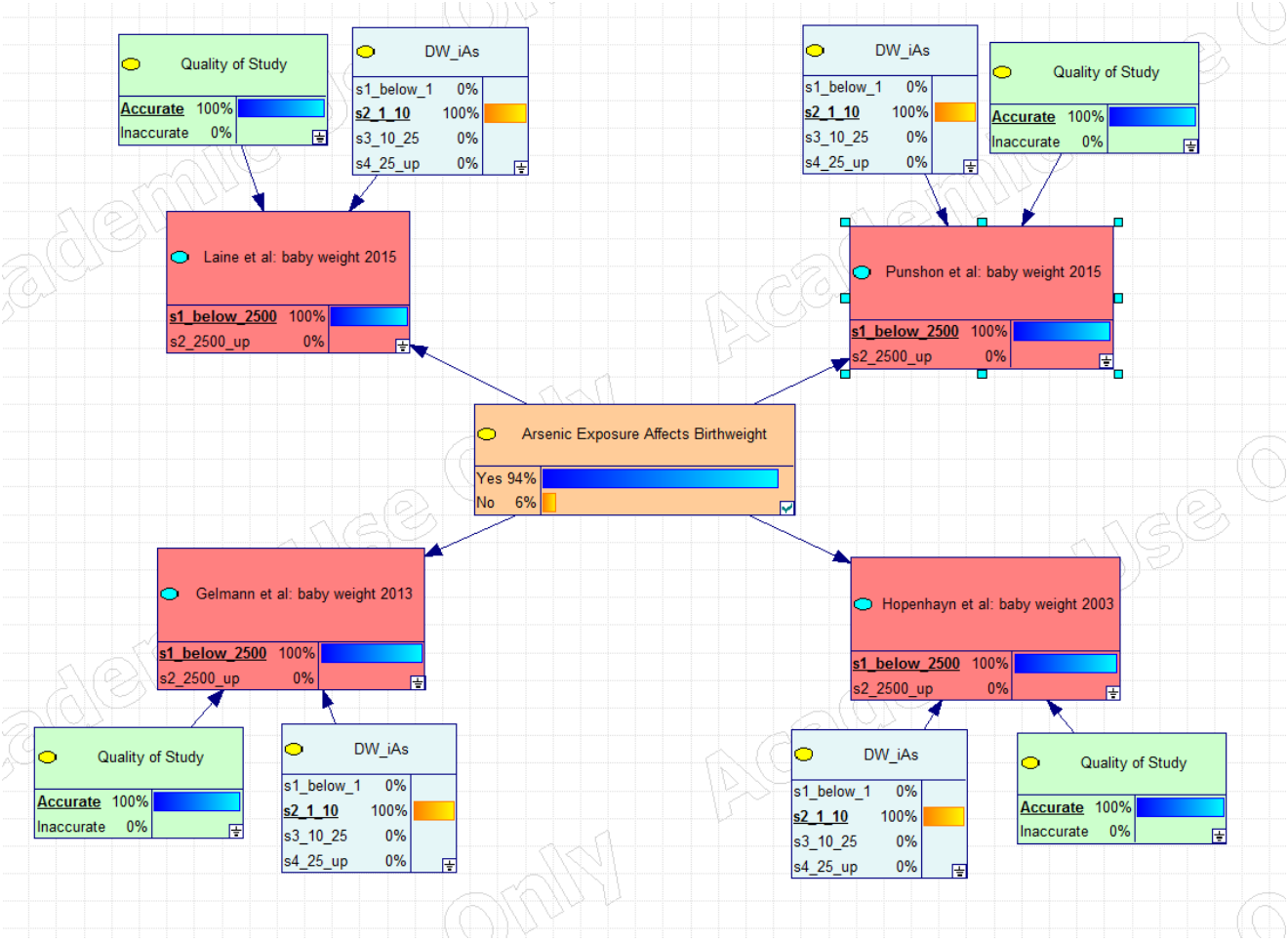


Figure 31. Updated Updated Arsenic Bayesian Network (A-BN), DW-iAs: 1-10 µg/L

These results can be used to design a better experimental plan and understand the value of information in each study. The joint risk tells more about a potential linkage between arsenic exposure and low birth weight. A-BN can be used for several potential scenarios to understand and evaluate each literature study to learn more about the arsenic-exposure network.

