A Method for Biomarker Validation and Biomarker-Based Dose Response: A Case Study with a Bayesian Network Model for Benzene

C. Eric Hack[†, #], Lynne T. Haber[†*], Andrew Maier[†], Paul Schulte[‡], Bruce A. Fowler[§], W. Gregory Lotz[‡] and Russell E. Savage Jr.[‡]

[†]*Toxicology Excellence for Risk Assessment (TERA), Cincinnati, Ohio 45211*
Lynne T. Haber*: haber@tera.org*
Andrew Maier*: maier@tera.org*

[#]*Current Affiliation: Henry Jackson Foundation, AFRL/HEPB, Wright-Patterson AFB, Ohio 45433*
C. Eric Hack*: Charles.Hack@WPAFB.AF.MIL*

[‡]*National Institute for Occupational Safety and Health (NIOSH), Cincinnati, Ohio 45226*
Paul Schulte: *PAS4@CDC.GOV*
W. Gregory Lotz: *WGL0@CDC.GOV*
Russell E. Savage: *ras6@cdc.gov*

[§]*Agency for Toxic Substances and Disease Registry (ATSDR), Atlanta, Georgia 30333*
Bruce Fowler: *BXF9@CDC.GOV*

[*]To whom correspondence should be addressed.
2300 Montana Avenue, Cincinnati, Ohio 45211
Phone: (513) 542-7475
Fax: (513) 542-7487
email: Haber@tera.org

ABSTRACT

For over two decades, scientists have been touting the importance of the application of biomarkers in reducing disease and protecting individuals from the harmful effects of exposure to occupational and/or environmental chemicals. While established guidelines for biomarker validation exist, a well documented systematic approach to validating biological endpoints is not always conducted before such endpoints are labeled as biomarkers. This pilot study seeks to develop and demonstrate the use of a system for integrating complex and multifaceted data, validating biomarkers, and incorporating the biomarkers into an occupational risk assessment. A biomarker database structure was developed and decision rules were identified to organize the diverse types of data to support a benzene risk assessment. A suite of biomarker validation approaches was applied to evaluate potential biomarkers of exposure and effect for key chronic cancer and noncancer endpoints, specifically the hematotoxicity endpoints of anemia, leukopenia, and leukemia. Traditional biomarker evaluation approaches based on the Hill criteria and regression analysis were coupled with a Bayesian network approach to test and validate (or discount) biomarkers along the exposure-disease continuum. Dose-response analyses using validated biomarkers were conducted to contrast effective concentration estimates based on various biomarker approaches. Recommendations for general use of biomarkers in risk assessment are discussed. Further refinement is needed of the quantitative validation techniques and its application to OEL development for other diverse datasets.

Key Words: Biomarker, Identification, Validation, Application, OELs, Risk Assessments

INTRODUCTION

While it is difficult to trace the origin of the term "biomarker," it is clear that biomarkers have found widespread use in contemporary biomedical sciences. The first documented merger of "molecular" with "epidemiology" occurred in 1973 (Kilbourne, 1973) with "biomarkers" serving as the cornerstone for this unique discipline. In 1987, the National Research Council on Environmental Studies and Toxicology defined a biomarker as any cellular or molecular indicator of toxic exposure, adverse health outcome or susceptibility. That definition paved the way for a further delineation of biomarkers into three categories – biomarkers of exposure, effects and susceptibility; these categories are often cited for their potential utility in environmental, occupational or other related toxicological applications. While various disciplines foster unique applications of biomarkers to their special needs, a shared aspect is that the transformation from a measured biological endpoint to a validated biomarker with a specific public health application is a difficult and tedious process. Albertini (2001) noted that the mere availability of biomarkers does not mean that they will be useful for human studies directed at public health issues, and that although many biomarkers have been identified, few have been validated to the point of known usefulness (Albertini, 1998). Perera (2000) pointed out that although biomarkers can play a key role in cancer epidemiological studies, many studies failed to use validated biomarkers or employed study designs that did not adequately consider the biology of the endpoints. Established guidelines for biomarker validation exist (Barker, 2003; Schulte, 2005) and Schulte (2005) emphasized the importance of validation prior to use of biomarkers. However, lack of validation has continued to be a key barrier to increased biomarker application in human health risk assessment (Maier et al., 2004).

To attempt to address this barrier, we describe a biomarker decision support system to facilitate the systematic identification, documentation, and validation of biomarkers. We focus on application of the concept to occupational risk assessment, since biomarkers of exposure have a rich history of use in this area and because there is significant potential for biomarkers to enhance the scientific underpinnings in the development of occupational exposure limits.

To test and demonstrate the decision support system for occupational safety and health applications, benzene was chosen as the case study chemical, because it is an occupationally-important chemical that had an occupational exposure limit (OEL) that may benefit from an updated assessment based on newer data, and data on an array of potential biomarkers are available. In this case study, we extracted biomarker data from public databases or the appropriate literature, integrated the data into a user-friendly biomarker database of our design, employed a decision support system to analyze the biomarker data using a suite of validation tools, and demonstrated the contribution of biomarker data in the dose-response assessment for OEL development. The decision support system includes a suite of analytical tools, with the appropriate tools chosen based on the data available for the chemical of interest. For the benzene case study, we used a relatively novel approach of using a Bayesian network to quantitatively incorporate data on biomarkers of effect in determining the point of departure, as opposed to the conventional approach of using either effect or biomarker data independently as the point of departure. A goal of the use of the Bayesian network is to provide a quantitative approach for linking changes in biomarkers of effect (also called precursors to the adverse effect) both to exposure information and to changes in disease response, thus providing a scientifically valid point of departure that incorporates precursor dose-response information without being dependent on the difficult issue of a definition of adversity for precursors.

METHODS AND RESULTS

Biomarker Database

This project developed and implemented a systematic approach for biomarker identification, documentation, validation, and incorporation into an occupational risk assessment. Our study approach involves identifying a case-study chemical, developing a biomarker database, and applying a decision support system for data evaluation. The steps in the decision support system are a) categorization of inputs; b) data evaluation; c) analysis and validation of biomarkers and dose-response; d) development of the risk assessment output (Figure 1).

Benzene was chosen as the case study chemical based on (1) the potential for occupational exposure, (2) severity of toxic response, (3) availability and continuity of biomarker data (i.e., linkage to effects), (4) adequate knowledge of mode of action, and (5) the perceived need for an updated OEL, based on the availability of new data. Based on a review of the data on benzene biomarkers, benzene toxicology, and information on the underlying biology, a list of potential biomarkers was developed. We initially sought to evaluate a diversity of biomarkers, in light of the pilot nature of the study. As a result of the wealth of data for benzene, we were able to exclusively study biomarkers derived from human studies as a simplified approach for this case study.

Relevant studies were identified by searching the literature using the National Library of Medicine TOXNET cluster of databases and the Medline database, from recent comprehensive toxicology reviews, and by contacting researchers familiar with benzene toxicology. We

developed a customized biomarker database in Microsoft Access® to house data related to benzene biomarkers and document study selection. The relational database included four levels: 1) general study information; 2) species or cell type investigated; 3) specific information on biomarkers reported; and 4) dose-response data. Copies of the database are available from the authors upon request. Although only human data were used for this demonstration project, the biomarker database was developed to accommodate data from animal and in vitro studies as well. Studies were categorized as providing information on one or more biomarker types: exposure, internal dose, effective dose, early effects, mild effects, severe effects, or susceptibility. These categories have been well described elsewhere and we applied definitions consistent with those provided in risk assessment guidance documents (e.g., EPA, 1994). This initial categorization step provided documentation of the range of studies identified and served as an initial screening point to allow for additional targeted literature searches that focused on biomarker areas of most interest.

In order to assure adequacy of the biomarker data, the full list of studies identified was culled prior to further analysis. Studies were retained if: 1) measured endpoints had accompanying exposure estimates; 2) multiple endpoints spanning the exposure-disease continuum in the same population were evaluated; 3) the endpoints evaluated were those most relevant to the diseases (adverse effects) under consideration; and 4) study quality was adequate to yield reliable estimates.

A significant challenge in developing quantitative evaluation approaches that combine diverse data types is the need to format data in a manner that maintains the integrity of the data, but allows for combining data sets. For studies meeting the study adequacy criteria described in the previous paragraph, biomarker data were normalized to facilitate inter-study data

comparisons and the combining of studies for the validation analysis. This was done by converting common endpoints reported in different studies to the same response units. For example data were converted to consistent units, or were converted to similar metrics, such number versus percentage of cells affected. For comparing data for different endpoints, the data were normalized relative to some nominal response level.