4. DISCUSSION AND CONCLUSION

The objective in conducting this study is to promote a broader understanding of BNs as one health risk-assessment tool. There is limited evidence in the literature about the effect of arsenic exposure on infant birth weight. There is a need to combine different outcomes from literature studies to understand and predict the potential risks. I develop A-BN to demonstrate how to implement BN to quantify the HBWoE method. The results show that this framework provides

objective and comparable outcomes to evaluate a large amount of data from the literature. This framework can be implemented in several other environmental contaminant exposure health-risk problems.

The next step for this task is combining more studies in A-BN because arsenic exposure is a well-studied problem, and I can potentially find more evidence about the effects of arsenic on human health. Also, another task will be analyzing arsenicals to try to find evidence on whether monomethylated arsenic can be an indicator of low birthweight risk.

5. ACKNOWLEDGEMENTS

I would like to thank Dr. Jacqueline MacDonald Gibson, Dr. Rebecca Fry, and Jessica Laine for providing expert knowledge on arsenic metabolism.

6. APPENDIX C

WoE for Each Hy...	Negative_birthweight							
Quality of Study	Accurate				Inaccurate			
DW_iAs	s1_below_1	s2_1_10	s3_10_25	s4_25_up	s1_below_1	s2_1_10	s3_10_25	s4_25_up
s1_below_2500	0.2	0.6	0.7	0.85	0.1	0.5	0.6	0.7
s2_2500_up	0.8	0.4	0.3	0.15	0.9	0.5	0.4	0.3

Positive_birthweigt								
Accurate					Inaccurate			
s1_below_1	s2_1_10	s3_10_25	s4_25_up	s1_below_1	s2_1_10	s3_10_25	s4_25_up	
0.02	0.25	0.45	0.6	0.15	0.2	0.4	0.55	
0.98	0.75	0.55	0.4	0.85	0.8	0.6	0.45	

Figure 32. Assumed CPT, baby weight, Laine et al.

WoE for Each Hy...	Negative_birthweight							
Quality of Study	Accurate				Inaccurate			
DW_iAs	s1_below_1	s2_1_10	s3_10_25	s4_25_up	s1_below_1	s2_1_10	s3_10_25	s4_25_up
s1_below_2500	0.2	0.7	0.8	0.9	0.1	0.5	0.6	0.8
s2_2500_up	0.8	0.3	0.2	0.1	0.9	0.5	0.4	0.2

Positive_birthweigt								
Accurate					Inaccurate			
s1_below_1	s2_1_10	s3_10_25	s4_25_up	s1_below_1	s2_1_10	s3_10_25	s4_25_up	
0.02	0.3	0.7	0.75	0.1	0.4	0.8	0.9	
0.98	0.7	0.3	0.25	0.9	0.6	0.2	0.1	

Figure 33. Assumed CPT, baby weight, Punshon et al.

WoE for Each Hy...	Negative_birthweight							
Quality of Study	Accurate				Inaccurate			
DW_iAs	s1_below_1	s2_1_10	s3_10_25	s4_25_up	s1_below_1	s2_1_10	s3_10_25	s4_25_up
s1_below_2500	0.3	0.75	0.9	0.95	0.2	0.6	0.7	0.85
s2_2500_up	0.7	0.25	0.1	0.05	0.8	0.4	0.3	0.15

Positive_birthweigt								
Accurate					Inaccurate			
s1_below_1	s2_1_10	s3_10_25	s4_25_up	s1_below_1	s2_1_10	s3_10_25	s4_25_up	
0.1	0.4	0.6	0.8	0.05	0.3	0.7	0.75	
0.9	0.6	0.4	0.2	0.95	0.7	0.3	0.25	

Figure 34. Assumed CPT, baby weight, Gelmann et al.