Data were collected from several studies of different groups with different exposure scenarios and experimental designs. To enable quantitative analysis, the data were grouped into exposure categories determined by examining the range of exposure associated with the various candidate biomarker and disease observations. The data were ordered by increasing exposure concentration, and cut-points for the exposure categories were chosen with the goal of spreading the observations as evenly as possible over the exposure categories. For example, the exposure ranges were chosen so that observations of leukemia were in different exposure groups, rather than being lumped together. Using the ranges of exposures also facilitated the evaluation of the linkage across the exposure-disease continuum since various biomarkers and effects were grouped together in a common exposure category. The eight resulting exposure groups were 0 up to 0.01 ppm, 0.01 up to 0.5 ppm, 0.5 up to 1.5 ppm, 1.5 up to 3 ppm, 3 up to 10 ppm, 10 up to 25 ppm, 25 up to 50 ppm, and 50 ppm and greater. While the selection of cut-points was a qualitative exercise for this demonstration project, it is a typical meta-analysis approach for such efforts.

Table 1 lists the studies that were used for the biomarker validation analysis after applying initial data adequacy screening criteria to the larger pool of available studies identified in the published literature.[1]

Biomarker Analysis and Validation

The generic decision support system was designed to accommodate the use of a suite of validation tools, and then applied to benzene. Data sets were evaluated using a variety of qualitative and quantitative approaches, including use of the Hill criteria, for which EPA (2005) gives some of the most detailed discussion of the applications for risk assessment. The particular analytical tools used for a given chemical will vary depending on the nature of the data set available, and might include linear regression, qualitative graphical analysis approaches, or more complex techniques such as physiologically-based pharmacokinetic modeling or Bayesian network modeling. For the benzene case study, the analytical tools were qualitative analysis using the Hill criteria, qualitative graphical analysis, linear regression, and Bayesian network modeling.

Although the mechanism of action of benzene's effects on hematological outcomes is not fully known, there are several potential biomarkers of benzene exposure or of effects of exposure, and several have been linked to clinical disease. The disease endpoints considered in this analysis were anemia, leukopenia, and leukemia. Central nervous system (CNS) effects were not included in the analysis, since they are most relevant following acute high-level exposures and result from a different pathway of precursors.

The identified biomarkers were classified as biomarkers of exposure (Table 2a) or effect (Table 2b), and ordered into pathways leading to the relevant endpoints, including alternative pathways, when relevant. While there were multiple human studies providing data for most of the biomarkers of exposure (except for the albumin adducts and urinary benzene triol), the analysis of most of the biomarkers of effect was limited by the availability of only one useful human study that tied the biomarker data to exposure measurements. In addition, for two of the

effect categories, all or most of the available biomarker data came from only one study. Possible exposure biomarkers considered included benzene in the blood, urine, and exhaled air; urinary levels of the benzene metabolites trans,trans-muconic acid (ttMA), S-phenylmercapturic acid (SPMA), catechol, benzene triol, and phenol; and protein adducts formed by benzene metabolites. The initial identification of mechanistic pathways allowed for several pathways to benzene-induced anemia and leukopenia, as well as for leukemia. Three possible pathways to leukemia were considered with multiple biomarkers considered for each pathway. The pathways considered involved (1) genotoxicity, (2) oxidative stress (which could lead indirectly to DNA damage), and (3) changes in gene expression related to cell proliferation. As noted earlier, the choices of potential biomarkers were limited by the decision to use only in vivo human data. Two pathways to anemia were considered. One pathway involved disruption of heme synthesis, and several biomarkers in the heme synthesis pathway were considered. The other possible pathway for benzene-induced anemia begins with decreases in blood progenitor cells, which may be caused directly by cytoxicity mechanisms or secondary to genotoxicity. Decreases in progenitor cells could also lead to leukopenia. This initial set of pathways was refined as part of the validation approaches described in the following paragraphs, ultimately leading to the network presented in Figure 5.

It should be noted that the potential pathways identified were developed and refined by combining current information on mode of action hypotheses with biomarkers for which data amenable to analysis were available. This approach does not preclude the involvement of other relevant biological events, which were not included in the analysis due to the lack of adequate dose-response data.

A suite of validation approaches was used in choosing the biomarkers that ultimately served as the basis for the dose-response analysis. The first validation approach involved qualitative application of the Hill criteria of (1) consistency, (2) strength, (3) specificity, (4) temporal relationship, (5) biological plausibility/coherence, and (6) dose-response relationship (EPA 2005) to evaluate the relationship between the biomarker and the disease endpoint of interest (leukemia, anemia, or leukopenia). Biological plausibility/coherence was evaluated against current hypotheses in benzene mechanisms as articulated in several recent comprehensive toxicity reviews (Bird et al., 2005; Eastmond et al., 2001; Smith, 1996). All of the biomarkers under consideration passed this criterion. To evaluate the Hill criterion of temporality, one would ideally compare the duration of exposure required for changes in biomarker levels compared with the duration required for effects on disease incidence. This was not possible, however, given the nature of the datasets evaluated (human epidemiology studies with undefined or only longer-term exposures). In lieu of this information on temporal patterns of exposure and response, a qualitative approach was taken, using a general mode of action understanding to judge whether each biomarker would be expected to be affected prior to the clinical outcome of concern.

Specificity is the weakest of the Hill criteria, and the least likely to be fulfilled. As is commonly seen with effect endpoints and biomarkers, none of the effect biomarkers are specific to benzene. The specificity of several of the exposure biomarkers was reviewed by ACGIH (2001) in its consideration of biological exposure indices (BEIs). The measures of benzene internal dose, benzene in blood, urine, and breath, are specific to benzene. ACGIH (2001) considered the urinary levels of the benzene metabolites ttMA and SPMA to be sufficiently specific that both were adopted as recommended BEIs. Neither of these compounds are

themselves toxic agents, although ttMA is produced by the oxidation of the toxic intermediate trans,trans-muconaldehyde. In contrast to SPMA and ttMA, urinary levels of the benzene metabolites hydroquinone, catechol, 1,2,4-benzene triol, and phenol are not specific to benzene, because, in addition to reflecting benzene exposure, these levels may also reflect direct exposure to these chemicals or exposure to other chemicals that are metabolized to the same compounds. Albumin adducts of benzene oxide (BO) and 1,4-benzoquinone (BQ) are fairly specific, and reflect internal dose of reactive metabolites that have been proposed to cause the hematotoxic and leukemogenic effects. Because specificity is a weak criterion, this criterion alone was not used to exclude potential biomarkers.

In considering the dose-response, a key decision in conducting the quantitative analyses was choosing the correct dose metric. Leukemia incidence is generally analyzed as a function of cumulative exposure (e.g., ppm-years of exposure), but cumulative exposure information was not reported in most of the biomarker studies. Since a common exposure metric was needed in order to analyze the relationships among the various endpoints, chronic, time-weighted average (TWA) exposure was used. The measured value of exposure biomarkers is more likely associated with shorter term exposure than with cumulative exposure, due to clearance of metabolites from the body, the turnover rate of blood cells where the markers where observed, or other recovery mechanisms during periods with no exposure, such as weekends or other times off work. The ideal approach to characterizing exposure would be to use a dynamic model, but the exposure data were not adequately characterized to implement this approach in most cases. However, assuming that exposures are relatively constant, using chronic TWA exposures is a good surrogate for cumulative exposure.

To evaluate the Hill criteria of consistency and dose-response, the values of markers from multiple studies were plotted together to determine whether they reflected a consistent dose-response relationship. However, for many of the biomarkers, particularly biomarkers of effect, only one appropriate human study was available, limiting the ability to evaluate consistency. Graphical comparisons were used to evaluate the various candidate precursors to leukemia (Figure 2), anemia (Figure 3), and leukopenia (Figure 4). In this analysis, normalized benzene dose-response graphs for biomarkers of effect were visually evaluated for consistency in response. Several of the biomarkers considered had dose-response curves that were not consistent across studies or across dose groups within a study, and so were excluded from further validation. For clarity, these biomarkers are generally not shown in Figures 2-4. The results of these analyses for several illustrative individual biomarkers for each disease endpoint are presented in the following paragraphs.

As the next step in the decision support system, regression analyses were used to determine the association between environmental exposure measures for benzene and biomarkers of benzene exposure, as measured by the coefficient of determination[2] (r^2, Table 3). The selection of potential biomarkers based on correlation coefficients is somewhat arbitrary, but when coupled with other aspects of the Hill criteria can be informative. The regression analysis was conducted both for the entire concentration-response range observed, and for exposures only at lower concentrations in the range of regulatory or toxicological significance (i.e., up to approximately 1 ppm). Regression analysis was not performed for albumin adducts with 1,4-benzoquinone or benzene oxide, or for urinary benzenetriol because there were only 3 observations and no observations below 1 ppm except for the control values. The benzene metabolites were generally much better markers of exposure than benzene in urine, breath, or

blood (Table 3). Both urinary ttMA and SPMA correlated well with benzene exposure, consistent with the ACGIH (2001) recommendation of these metabolites as biological exposure indices for benzene exposure. However, correlation with benzene exposure at low concentrations was better for urinary ttMA than for urinary SPMA (Table 3). Correlation at low benzene exposures was poor for other metabolites, consistent with the ACGIH (2001) conclusion that these metabolites were not specific and sensitive enough in the low exposure range. Overall, only two biomarkers of exposure had acceptable correlation coefficients both overall and in the low concentration range, urinary ttMA and benzene in blood (BB). The selection of these two candidates allowed us to test the impacts of parent compound versus metabolites in the dose-response predictions. Based on these considerations, urinary ttMA and BB were included in the Bayesian network modeling described below. ACGIH (2001) considered benzene in blood to be a specific and sensitive biomarker, but noted that it is markedly affected by current or recent exposure, rather than average or cumulative exposure.