WoE for Each Hy...	Negative_birthweight							
Quality of Study	Accurate				Inaccurate			
DW_iAs	s1_below_1	s2_1_10	s3_10_25	s4_25_up	s1_below_1	s2_1_10	s3_10_25	s4_25_up
s1_below_2500	0.2	0.65	0.8	0.85	0.1	0.55	0.7	0.75
s2_2500_up	0.8	0.35	0.2	0.15	0.9	0.45	0.3	0.25

Positive_birthweigt								
Accurate					Inaccurate			
s1_below_1	s2_1_10	s3_10_25	s4_25_up	s1_below_1	s2_1_10	s3_10_25	s4_25_up	
0.1	0.4	0.6	0.8	0.05	0.3	0.7	0.75	
0.9	0.6	0.4	0.2	0.95	0.7	0.3	0.25	

Figure 35. Assumed CPT, baby weight, Hopenhayn et al.

Table 17. Literature study 1: Laine et al, 2015

Characteristic	*n*[a] (%) or mean, median [range]
Maternal age at delivery (years)	24, 23 [18–41]
Race	
Hispanic	199 (99.5)
Education	
< High school	50 (25.1)
High school	95 (47.7)
College	41 (20.6)
Post-college	13 (6.5)
Time living at residence (years)	20, 21 [1–41]
Smoking status	
Nonsmokers	186 (93.0)
Current smokers	13 (7.0)
Alcohol consumption	
None	159 (79.5)
Some	41 (20.5)
Prenatal vitamin daily intake	192 (97.0)
Seafood consumption	
None	155 (78.3)
Some	43 (21.7)
Previous pregnancies	
0	70 (35.0)
1	50 (25.0)
≥ 2	80 (40.0)
Previous pregnancy loss	
0	176 (88.0)
1	18 (9.0)
≥ 2	6 (3.0)
Method of delivery	
Vaginal	118 (59.0)
Cesarean section	82 (41.0)
Gestational age (weeks)	
All	39, 40 [34–42]
< 37	3 (1.5)
≥ 37	197 (98.5)
Newborn sex	
Male	104 (52.0)
Female	96 (48.0)
Birth weight (g)	
All	3339, 3355 [1800–5120]
Male	3453, 3490 [2100–5120]*
Female	3215, 3150 [1800–4200]
LBW	4 (2.0)
SGA	28 (14.0)
LGA	19 (9.5)
Placental weight (g)	648, 640 [390–1070]
Newborn length (cm)	50, 50 [40–59]
Head circumference (cm)	35, 35 [31–38]
APGAR score	9, 9 [8–10]
Exposure measures	
DW-iAs (µg As/L)	24.6, 13.0 [< LOD[b]–236.0]
Urinary arsenicals[c]	
U-tAs (µg/L)	37.5, 23.3 [4.3–319.7]
U-iAs (µg/L)	2.1, 1.3 [0.14–23.0]
U-MMAs (µg/L)	2.3, 1.4 [0.12–18.2]
U-DMAs (µg/L)	33.1, 20.6 [1.4–292.5]
iAs (%)	6.1, 5.3 [0.77–45.1]
MMAs (%)	6.4, 6.0 [0.68–24.9]
DMAs (%)	87.6, 88.5 [32.7–96.7]
MMAs/iAs	1.2, 1.2 [0.13–5.5]
MMAs/DMAs	0.077, 0.069 [0.0072–0.68]
DMAs/MMAs	17.6, 14.6 [1.5–140.0]
MMAs + DMA/iAs	19.6, 18.1 [1.2–129.9]

[a]Differences in *n* based on missing demographic data. [b]LOD for DW-iAs = 0.456 µg As/L. [c]All urinary values were adjusted by SG. *Significant difference in means

Table 18. Literature study 2: Punshon et al, 2015

Characteristics	N *(%)*	*Placenta As (ng/g) Mean (± SD)*
Maternal:		
Age at enrollment (years):		
< 30 years	308 (40)	1.08 (1.52)
≥30 years	456 (60)	1.15 (1.64)
BMI (kg/m^2)		
Normal (BMI < 25)	397 (58)	1.16 (1.63)
Overweight (≥ 25 to < 30)	170 (25)	1.06 (1.25)
Obese (≥30)	114 (17)	0.92 (0.93)
Smoking status		
Smoker	44 (6)	0.96 (0.85)
Non-smoker	649 (94)	1.10 (1.42)
Parity		
First live birth	307 (41)	1.16 (1.42)
1 or more live birth	447 (59)	1.12 (1.72)
Infant:		
Sex		
Female	378 (50)	1.13 (1.55)
Male	385 (50)	1.15 (1.64)
Birth weight (g)		
Low (<2500)	28 (4)	1.04 (1.22)
Normal (≥2500)	725 (96)	1.14 (1.62)

Arsenic Variable	*% Increase in placental Arsenic (95% CI)*	N	P
Biomarkers			
Maternal urinary[a,b] (μg/l)	31.0 (15.4, 48.7)	431	< 0.0001
Maternal toenail[c] (μg/g)	13.2 (2.0, 25.7)	579	0.0196
Infant toenail[d] (μg/g)	18.9 (1.8, 38.8)	151	0.0293
Exposure[e]			
Household drinking water[c] (μg/l)	2.1 (1.3, 3.9)	716	< 0.0001

[a]Excluding arsenobetaine. [b]Adjusted for creatinine concentration (mg/dl). [c]No factors appreciably altered the estimates (see text). [d]Adjusted for parity. [e]Based on a doubling of exposure.

Table 19. Literature study 3: Gelmann et al, 2013

Unexposed Study Participants			
Characteristic	**NBW[b] (n=10)**	**LBW[c] (n=9)**	**p-value[d]**
Age (years)	30.2 ± 3.0	30.6 ± 3.1	0.80
Birth weight of child (g)	3352.0 ± 352.5	2355.6 ± 104.4	<0.01
BMI (kg/m^2)	22.7 ± 3.8	22.8 ± 3.9	0.94
Education (years)	13.9 ± 2.4	11.9 ± 4.8	0.25
Fasting blood glucose, 3rd trimester (mg/dL)	85.4 ± 17.1	86.0 ± 11.2	0.93
Number of prior pregnancies	1.2 ± 1.0	1.9 ± 2.0	0.35
Use of prenatal vitamins (%)	80.0 (8)	77.8 (7)	1.00
Current smoker (%)	10.0 (1)	33.3 (3)	0.30
Smoked during pregnancy (%)	10.0 (1)	33.3 (3)	0.30
Consumes alcohol (%)	0.0 (0)	0.0 (0)	NA

Exposed Study Participants			
Characteristic	**NBW[b] (n=10)**	**LBW[c] (n=9)**	**p-value[d]**
Age (years)	31.1 ± 5.5	26.6 ± 5.4	0.86
Birth weight of child (g)	3470.0 ± 222.6	2405.6 ± 212.8	<0.01
BMI (kg/m^2)	24.5 ± 3.2	21.9 ± 3.4	0.11
Education (years)	10.6 ± 2.4	6.6 ± 3.8	0.02
Fasting blood glucose, 3rd trimester (mg/dL)	74.0 ± 5.2	76.3 ± 3.4	0.27
Number of prior pregnancies	1.1 ± 1.3	1.6 ± 2.2	0.58
Use of prenatal vitamins (%)	80.0 (8)	66.7 (6)	0.63
Current smoker (%)	20.0 (2)	11.1 (1)	1.00
Smoked during pregnancy (%)	10.0 (1)	11.1 (1)	1.00
Consumes alcohol (%)	0.0 (0)	0.0 (0)	NA

[a] Values are mean ± SD for continuous variables and column % (n) for categorical variables.

[b] NBW = Normal Birth Weight

[c] LBW = Low Birth Weight

[d] P-value is for t-test (continuous variables) or χ^2 test/Fisher's exact test (categorical variables).