The Hill criterion of strength of association was evaluated based on visual inspection of the data and by examining the slope parameters in the Bayesian network described at the end of this section. Larger slopes in the dose-response and larger Bayesian slope parameters indicate a stronger association. The Bayesian slope parameters are included as tables A-1 and A-2 of Appendix A.

To facilitate the next steps of the analysis, the results of these initial analyses using the Hill criteria were then evaluated in greater depth using the graphical comparisons in order to identify the most appropriate biomarkers for further analysis. This qualitative analysis of the adequacy of a biomarker relies on evaluation of the shape and position of the biomarker response curve relative to the disease response curve. The biomarker curve should rise before, that is to

the left of the disease response curve, but in order to be predictive of disease, it must not plateau or level off before an increase in the disease response is apparent. Furthermore, an apparent plateau in the biomarker curve is desirable for the use of the biomarker for making quantitative disease predictions. Although the absence of a plateau in the biomarker curve does not necessarily imply that it is not predictive, it nonetheless improves the ability to use the biomarker for making quantitative disease predictions.

Based on inspection of Figure 2, the genotoxicity markers hyperdiploidy in chromosome 7, micronuclei, and the oxidative stress marker 8-OHdG in lymphocytes are the most adequate precursors to leukemia. Each of the curves for these three markers exhibited smooth dose-response curves that appear to rise above background levels slightly before an increase in leukemia incidence is observed. The micronuclei and 8-OHdG curves also have some data at higher doses where these markers appear to plateau or level off, which improves the dose-response characterization, since steep dose-response curves at high doses may not be biologically plausible, and may make predictions at high doses difficult. The hyperdiploidy and aneusomy data are somewhat limited, with only three data points, missing both a plateau and data in the range where leukemia incidence begins to increase. This lack of key data introduces uncertainty and degrades the dose-response characterization. Furthermore, aneusomy appears to begin increasing at doses higher than those that cause leukemia, and the dose-response curves for aneusomy and leukemia clearly cross, which would not be expected if aneusomy causes leukemia. The single strand breaks (SSB) and breakage at chromosome 9 had dose-response curves that were not inconsistent with leukemia, but these curves increased very sharply (and before any apparent increase in leukemia incidence in the case of SSB) and were not observed to peak or level off, limiting their utility for dose-response modeling. Based on this application of

the Hill criteria, micronuclei, hyperdiploidy, and 8-OHdG were identified as the best leukemia precursors from the available datasets for use as candidates for the next phase of the validation scheme using the Bayesian network.

Figure 3 shows the various potential precursors to anemia as measured by a decrease in red blood cell counts (RBC) or by RBC abnormalities. Red blood cell count was chosen to represent the anemia endpoint since the red blood cell abnormalities did not have a consistent dose-response. Based on inspection of the figure, the blood progenitor cells behave in a way consistent with the biological knowledge of erythropoiesis. The "colony-forming unit -- granulocyte, erythrocyte, monocyte, megakaryocyte" cells (CFU-GEMM), which differentiate into the "burst-forming unit – erythroid" cells (BFU-E) are reduced at lower doses than those that decrease BFU-E. Similarly, BFU-E cells differentiate into red blood cells, and BFU-E decreases occur at lower doses than RBC decreases. There were only two data points for each of the biomarkers related to heme synthesis (e.g., delta- aminolevulinic acid [ALA] in erythrocytes or lymphocytes). Thus, the graphical comparison of these markers could not be used to distinguish among them or to evaluate whether they are good biomarkers for anemia. However, based on biological knowledge of heme synthesis formation, changes in the heme synthesis biomarkers would be expected to affect hemoglobin levels, so ALA in erythrocytes was chosen to represent the heme synthesis biomarkers as a precursor to hemoglobin in the Bayesian network. There is a reduction in hemoglobin levels at slightly lower doses than those that cause RBC levels to fall, so hemoglobin is also confirmed as a potential precursor to anemia. Based on this analysis, the Bayesian network model included one marker from the heme synthesis pathway (ALA in erythrocytes), hemoglobin, and the progenitor cells.

Figure 4 shows the various potential precursors to leukopenia as measured by a decrease in white blood cell (WBC) counts. There are three WBC curves in the figure, representing data from three separate studies. These data are shown as separate curves to illustrate the inter-study variability in the observations of WBC. As with the analysis for anemia, the blood progenitor cells behaved in a manner consistent with the biology of hematopoiesis. In the hematopoiesis pathway, CFU-GEMM differentiates into "colony-forming units – granulocyte, macrophage" (CFU-GM), which then differentiates into white blood cells. Similarly, as dose increased, CFU-GEMM was reduced first, followed by a reduction in CFU-GM, followed by a reduction in WBC.

Bayesian network

Ideally, exposure and effect biomarkers can be linked to predict dose-response response behavior using biologically-based dose-response models that are structured based on detailed quantitative knowledge of the mode of action. In most cases, however, such a quantitative level of understanding is not available, and alternative approaches are needed. Since data were available on endpoints spanning the exposure-response continuum, with multiple plausible pathways, a powerful validation and dose response approach using Bayesian Networks was used. The Bayesian Network is described in detail in Appendix A and was constructed using MCSim (Bois et al., 2005) with supporting statistical analyses in Microsoft Excel®; the model code is presented in Appendix B. The network (Figure 5) was based on the hypothesized mode of action of hematotoxicity, the availability of adequate data from the screened pool of potential studies, and the results of the biomarker evaluation against the Hill criteria. Benzene concentration and exposure biomarkers (i.e., urinary ttMA and benzene in blood) are combined into a single node

in the schematic shown in Figure 5 for clearer presentation of the network, though separate arrows from benzene concentration and the exposure markers to the next level of early effect markers are present in the actual network. In addition, the coded network allows for benzene concentration to act directly upon the effect markers so that it is possible for the validation analysis to reject poor biomarkers of exposure.

The Bayesian network was calibrated and diagnostics were performed to choose among competing biomarkers (e.g., different markers of genotoxicity) and to identify elements of the network of potential biomarkers that contributed to the model's predictions. The network must be calibrated, or fit to the data, to find values for the parameters describing the responses of and relationships between the biomarkers and disease endpoints. Bayesian analysis using MCMC simulation was used to calibrate the network. Details of the calibration are given in Appendix A. Diagnostic evaluations included evaluation of the slope parameter in the dose-response function (the probability that slope parameter is positive, the sign and magnitude of the slope parameter, and the standard error or uncertainty in the slope parameter) and the correlation coefficient between the disease and the various biomarkers. The resulting network of biomarkers with the strongest linkages to disease, from the database of biomarkers examined, is shown in Figure 6. Figures 7 and 8 illustrate one diagnostic test that was performed. This test examines the probability that the slope from a precursor to disease is greater than zero, with a probability of 50% indicating that the parameter is most likely zero. This diagnostic indicates which biomarkers are most likely causing disease, and which potential biomarkers appear less likely to be contributing to disease when all potential biomarkers are considered simultaneously in the network. In this case, the test indicates 8-OHdG (Figure 8) is a better biomarker than micronuclei (Figure 7), because the slope parameter from 8-OHdG to leukemia is more likely to

be positive (i.e., probability of 85% vs. 61%). The standard deviation of the slope parameters relative to their means were also compared to help choose among competing biomarkers. Although the slope from 8-OHdG to leukemia did not have the greatest magnitude, it did have the smallest standard deviation relative to the mean, and was again indicated to be the best early effect biomarker for leukemia. Thus, based on the data base and decision support system used, 8-OHdG was determined to be the best precursor of leukemia amongst the biomarkers examined.

The diagnostics were less conclusive for validating biomarkers along the pathways to anemia and leukopenia. This is likely due to sparse data along the pathway to anemia (e.g., for the heme synthesis pathway) and inconsistent dose-response data for the white blood cell counts (see Figure 4). Therefore, based on the graphical comparison and knowledge of the mode of benzene toxicity, the pathways through the progenitor cells were also included in the network of validated biomarkers. 8-OHdG was also linked to CFU-GEMM in the validated biomarker network based on the biological argument that oxidative stress or genotoxicity (for which 8-OHdG may be an indicator) could lead to reduced blood progenitor cell colony formation.

Dose-Response Analyses

Dose-response analyses incorporating the validated biomarkers of exposure and effects were conducted using four approaches. First, a logistic model was used to directly evaluate the relationship between the concentration of benzene in air and the disease. This dose-response was conducted to provide a frame of reference for evaluating the impact of incorporating the biomarkers. Second, a logistic model was used to evaluate the relationship between exposure biomarkers and disease, and to derive the effective exposure biomarker value corresponding to the target level of disease (e.g., 1/1000 incidence). The benzene exposure level corresponding to

this value of the effective exposure marker was then back-calculated, using a power model regression of the exposure biomarker on benzene concentration in air. In the third approach, a logistic dose-response function was fit to the benzene concentration in air versus precursor data for the validated precursor to each disease endpoint under consideration (i.e., the biomarker of effect that immediately precedes disease in the network model), based on a precursor level that was defined as an adverse response. For this case study, an adverse response was defined as one control standard deviation from the control mean response for continuous endpoints, and 10% extra risk above the control response for dichotomous (i.e., yes/no, or incidence) endpoints, consistent with the standard approach in the absence of a consensus among the toxicological risk assessment community on the definition of an adverse response for a given endpoint based on biological criteria (EPA, 2000). The power and logistic models for the second and third approaches were fitted using maximum likelihood techniques in EXCEL, an approach that is also used in EPA's Benchmark Dose Software (EPA, 2000). In the fourth approach, a Monte Carlo analysis of the Bayesian network model was conducted using MCSim (Bois et al., 2005) to calculate the benzene concentration that, when propagated through the biomarker network, produces the target disease response rate of 1:1000.

These four approaches were used to obtain estimates of the effective concentration (EC) corresponding to a specified response. The results of these alternative biomarker-based dose-response analyses are shown in Tables 4 through 6.

Table 4 shows the dose-response analyses for leukemia. The ECs derived using the biomarker of exposure (i.e., urinary tt-MA) and the precursor to disease (8-OHdG) are within a factor of two of the reference EC from the model of benzene in air versus disease. However, the network analysis results in an effective benzene concentration (EC) that is significantly lower

then the reference EC from modeling benzene in air versus disease. The network dose-response modeling approach was also attempted using hyperdiploidy as the key precursor to leukemia. The resulting dose-response (data not shown) was much more uncertain than that obtained using 8-OHdG. This is further support of the conclusion that 8-OHdG is the preferred biomarker based on these data, although a definitive conclusion is not possible, due to the limitations on biomarkers of genotoxicity used in the model.

The estimates of the EC dose-response analyses for anemia are shown in Table 5, and for leukopenia are shown in Table 6. The ECs derived using the biomarker of exposure is again very similar to the reference ECs. However, the ECs derived using the precursor, progenitor cell colonies, are 5-fold lower than using RBC and approximately 40-fold lower than using WBC. The Bayesian network dose-response analysis was inconclusive for these endpoints, likely due to sparse data on precursor effects and inconsistent data for WBC.

DISCUSSION

In this pilot study we demonstrate the use of a biomarker decision support system based on the Hill criteria, and using a suite of analytical tools for compiling, organizing, and validating the linkages among diverse biomarkers from across the exposure-disease continuum. We demonstrate a systematic approach for comparing and choosing among exposure or effect biomarkers along the continuum, and incorporating the biomarkers in an exposure-biomarker-disease response analysis to compute biomarker-based points of departure for OEL derivation.

Benzene was used as a case study due to its importance as a relevant occupational exposure and based on the availability of potential biomarker information. We computed ECs

for three potential hematological endpoints; leukemia, anemia, and leukopenia. For each endpoint, a reference EC was estimated using the conventional approach of regressing disease response on benzene concentration in air for comparison with the biomarker-based approaches. The values we report here are not intended to correspond directly to current OELs for several reasons. We calculated ECs that are best estimates rather than lower confidence bounds, as would often be used in a full risk assessment. Additional aspects related to the application of uncertainty factors were not considered – our analyses are more akin to point of departure estimates. In addition, alternative biological models were not considered. Finally, our values arise directly from data analysis and do not consider risk policy considerations often embedded in OEL determinations. Nevertheless, the methods demonstrated here showcase the potential impact of using biomarker information to estimate points of departure in support of the OEL development process. Our results are suitable for comparisons of the relative magnitude of point of departure estimates derived with and without the use of biomarker information. In addition, using the approaches we developed here to conduct a more complete analysis of the large database for benzene, including both experimental animal data and in vitro data, could further inform current risk assessments for benzene.

For all endpoints, the precursor-based EC, which was based on directly modeling the air concentration vs. the precursor data, is smaller than the reference EC. This is the expected result for an evaluation of the concentration vs. precursor, since the precursor event often occurs at a lower dose than the disease, and cannot have a higher threshold. The air concentration vs. precursor relationship could be refined by a biologically-based definition of adversity for the precursor. Issues surrounding the identification of relevant precursor events for direct use in risk assessment, and the definition of adversity have been a hindrance in developing consensus on

risk values. Recent risk assessments, such as for perchlorate (Ginsberg and Rice, 2005; Johnston et al., 2005; Strawson et al., 2005), provide an example of the debate surrounding this issue. Our work presented here does not remove the need for clear articulation of causal proximity of a validated biomarker to the adverse effect of interest or consideration of the degree of change in the ultimate endpoint of interest that is adverse. Nevertheless, we do provide a systematic evaluation approach to determine whether a specific precursor event can be considered validated and thus becomes a candidate for this type of evaluation. The direct use of validated exposure and effect biomarkers in dose-response assessment is optimally done using physiologically-based pharmacokinetic models (for exposure biomarkers) or biologically-based dose-response models (for early effect biomarkers) (EPA, 1994; EPA, 2005).

While direct incorporation of biomarkers using biologically-based models is optimal, sufficient mechanistic understanding or data on key precursors are not usually available. We sought to develop an alternative approach using Bayesian network models that would be workable for inclusion of biomarkers in the dose response for chemicals lacking complex biological models. In particular, a goal of the use of the Bayesian network is to provide a quantitative approach for linking changes in precursors both to exposure information and to changes in disease response, thus providing a scientifically valid point of departure that incorporates precursor dose-response information without being dependent on the difficult issue of a definition of adversity for precursors. For leukemia, the network dose-response model yielded a median estimate of the EC that was approximately an order of magnitude smaller than the reference EC based on the traditional dose-response of benzene in air versus leukemia. This was apparently due to the presence of more information in the low-dose region where changes in 8-OHdG are detectable but effects on leukemia mortality are not. The network incorporates all

of the information from the exposure-response continuum, and the result suggests that incorporation of the more sensitive early effect markers (i.e., 8-OHdG) may help inform the leukemia response at lower exposures and, in this case, result in a smaller estimate of the EC.

There were several uncertainties in the assessment. The data were too limited to evaluate the impact of several other biologically-plausible precursors to leukemia in the overall model. This result itself is useful as a validation tool as it identifies key biomarkers as candidates for additional data collection. Even for 8-OHdG, only one set of adequate data in humans that linked biomarker and exposure data was available, limiting the degree to which consistency could be evaluated.

The network dose-response analyses of anemia and leukopenia were not successful. The estimates of the EC were highly uncertain, indicating that the network model was not able to predict decreases in RBC and WBC. However, this too is valuable information, suggesting that either more data collection or better dose-response methods are needed to model these toxicity pathways adequately. The failure of the model to predict leukopenia may be due to inconsistent WBC results observed in three separate studies (Lan et al., 2004; Liu et al., 1996; Rothman et al., 1996), as shown in Figure 4. Lan et al. (2004) reported much greater reductions in WBC at lower exposures than the other studies, concluding that the threshold for WBC decreases is less than 1 ppm. The Liu et al. (1996) data suggested that the threshold is actually somewhere between 1 and 30 ppm, while the Rothman et al. (1996) study indicated it is approximately 14 ppm. Combining these data in the analysis may have introduced too much inter-study variability to obtain an adequate fit to the data, and resulted in a highly uncertain estimate of the EC for WBC decreases that is much larger than 1 ppm. The RBC data reported by Rothman et al. (1996) appear to be consistent within this study, but there were only three RBC measurements

reported, and no other studies with RBC observations were entered in the database, so the analysis of RBC had little power.

Bayesian networks have been used extensively in the ecology field for evaluating causal relationships (e.g., Orme-Zavaleta et al., 2006). This tool has also received much attention recently in the field of human health toxicology for application in the identification and analysis of gene regulatory networks (Dojer et al., 2006; Imoto et al., 2003; Toyoshiba et al., 2004; Toyoshiba et al., 2006). However, to our knowledge, our application for identifying relevant linkages of biomarkers along the exposure disease continuum to support dose-response assessments has not been adequately explored. We demonstrated the utility of this approach in a limited case study. There are a variety of additional extensions of the concept that would be needed for testing the broader applicability of this method. For example, we limited the analysis to in vivo studies, studies with human data, and exposure and effect biomarkers (but not susceptibility biomarkers). It is possible that endpoints not analyzed in this case study due to inadequate human exposure-response data for validation are actually important events in the exposure-response continuum, and therefore could be good biomarkers of exposure or effects for benzene. Significant data on benzene biomarkers has been collected in animals and in mechanistic in vitro studies (e.g., Au et al., 2002; Bird et al., 2005; Faiola et al., 2004;), and therefore, there is opportunity for an enhancement to the current method related to combining data from these diverse sources. PBPK models are likely to play a role in this regards. In addition, in some cases PBPK models have been shown as an appropriate method for evaluating information on biomarkers of susceptibility (Gentry et al., 2002; Haber et al., 2002). Further work could include strategies for quantitative linkages between PBPK models to the network analyses such as demonstrated here.