Table 20. Literature study 4: Hopenhayn et al, 2003

Characteristic	Antofagasta (N = 424)	
Maternal age at interview (years) (No. and %)		
18–20	108	26
21–25	126	30
26–35	157	37
36+	33	8
Maternal education (years in school) (No. and %)		
0–8	102	24
9–12	230	54
>12	92	22
Marital status (No. and %)		
Married/co-habitating	310	73
Single/separated/widowed	114	27
Missing	0	
Maternal ethnicity* (No. and %)		
Indigenous	42	10
Nonindigenous	382	90
Monthly household income (US$) (No. and %)		
0–≤ $400	254	60
$401–≤ $700	121	29
>$701	47	11
Missing	2	1
Parity (No. and %)		
0	159	38
1	137	32
2	81	19
3 or more	50	12
Maternal body mass index (BMI) (No. and %)		
1st tertile (16.21–23.5)	159	38
2nd tertile (23.5–27.1)	122	29
3rd tertile (27.2–48.99)	132	31
Missing	11	3
Gestational age at first prenatal care visit (No. and %)		
≤13 weeks	306	72
14–27 weeks	101	24
≥28 weeks	7	2
Missing	10	2
Number of prenatal care visits (No. and %)		
0–5	19	5
6–10	175	41
>10	220	52
Missing	10	2
Adequacy of prenatal care (No. and %)		
Inadequate	14	3
Intermediate	75	18
Adequate	325	77
Missing	10	2
Maternal height (meters) (Mean and SD)	1.56	0.06
Maternal weight gain during pregnancy (kilograms) (Mean and SD)	10.9	4.8

Table 21. Literature study 4: Hopenhayn et al, 2003 (continue)

Characteristic	Antofagasta (N = 424)	
Alcohol* consumption during pregnancy		
Yes (No. and %)	62	15
No (No. and %)	362	85
Missing (No. and %)	0	
Alcoholic drinks per week. (Mean and SD)	0.8	0.9
Cigarette smoking prior to pregnancy		
None (No. and %)	195	46
< 5 cigarettes/day (No. and %)	147	35
≥ 5 cigarettes/day (No. and %)	78	18
Missing (No. and %)	4	1
Cigarettes per day for smokers (Mean and SD)	4.9	5.1
Cigarette smoking at time of interview		
None (No. and %)	385	91
< 5 Cigarettes/Day (No. and %)	35	8
≥ 5 Cigarettes/Day (No. and %)	2	1
Missing (No. and %)	2	1
Cigarettes per day for smokers (Mean and SD)	1.3	1.4
Coffee consumption during pregnancy (No. and %)		
Yes	50	12
No	374	88.2
Caffeinated† beverage consumption during pregnancy (No. and %)		
Yes	406	96
No	18	4
Vitamins or supplements taken during pregnancy (No. and %)		
Vitamin complex	85	20
Iron	261	62
Folic acid	94	22
Calcium	195	46
Index of Everyday Stressors‡ (Mean and SD)	19.1	9.5
Index of Major Life Stressors§ (Mean and SD)	1.8	1.3
Walking on a daily basis (No. and %)		
≤ 9 Blocks/Day	222	52
> 9 Blocks/Day	202	48
Missing	0	
Fats (grams) (Mean and SD)	66.6	28.9
Carbohydrates (grams) (Mean and SD)	355.0	126.7
Proteins (grams) (Mean and SD)	85.6	27.7

7. REFERENCES

Ahmad, SA, Sayed, MHSU, Barua, S, *et al.* 2001. Arsenic in Drinking Water and Pregnancy Outcomes. Environmental health perspectives. 109.

Bailey, KA and Fry, RC. 2015. Health Effects of Prenatal and Early-Life Exposure to Arsenic. Handbook of arsenic toxicology. 405–28.

Bailey, LA, Nascarella, MA, Kerper, LE, *et al.* 2016. Hypothesis-based weight-of-evidence evaluation and risk assessment for naphthalene carcinogenesis. Critical reviews in toxicology. 46: 1-42.

bayesfusion.com, 2016b, GeNIe 2016.

Caldwell, KE, Labrecque, MT, Solomon, BR, *et al.* 2015. Prenatal arsenic exposure alters the programming of the glucocorticoid signaling system during embryonic development. Neurotoxicology and teratology. 47: 66-79.