In conclusion, this paper has demonstrated a systematic approach for evaluating potential biomarkers and incorporating them into dose-response assessments for risk assessment applications. Application of a biomarker decision support system proved useful in organizing the data evaluation. This decision support system embodies a suite of biomarker validation approaches. One novel approach developed in this study was the application of a Bayesian network for validation and use of biomarkers to enhance dose-response analysis. The systematic and sequential approach described here provides a scientific and mathematically sound rationale for selection of a suitable biomarker(s) for occupational safety and health applications. We demonstrate the approach with a limited case study using benzene, and extension of the methods will be needed to address additional issues likely to be encountered for diverse chemical datasets. Future work is aimed at identifying and expanding the applications of the proposed approach.

REFERENCES

ACGIH (American Conference of Governmental Industrial Hygienists). (2001). Documentation of the biological exposure indices. 7th Edition. ACGIH Worldwide.

Albertini, R. J. (1998). The use and interpretation of biomarkers of environmental genotoxicity in humans. *Biotherapy* **11,** 155-167.

Albertini, R. J. (2001). Developing sustainable studies on environmental health. *Mutat. Res.* **480-481,** 317-331.

Au, W. W., Oberheitmann, B., Harms, C. (2002). Assessing DNA damage and health risk using biomarkers. *Mutat. Res*. **509**, 153-163.

Barker, P. E. (2003). Cancer biomarker validation: standards and process: roles for the National Institute of Standards and Technology (NIST). *Ann. N. Y. Acad. Sci.* **983**,142-150.

Bird, M.G., Grerim, H., Snyder, R., Rice, J.M. (2005). International symposium: Recent advances in benzene toxicity. *Chem. Biol. Interact*. **153-154**, 1-5.

Bogadi-Sare, A., Brumen, V., Turk, R., Karacic, V., Zavalic, M. (1997). Genotoxic effects in workers exposed to benzene: with special reference to exposure biomarkers and confounding factors. *Ind. Health*. **35**, 367-373.

Bois, F., Maszle, D., Revzan, K., Tillier, S., and Yuan, Z. (2005). MCSIM: logiciel de simulation **[**MCSim Simulation Package]. Experimental Toxicology Unit, INERIS, Verneuil-en-Halatte, France. http://toxi.ineris.fr/en/activites/toxicologie_quantitative/mcsim/mcsim.php

Boogaard, P. J., and Van Sittert, N. J. (1995). Biological monitoring of exposure to benzene: a comparison between S-phenylmercapturic acid, trans,trans-muconic acid, and phenol. *Occup Environ Med*. **52**(9), 611-620.

Crump, K. and Allen, B. (1985). Methods for quantitative risk assessment using occupational studies. *Amer. Stat.* **39**, 442-450.

Committee on Biological Markers of the National Research Council. (1987). Biological markers in environmental health research. *Environ. Health Perspect.* **74**, 3-9.

Dojer, N., Gambin, A., Mizera, A., Wilczynski, B., Tiuryn, J. (2006). Applying dynamic Bayesian networks to perturbed gene expression data. *BMC Bioinformatics* **8**, 7:249.

Eastmond D. A., Schuler, M., Frantz, C., Chen, H., Parks, R., Wang, L., Hasegawa, L. (2001). Characterization and mechanisms of chromosomal alterations induced by benzene in mice and humans. HEI Research Report.

Egeghy, P. P., Tornero-Velez, R., and Rappaport, S. M. (2000). Environmental and biological monitoring of benzene during self-service automobile refueling. *Environ. Health Perspect.* **108** (12), 1195-1202.

EPA (United States Environmental Protection Agency). (1994). Methods for derivation of inhalation reference concentrations and application of inhalation dosimetry. EPA/600/8-90/066F.

EPA (United States Environmental Protection Agency). (2000). Benchmark dose technical guidance document [external review draft]. Risk Assessment Forum, Washington, DC; EPA/630/R-00/001. Available from: <http://www.epa.gov/iris/backgr-d.htm>.

EPA (United States Environmental Protection Agency). (2005). Guidelines for carcinogen risk assessment. Risk Assessment Forum, Washington, DC; EPA/630/P-03/001B.

Faiola, B., Fuller, E., Wong, V.A., Pluta, L., Abernethy, D.J., Rose, J., Recio, L. (2004). Exposure of hematopoietic stem cells to benzene or 1,4-benzoquinone induces gender-specific gene expression. *Stem Cells*. **22**, 750-758.

Forrest, M. S., Lan, Q., Hubbard, A. E., Zhang, L., Vermeulen, R., Zhao, X., Li, G., Wu, Y., Shen, M., Yin, S., Chanock, S. J., Rothman, N., and Smith, M. T. (2005). Discovery of Novel Biomarkers by Microarray Analysis of Peripheral Blood Mononuclear Cell Gene Expression in Benzene-Exposed Workers. *Environ. Health Perspect*. **113**(6), 801-807.

Fustinoni, S. Buratti, M., Campo, L., Colombi, A. Consonni, D. Pesatori, A. C., Bonzini, M., Farmer, P., Garte, S., Valerio, F., Merlo, D. F., Bertazzi, P. A. (2005). Urinary t,t-muconic acid, S-phenylmercapturic acid and benzene as biomarkers of low benzene exposure, *Chem. Biol. Interact.* **153-154**, 253-256.

Garte, S., Popov, T., Georgieva, T., Bolognesi, C., Taioli, E., Bertazzi, P., Farmer, P., and Merlo, D. (2005). Biomarkers of exposure and effect in Bulgarian petrochemical workers exposed to benzene. *Chem. Biol. Interact*. **153-154**, 247-251.

Gentry, P. R., Hack, C. E., Haber, L., Maier, A., and Clewell, H.J. 3rd. (2002). An approach for the quantitative consideration of genetic polymorphism data in chemical risk assessment: examples with warfarin and parathion. *Toxicol. Sci*. **70**(1), 120-39.

Ginsberg, G., and Rice, D. (2005). The NAS perchlorate review: questions remain about the perchlorate RfD. *Environ Health Perspect*. 2005 Sep;**113**(9), 1117-1119. Erratum in: *Environ Health Perspect*. 2005 Nov;**113**(11), A732.

Goldstein, B. D., Witz, G., Javid, J. Amoruso, M. A., Rossman, T., Wolder, B. (1981). Muconaldehyde, a potential toxic intermediate of benzene metabolism. Biological Reactive Intermediates II. Proceedings of the Second International Symposium, Chemical Mechanisms and Biological Effects, Part A, Plenum Press, New York, pages 331-339

Haber, L. T., Maier, A., Gentry, PR., Clewell, H. J., and Dourson, M.L. (2002). Genetic polymorphisms in assessing interindividual variability in delivered dose. *Regul. Toxicol. Pharmacol*. **35**(2 Pt 1), 177-97.

Hotz, P., Carbonnelle, P., Haufroid, V., Tschopp, A., Buchet, J.P., Lauwerys, R. (1997). Biological monitoring of vehicle mechanics and other workers exposed to low concentrations of benzene. *Int. Arch. Occup. Environ. Health*. **70,** 29-40.

Imoto, S., Higuchi, T., Goto, T., Tashiro, K., Kuhara, S., Miyano, S. (2003). Combining microarrays and biological knowledge for estimating gene networks via Bayesian networks. *Proc. IEEE Comput. Soc. Bioinform. Conf*. **2**, 104-13.

Javelaud, B., Vian, L., Molle, R., Allain, P., Allemand, B., Andre, B., Barbier, F., Churet, AM., Dupuis, J., Galand, M., Millet, F., Talmon, F., Touron, C., Vaissiere, M., Vechambre, D., Vieules, M., and Viver, D. (1998). Benzene exposure in car mechanics and road tanker drivers. *Int. Arch. Occup. Environ. Health.* **71**, 277-283.

Johnston, R. B. Jr, Corley, R., Cowan, L., Utiger, R. D. (2005). The NAS perchlorate review: Adverse effects? *Environ. Health. Perspect.* **113**(11), A728-9.

Kilbourne, E. D. (1973). The molecular epidemiology of influenza. *J. Infect. Dis*. **127**(4), 478-487.

Lan, Q., Zhang, L., Li, G., Vermeulen, R., Weinberg, R.S., Dosemeci, M., Rappaport, S.M., Shen, M., Alter, B.P., Wu, Y., Kopp, W., Waidyanatha, S., Rabkin, C., Guo, W., Chanock, S., Hayes, R.B., Linet, M., Kim, S., Yin, S., Rothman, N. and Smith, M.T. (2004). Hematotoxicity in workers exposed to low levels of benzene. *Science* Dec 3; **306**(5702):1774-1776.

Liu, L., Zhang, Q., Feng, J., Deng, L., Zeng, N., Yang, A. and Zhang, W. (1996). The study of DNA oxidative damage in benzene-exposed workers. *Mutat. Res.* **370**(3-4): 145-150.

Maier, A., Savage, R. E. Jr., Haber L. T.. (2004). Assessing biomarker use in risk assessment - a survey of practitioners. *J. Toxicol. Env.Health* Part A, **67**, 687-695.

Muzyka, V., Bogovski, S., Viitak, A., and Veidebaum, T. (2002). Alterations of heme metabolism in lymphocytes and metal content in blood plasma as markers of diesel fuel effects on human organism. *Sci. Total Environ.* **286**(1-3), 73-81.

Nilsson, R. I., Nordlinder, R. G., Tagesson, C., Walles, S., and Jarvholm, B.G. (1996). Genotoxic effects in workers exposed to low levels of benzene from gasoline. *Am. J. Ind. Med.* **30**(3):317-324.

Ong, C. N., Kok, P. W., Lee, B. L., Shi, C. Y., Ong, H. Y., Chia, K. S., Lee, C. S. and Luo, X. W. (1995). Evaluation of biomarkers for occupational exposure to benzene. *Occup. Environ. Med.* **52**(8), 528-533.

Ong, C. N., Kok, P. W., Ong, H. Y., Shi, C. Y., Lee, B. L., Phoon, W. H., Tan, K. T. (1996). Biomarkers of exposure to low concentrations of benzene: a field assessment. *Occup Environ Med.* **53**, 328-333.

Orme-Zavaleta, J., Jorgensen, J., D'Ambrosio, B., Altendorf, E., and Rossignol, P. A. (2006). Discovering spatio-temporal models of the spread of West Nile virus. *Risk Anal*. **26**(2),413-22.

Perera, F. P. (2000). Molecular epidemiology: on the path to prevention? *J. Natl. Cancer Inst.* **92**(8), 602-612.

Popp, W., Rauscher, D., Muller, G., Angerer, J., and Norpoth, K. (1994). Concentrations of benzene in blood and S-phenylmercapturic and t,t-muconic acid in urine in car mechanics. *Int. Arch. Occup. Environ. Health.* **66**(1), 1-6.

Qu, Q., Melikian, A., Li, G., Shore, R., Chen, L., Cohen, B., Yin, S., Kagan, M., Li, H., Meng, M., Jin, X., Winnik, W., Li, Y., Mu, R., and Li, K.. (2000). Validation of biomarkers in humans exposed to benzene: Urine metabolites. *Am J Ind Med*. **37**, 522-531.

Rappaport, S., Waidyanatha, S., Yeowell-O'Connell, K., Rothman, N., Smith, M., Zhang, L., Qu, Q., Shore, R., Li, G., and Yin, S. (2005). Protein adducts as biomarkers of human benzene metabolism. *Chem. Biol. Interact.* **153-154**, 103-109.

Rothman, N., Haas, R., Hayes, R., Li, G., Wiemels, J., Campleman, S., Quintana, P., Xi, L., Dosemeci, M., Titenko-Holland, N., Meyer, K., Lu, W., Zhang, L., Bechtold, W., Wang, Y., Kolachana, P., Yin, S., Blot, W., and Smith, M.. (1995). Benzene induces gene-duplicating but not gene-inactivating mutations at the glycophorin A locus in exposed humans. *Proc. Natl. Acad. Sci.* **92**, 4069-4073.

Rothman, N., Li, G.L., Dosemeci, M., Bechtold, W.E., Marti, G.E., Wang, Y.Z., Linet, M., Xi, L.Q., Lu, W., Smith, M.T., Titenko-Holland, N., Zhang, L.P., Blot, W., Yin, S.N. and Hayes,

R.B. (1996). Hematotoxicity among Chinese workers heavily exposed to benzene. *Am. J. Ind. Med.* **29**(3), 236-246.

Sarto, F., Cominato, I., Pinton, A. M., Brovedani, E. M., Peruzzi, M., Bianchi, V., Levis, AG. (1984). A cytogenic study on workers exposed to low concentrations of benzene. *Carcinogenesis* **5**(6), 827-832.

Schulte, P. A. (2005). The use of biomarkers in surveillance, medical screening and intervention. *Mutat. Res.* **592**(1-2), 155-163.

Sherwood, R. J. (1988). Pharmacokinetics of benzene in a human after exposure at about the permissible limit. *Ann. N. Y. Acad. Sci.* **534**(11), 635-647.

Sherwood, R. J., and Sinclair, G. C. (1999). New PBPK model applied to old occupational exposure to benzene. *AIHA J*. **60**, 259-265.

Smith, M. T. (1996). The mechanism of benzene-induced leukemia: a hypothesis and speculations on the causes of leukemia. *Environ. Health Perspect.* **104** Suppl 6,1219-25.

Stommel, P., Muller, G., Stucker, W., Verkoyen, C., Schobel, S. and Norpoth, K. (1989). Determination of S-phenylmercapturic acid in the urine- an improvement in the biological monitoring of benzene exposure. *Carcinogenisis*, **10**(2), 279-282.

Strawson, J., Dourson, M.L., Zhao, Q.J. (2005). The NAS perchlorate review: is the RfD acceptable? Environ. Health Perspect. **113**(9), 1117-1119.

Toyoshiba, H., Yamanaka, T., Sone, H., Parham, FM., Walker, NJ., Martinez, J., Portier, CJ. (2004). Gene interaction network suggests dioxin induces a significant linkage between aryl hydrocarbon receptor and retinoic acid receptor beta. *Environ. Health. Perspect.* **112**(12), 1217-1224.

Toyoshiba, H., Sone, H., Yamanaka, T., Parham, F. M., Irwin, R. D., Boorman, G. A., Portier, C. J. (2006). Gene interaction network analysis suggests differences between high and low doses of acetaminophen. *Toxicol. Appl. Pharmacol*. **215**(3), 306-316.

Verdina, A., Galati, R., Falasca, G., Ghittori, S., Imbriani, M., Tomei, F., Marcellini, L., Zijno, A., Del Vecchio, V., Crebelli, R. (2001). Metabolic polymorphisms and urinary biomarkers in subjects with low benzene exposure. *J. Toxicol. Environ. Health Part A*. **64**, 607-618.

Waidyanatha, S., Rothman, N., Fustinoni, S., Smith, M.T., Hayes, R.B., Bechtold, W., Dosemeci, M., Guilan, L., Yin, S. and Rappaport, S.M. (2001). Urinary benzene as a biomarker of exposure among occupationally exposed and unexposed subjects. *Carcinogenesis.* **22**(2), 279-286.

Wiwanitkit, V., Suwansaksri, J., Nasuan, P. (2001). Urine trans,trans-muconic acid as a biomarker for benzene exposure in gas station attendants in Bangkom, Thailand. *Ann. Clin. Lab. Sci.* **31**(4), 399-401.

Zhang, L., Eastmond, D. A., Smith, M. T. (2002). The nature of chromosomal aberrations detected in humans exposed to benzene. *Crit Rev Toxicol.* **32**(1),1-42.

Footnotes

[1] Benzene exposure and leukemia mortality data from the pliofilm cohort (Crump and Allen, 1985) were generously provided by Dr. Crump and Dr. Van Landingham at Environ. The subjects with occupational exposure durations greater than 1 year were included in the analysis.

[2] The coefficient of determination is a goodness of fit statistic from ordinary least squares regression (range 0 to 1), while the correlation coefficient is a measure of how strongly one variate tends to vary with another variate (range -1 to 1). The coefficient of determination is interpreted as the proportion of variation observed in the dependent variable (e.g., urinary metabolite level) that can be explained by the simple linear regression model. In other words, what proportion of the error is explained by the model.

Figure Legends

Figure 1. Conceptual representation of the biomarker decision support system. The sorting process is an iterative one, involving both judgments made prior to the analysis, and quantitative analysis to affect judgments.

Figure 2. Graphical analysis of precursors to leukemia incidence. Selected biomarkers of effect potentially related to leukemia, as well as the leukemia incidence, were plotted versus the concentration of benzene in air. The responses have been normalized to put them on roughly the same scale to facilitate comparisons, so the vertical axis is arbitrary. For clarity, only selected biomarkers are shown. Data are shown for individual biomarkers (e.g., aneusomy and hyperdiploidy) that are combined into one category (e.g., chromosome aberrations) in Table 2B.