Druzdzel, M, 2003, Stochastic Sampling Algorithms: EPIS Sampling, Accessed 01.27 2016, https://dslpitt.org/genie/wiki/Stochastic_Sampling_Algorithms:_EPIS_Sampling.

Gardner, RM, Nermell, B, Kippler, M, *et al.* 2011. Arsenic methylation efficiency increases during the first trimester of pregnancy independent of folate status. Reproductive toxicology. 31: 210-8.

Gelmann, ER, Gurzau, E, Gurzau, A, *et al.* 2013. A pilot study: the importance of inter-individual differences in inorganic arsenic metabolism for birth weight outcome. Environmental toxicology and pharmacology. 36: 1266-75.

Hopenhayn, C, Ferreccio, C, Browning, SR, *et al.* 2003. Arsenic Exposure from Drinking Water and Birth Weight. Epidemiology. 14: 593-602.

Jarup, L. 2003. Hazards of heavy metal contamination. British Medical Bulletin. 68: 167-82.

Laine, JE, Bailey, KA, Rubio-Andrade, M, *et al.* 2015. Maternal arsenic exposure, arsenic methylation efficiency, and birth outcomes in the Biomarkers of Exposure to ARsenic (BEAR) pregnancy cohort in Mexico. Environmental health perspectives. 123: 186-92.

Mandal, BK and Suzuki, KT. 2002. Arsenic round the world: a review. Talanta. 58: 201-35.

McCann, RK, Marcot, BG, and Ellis, R. 2006. Bayesian belief networks: applications in ecology and natural resource management. Canadian Journal of Forest Research. 36: 3053-62.

NRC. 2001. Arsenic in Drinking Water: Update. National Research Council, Washington, D.C., U.S.

Punshon, T, Davis, MA, Marsit, CJ, *et al.* 2015. Placental arsenic concentrations in relation to both maternal and infant biomarkers of exposure in a US cohort. Journal of exposure science & environmental epidemiology. 25: 599-603.

Qi, Y, Li, H, Zhang, M, *et al.* 2014. Autophagy in arsenic carcinogenesis. Experimental and toxicologic pathology : official journal of the Gesellschaft fur Toxikologische Pathologie. 66: 163-8.

Rager, JE, Bailey, KA, Smeester, L, *et al.* 2014. Prenatal arsenic exposure and the epigenome: altered microRNAs associated with innate and adaptive immune signaling in newborn cord blood. Environmental and molecular mutagenesis. 55: 196-208.

Rhomberg, LR. 2013. Hypothesis-Based Weight of Evidence: An Approach to Assessing Causation and its Application to Regulatory Toxicology. Methods for Research Synthesis: A Cross-Disciplinary Workshop, Harvard Center for Risk Analysis.

Rhomberg, LR and Bailey, LA. 2015. The Accounts Table for Weighing Diverse Lines of Evidence to Assess Causality. Society for Risk Analysis

Stanton, BA, Caldwell, K, Congdon, CB, *et al.* 2015. MDI Biological Laboratory Arsenic Summit: Approaches to Limiting Human Exposure to Arsenic. Current environmental health reports. 2: 329-37.

Thomas, DJ, Styblo, M, and Lin, S. 2001. The cellular metabolism and systemic toxicity of arsenic. Toxicology and applied pharmacology. 176: 127-44.

USNRC. 2013b. Critical Aspects of EPA's IRIS Assessment of Inorganic Arsenic Interim Report. National Research Council of the National Academies, 978-0-309-29706-6, Washington D.C., U.S.

Vahter, M. 2002. Mechanisms of arsenic biotransformation. Toxicology. 181-182: 211-7.

Vahter, M, Li, L, Rahman, A, *et al.* 2006. Does metabolism of arsenic affect the toxicity during early human development? Toxicol Lett.

Wanibuchi, H, Salim, EI, Kinoshita, A, *et al.* 2004. Understanding arsenic carcinogenicity by the use of animal models. Toxicology and applied pharmacology. 198: 366-76.