Figure 3. Graphical analysis of precursors to anemia. Selected biomarkers of effect potentially related to anemia, as well as levels of red blood cells (RBC) and red blood cell abnormalities (RBC abnormalities – an effect), were plotted versus the concentration of benzene in air. The responses have been normalized to put them on roughly the same scale to facilitate comparisons, so the vertical axis is arbitrary. For clarity, only selected biomarkers are shown.

Figure 4. Graphical analysis of precursors to leukopenia. Selected biomarkers of effect potentially related to leukopenia, as well as levels of white blood cells (WBC), were plotted versus the concentration of benzene in air. Due to differences in the results of the studies of leukeopenia, separate curves are shown for each of the three WBC data sets. The responses have been normalized to put them on roughly the same scale to facilitate comparisons, so the vertical axis is arbitrary.

Figure 5. Network of candidate biomarkers of benzene exposure and effect. Open circles are biomarkers of exposure or effect, and shaded circles are disease endpoints. The structure of the network was based on the literature on benzene biology and hypothesized modes of action. ALA = delta- aminolevulinic acid; 8-OHdG = 8-hydroxyguanosine; CFU-GEMM = colony-forming unit -- granulocyte, erythrocyte, monocyte, megakaryocyte (a precursor cell type): BFU-E = burst-forming unit – erythroid; CFU-GM = colony-forming units – granulocyte, macrophage.

Figure 6. Final network of biomarkers of benzene exposure and effect with the strongest linkages to disease. Of the collection of biomarkers examined, the biomarkers shown in this figure represent those with the strongest linkage through the pathways to disease, based on analysis of the full network of candidate biomarkers. Benzene concentration is the chronic average benzene concentration in air. tt-MA = trans,trans-muconic acid; 8-OHdG = 8-hydroxyguanosine; CFU-GEMM = colony-forming unit -- granulocyte, erythrocyte, monocyte, megakaryocyte (a precursor cell type): BFU-E = burst-forming unit – erythroid; CFU-GM = colony-forming units – granulocyte, macrophage; RBC = red blood cell count; WBC = white blood cell count.

Figure 7. Histogram of the uncertainty distribution for the slope from % micronuclei to leukemia in the full Bayesian network of all candidate biomarkers. Pr (Slope > 0) = 61%, based on the posterior frequency distribution for this parameter from the MCMC calibration of the Bayesian network of potential biomarkers. The bins, or ranges of parameter values defined for the

histogram, are shown on the horizontal axis, and the number of times the parameter value falls within each bin is counted for the vertical axis values.

Figure 8. Histogram of the uncertainty distribution for the slope from 8-OHdG to leukemia in the full Bayesian network of all candidate biomarkers. Pr (Slope > 0) = 85%, based on the posterior frequency distribution for this parameter from the MCMC calibration of the Bayesian network of potential biomarkers. The bins, or ranges of parameter values defined for the histogram, are shown on the horizontal axis, and the number of times the parameter value falls within each bin is counted for the vertical axis values.

Table 1. Human Benzene Studies Selected for Biomarker Database

Population/Industry	Exposure	Molecular Endpoints (Biomarker Candidates)	Reference
49 shoe-makers and 27 healthy controls	Workers exposed an average of 17 years to benzene at a median of 6 ppm (ranging from 2 to 15 ppm) and toluene at a median of 25 ppm.	*Exposure:* blood benzene *Effects:* structural chromosome aberrations and sister chromatid exchanges in peripheral blood lymphocytes	Bogadi-Sare et al., 1997
188 European petrochemical industry workers and 52 controls	12 cohorts studied with a range of exposures from 0.01 to 211 mg/m^3	*Exposure :* urinary trans,trans-muconic acid (tt-MA) and S-phenylmercapturic acid (S-PMA)	Boogart and van Sittert, 1995
Review of 17 shale oil petrochemical plant workers and controls in Estonia, and 44 benzene-exposed workers and controls in China	Air benzene concentrations up to greater than 300 ppm, with median exposure groups of 0.3, 1, and 31 ppm.	*Exposure:* Blood benzene, urinary S-PMA, tt-MA, and phenol *Effects:* structural chromosome aberrations	Eastmond et al., 200
39 subjects exposed while refueling at self-service auto refueling stations	Benzene exposures averaged 2.9 ±5.8 mg/m^3 (median duration of 3 minutes), ranging from 0.076 to 36 mg/m^3.	*Exposure :* benzene in exhaled breath	Egeghy et al., 2000
28 workers in two shoe factories in China	15 control subjects exposed to very low levels of benzene (< 0.04 ppm), and 13 highly exposed subjects (mean 44 ppm, S.D. 24 ppm)	*Exposure:* Urinary benzene *Effects:* peripheral blood mononuclear cell gene expression	Forrest et al., 2005
415 workers including gasoline attendants, urban policemen, bus drivers, and two groups of referents in large Italian cities	Median benzene exposure was 61(gasoline attendants), 22 (policemen), 21 (bus drivers), 9 and 6 micrograms/m^3 (referents)	*Exposure :* urinary benzene, tt-MA and S-PMA	Fustinoni et al., 200
158 Bulgarian petrochemical industry workers and 50 matched controls	Air benzene concentrations averaged 1.8 ppm, with levels as high as 24 ppm among exposure workers, and an average of 0.02 ppm for controls.	*Exposure :* urinary tt-MA and S-PMA *Effects:* DNA single strand breaks (SSB) *Susceptibility:* NQO1 genotype had a significant effect on SSB	Garte et al., 2005
Healthy volunteers donated peripheral blood stem cells	Cell cultures incubated in vitro with 0.01 to 1.0 mM muconaldehyde	*Effects:*Colony counts, hemoglobin levels, glycerol hemolysis time, GSH levels	Goldstein et al., 198
410 male workers at garages, coke plants, and a by-product plant	Benzene concentrations in air ranging from non-detectable to 7 ppm	*Exposure :* muconic acid, *s*-phenylmercapturic acid, blood and breath benzene, urinary phenol and hydroquinone.	Hotz et al., 1997
66 car mechanics and 34 road-tanker drivers	Car mechanics exposed to 0.48±1.49 mg/m^3 benzene in air, and road-tanker drivers exposed to 1.88±4.18 mg/m^3	*Exposure :* urinary muconic acid *Effects:* leukocyte alkaline phosphatase activity	Javelaud et al., 199
250 shoe workers	Four exposure groups with benzene concentrations in air ranging from 0 to greater than 10 ppm	*Effects:* progenitor cell colony formation decrease *Susceptibility:*genetic variants in myeloperoxidase and NAD(P)H: quinone oxidoreductase	Lan et al., 2004
63 shoe workers and 24 car factory paint workers	Four exposure groups with benzene concentrations in air ranging from 0 to >200 mg/m^3	*Exposure :* urinary tt-MA *Effects:* 8-hydroxy-2-deoxyguanosine, micronuclei and white blood cell levels	Liu et al., 1996
22 diesel fuel distribution trade	Mean = 1.1±0.6 mg/m^3 ; Range = 0.1 to 5.1 mg/m^3	*Effects:* heme metabolism in lymphocytes and metal content in blood	Muzyka et al., 2002
33 workers occupationally exposed to benzene from gasoline	Mean = 0.13 ppm ; Range 0.003 to 0.6 ppm	*Effects:*urinary 8-hydroxy-2-deoxyguanosine, DNA SSB from leukocytes	Nilsson et al., 1996
9 car mechanics, 13 gasoline pump attendants, and 42 shoe manufacturing workers	Four exposure categories ranging from 0 to > 5 ppm	*Exposure :* blood benzene and urinary tt-MA	Ong et al., 1995
131 petroleum refinery workers in Singapore who never smoked	Exposure to a TWA benzene concentration of 0.25 ppm (range 0.01 to 3.5 ppm)	*Exposure :* blood benzene, urinary benzene and tt-MA, phenol, catechol and hydroquinone	Ong et al., 1996
26 car mechanics	Mean = 2.6 mg/m^3 ; Range = up to 13 mg/m^3	*Exposure :* blood benzene, urinary S-PMA and urinary tt-MA	Popp et al., 1994
25 glue and shoe making factory workers	Average daily personal benzene exposure of 31±26 ppm	*Exposure:* urine metabolites	Qu et al., 2000