WHO. 2011. Arsenic in Drinking-water. World Health Organization, WHO/SDE/WSH/03.04/75/Rev/1.

Wilson, JD. 2001b. Advanced Methods for Dose–Response Assessment: Bayesian Approaches—Final Report. Resources for the Future, Washington, D.C., U.S.

CHAPTER 4: THESIS CONCLUSIONS AND FUTURE WORK

The overall objective of this dissertation was to highlight the importance of statistical modeling and data driven knowledge for environmental-health risk assessment. This was accomplished by (1) reviewing the potential risks of missing chemical data and concentration variability by developing 27 occurrence scenarios. (2) Demonstrating how a performance analysis can be implemented for a Bayesian Network (BN) representation of a dose-response relationship. (3) Analyzing the risk factors of a prenatal arsenic exposure network by combining BN modeling and Hypothesis Based Weight of Evidence method as a tool for health risk assessment.

1. CONCLUSIONS

Chapter 1 of the thesis focused on a predictive occurrence model to analyze the effects of concentration variability in the system and correlated concentrations on mixture toxicities. I found that higher variability in concentrations causes higher effective (average) mixture toxicity when low to intermediate toxicity is associated with the median of the concentration distribution. The biggest enhancement of toxicity due to variability occurs at lower concentrations, however, if the variability is very high, toxicity is slightly lower for high median concentration values. Our findings showed that correlated concentrations do not systematically lead to increased mixture toxicity. The increased toxicity effect due to concentration variability is similar for concentration addition and independent action models. These results show that mixture occurrence and toxicity estimation should be explored to prioritize exposure sampling and mixture toxicity studies.

Chapter 2 explored the effect of different sample sizes on predicting the strength of the relationship between true responses and true doses of environmental toxicants. Our findings show that increasing actual strength of relationship increases the probability of accurately predicting *relationship (R)* classification. Also, increasing the sample size increases the

accuracy level for the predicted *R* for all scenarios. Increasing the experimental accuracy level significantly contributes to the efficiency of *R* prediction. The findings can guide the use of dose-response studies in regulatory decision-making by determining if data analysis is valid in certain cases, according to the strength of interactions between variables and the sample size.

Chapter 3 explored the prenatal, inorganic arsenic-exposure network, the strength of interactions, and the potential causal relationships by combining Hypothesis Based Weight of Evidence (HBWoE) and Bayesian Network (BN) modeling. Our analysis demonstrates how BN can be used to quantify HBWoE evaluation. I used four literature studies to predict the risk of low birth weight. Results can be used to design a better experimental plan and understand the value of information in each study. There is limited evidence in the literature about the effect of arsenic exposure on infant birth weight. This approach can be used for several potential scenarios to understand and evaluate each literature study to learn more about the arsenic-exposure network.

2. FUTURE WORK

Although this study advances the understanding of statistical methods in environmental-health risk assessment, this research could be further improved by additional work not addressed in this study. An important limitation of Chapter 1 is that the models are based on a specific set of 10 chemicals, with results that may not be representative of other mixtures. Also, for this study, mixture toxicity is dominated by the most toxic compound, ofloxacin, but this result could be different if no chemical dominates the mixture toxicity. Some advancements include the addition of different chemicals in the model and trying different mixtures.

Chapter 2 could be improved by including more categories in each node and repeating the process for different conditional dependencies. Also, our model outcomes are based on matching exposure-response accuracy levels, so an important future step could be trying

different combinations of accuracy levels. Findings could change for different prior distributions. Moreover, definition of medium and strong relationship could potentially change the outcomes, so different scenarios could help learning more information.

The next step for Chapter 3 could be combining more studies in A-BN because arsenic exposure is a well-studied problem, and I can potentially find more evidence about the effects of arsenic on human health. Also, another task will be analyzing arsenicals to try to find evidence on whether monomethylated arsenic can be an indicator of low birth weight risk. The A-BN could include more variables to objectively and systematically prioritize the variables in the network. This methodology could be applied to other environmental contaminant exposures.