160 workers in various Chinese industries	Two studies – 1) NCI GMA – 17.2 ppm ; 2) HEI – 3.7 ppm ; Combined – 5.7 ppm	*Exposure :* protein (serum albumin)adducts	Rappaport et al., 20
	Mean = 72.2 8 hr TWA	*Effects:* glycophorin A (GPA) gene loss mutation assay	Rothman et al., 199
55 workers in various Chinese industries	Workers exposed to benzene, but not to other solvents, at TWA of 24 (low), 44, below median or 294 (above median) mg/m^3 for an average of 6.3 years; doses were measured by personal dosimetry for 2 weeks	*Effects:* Significant increase in mean corpuscular volume and significant decreases in ALC,WBCs,RBCs,HCT and platelets in above median group; significant reductions in ALC,RBCs and platelet count in below-median group; but only reduction in ALC was significant in low-exposure group	Rothman et al., 199
22 benzene production workers	0.2 to 12.4 ppm	*Effects:* structural chromosomal aberrations, sister chromatid exchange	Sarto et al., 1984
1 experimental subject	Individual exposed to benzene in air at concentrations ranging from 6 to 100 ppm	*Exposure:*Benzene in exhaled air, and phenol and benzene in urine	Sherwood 1988
Workers loading benzene or gasoline or performing other operations on barges and sea-going vessels	Individuals exposed to benzene in air at concentrations ranging from 0.5 up to 1100 ppm	*Exposure:*Benzene in exhaled air and phenol in urine	Sherwood and Sinclair, 1999
48 benzene exposed	1 group – up to 0.15 ppm for 8 hours 1 group – up to 1.13 ppm for 12 hours	*Exposure :* S-PMA in urine	Stommel et al., 198
169 policemen	$< 10mg/m^3$	*Exposure :* urinary S-PMA urinary tt-MA *Susceptibility:*metabolic polymorphisims	Verdina et al., 2001
42 natural rubber factory workers, adhesive tape factory workers, and benzene-based paint and varnish users	$22 \leq 31$ ppm; $20 \geq 31$ ppm	*Exposure :* urinary benzene	Waidyanatha et al., 2001
30 gas station attendants and 49 healthy controls in Thailand	Stations located in the business area of Bangkok where the monitored level of air benzene was 0.76 to 4.1 ppm.	*Exposure :* urinary tt-MA	Wiwanitkit et al., 2
Review of 43 highly exposed workers and controls in China, and workers exposed to low levels of benzene	Air benzene concentrations ranged from low levels of 0.1 ppm and 1 to 2 ppm, to high levels with a median of 31 ppm.	*Effects:* chromosomal aberrations	Zhang et al., 2002

Table 2A. Candidate Biomarkers of Exposure Included in the Validation Analysis

Category	Marker	Sources	Qualitative Hill Criteria	Regression	Bayesian Network	Reason Excluded
Benzene internal dose						
	Blood benzene	Ong et al., 1995, Eastmond et al., 2001, Bogadi-Sare et al., 1997, Ong et al., 1996, Hotz et al., 1997, Popp et al., 1994, Verdina et al., 2001	-	-	X[a]	Outperformed by ttMA in Bayesian network, and does not represent metabolite in toxic pathway.
	Breath benzene	Hotz et al., 1997, Egeghy et al., 2000, Sherwood and Sinclair, 1999 Lan et al., 2004,	-	X		Poor correlation at low exposure (probably due to outlier).
	Urinary benzene	Ong et al., 1995, Waidyanatha et al., 2001, Ong et al., 1996, Qu et al., 2000, Fustinoni et al., 2005	-	X		Per L of urine (high interindividual variability), poor overall correlation.
Metabolite internal dose						
	Urinary ttMA	Ong et al., 1995, Rothman et al., 1996a, Liu et al., 1996, Ong et al., 1996, Hotz et al., 1997, Popp et al., 1994, Qu et al., 2000, Boogaard and Sittert, 1995, Javelaud et al., 1998, Fustinoni et al., 2005, Verdina et al. 2001, Wiwanitkit et al., 2001	-	-	-	Validated, best correlation with high and low exposure, best in Bayesian network.
	Urinary SPMA	Eastmond et al., 2001, Hotz et al., 1997,	-	X		Poor correlation at low exposure.

	Popp et al., 1994, Qu et al., 2000, Boogaard and Sittert, 1995, Rappaport et al., 2005, Stommel et al., 1989, Verdina et al., 2001,				
BQ-Albumin adducts	Rappaport et al., 2005	X			Only 3 data points from a single study.
BO-Albumin adducts	Rappaport et al., 2005	X			Only 3 data points from a single study.
Urinary hydroquinone	Rothman et al., 1996a, Ong et al., 1996, Hotz et al., 1997, Qu et al., 2000	-	X		Poor correlation at low exposure.
Urinary catechol	Rothman et al., 1996a, Ong et al., 1996, Hotz et al., 1997, Qu et al., 2000	-	X		Poor correlation at low exposure.
Urinary benzene triol	Qu et al., 2000	X	-		Negative dose-response, only 3 data points from single study.
Urinary phenol	Rothman et al., 1996a, Rothman et al., 1995, Ong et al., 1996, Hotz et al., 1997, Popp et al., 1994, Qu et al., 2000, Stommel et al., 1989	-	X		Poor correlation and not specific to benzene exposure at low exposure.

[a] The "X" indicates the candidate marker was excluded in that stage of the validation process.

Table 2B. Candidate Biomarkers of Effect Included in the Validation Analysis

Category	Marker	Sources	Qualitative Hill Criteria	Graphical Comparison	Bayesian Network	Reason Excluded
Oxidative Stress						
	Urinary 8-OHdG	Rothman et al., 1995	-	X[a]		Inconsistent dose-response
	Lymphocyte 8-OHdG	Liu et al., 1996	-	-	-	Validated
Toxicogenomics						
	CXCL16	Forrest et al., 2005	X			JUN expression chosen
	ZNF331	Forrest et al., 2005	X			JUN expression chosen
	JUN	Forrest et al., 2005	-	-	X	Outperformed by 8-OHdG
	PF4	Forrest et al., 2005	X			JUN expression chosen
Genotoxicity						
	N0 mutation	Rothman et al., 1995	X			only 2 observations
	NN mutation	Rothman et al., 1995	X			only 2 observations
	Chromosomal aberrations	Eastmond et al., 2001, Bogadi-Sare et al., 1997, Sarto et al., 1984, Zhang et al., 2002	-	-	X	(hyperdiploidy) outperformed by 8-OHdG
	Lymphocyte micronuclei	Liu et al., 1996	-	-	X	Outperformed by 8-OHdG
	Single DNA strand breaks	Rothman et al., 1995, Garte S et al., 2005	-	X		Response increased sharply too early
Stem cell cytotoxicity						
	Colony forming units (CFU) (GEMM)[b]	Lan et al., 2004,	-	-	-	Validated
	Colony forming units (GM)	Lan et al., 2004,	-	-	-	Validated
	Burst forming units (E)	Lan et al., 2004,	-	-	-	Validated
	LAPA (leukocyte alkaline phosphatase activity)	Javelaud et al., 1998	X			Only measured at <= 0.2 ppm
Anemia precursors						
	Hemoglobin	Lan et al., 2004, Fustinoni et al., 2005	-	-	X	Outperformed by stem cell pathway
	Protoporphyrin	Muzyka et al., 2002	-	X		ALA in erythrocytes chosen
	Protoporphyrin-DNA adducts	Muzyka et al., 2002	-	X		ALA in erythrocytes chosen

	Heme synthetase	Muzyka et al., 2002	-	X		ALA in erythrocytes chosen
	ALA synthetase	Muzyka et al., 2002	-	X		ALA in erythrocytes chosen
	ALA in erythrocytes[c]	Muzyka et al., 2002	-	-	X	Outperformed by stem cell pathway
	ALA in lymphocytes	Muzyka et al., 2002	-	X		ALA in erythrocytes chosen

[a] The "X" indicates the candidate marker was excluded in that stage of the validation process.

[b] Abbreviations: ALA = delta- aminolevulinic acid; CFU-GEMM = colony-forming unit -- granulocyte, erythrocyte, monocyte, megakaryocyte (a precursor cell type); CFU-GM = colony-forming units – granulocyte, macrophage; GSH = reduced glutathione; NBT = Nitroblue Tetrazolium; . To simplify the presentation, some of the markers shown reflect composites of multiple related markers. For example, multiple markers of chromosomal aberrations were considered, including hyperdiploidy, aneusomy, and aberrations in several individual chromosomes.

[c] Included in the Bayesian network based on biological plausibility even though failed in graphical analysis due to limited data.

Table 3. Correlation Coefficients from Linear Regression of Potential Biomarkers of Exposure

Endpoint	R^2 (all)	R^2 (< ~1 ppm)
Urinary ttMA	0.94	0.12
Blood benzene	0.37	0.25
Urinary SPMA	0.95	0.05
Urinary catechol	0.93	<0.01
Urinary hydroquinone	0.98	<0.01
Urinary phenol	0.99	0.02
Urinary Benzene	0.07	0.93
Breath benzene	0.77	0.02

Table 4. Comparison of Biomarker-based Dose-response Approaches for Acute Myeloid Leukemia (AML)

Approach	BMR[1]	EC[2] (ppm)
Benzene in Air (ppm) vs. AML	1/1000	7
Urinary ttMA (mg/g Cre) vs. AML	1/1000	4
Benzene in Air (ppm) vs. 8-OHdG	1 sd	5
Network Analysis of AML	1/1000	0.7

[1] Benchmark response, 1/1000 extra risk above background for AML or 1 control standard deviation from the control mean for 8-OHdG

[2] Effective concentration of benzene in air leading to the specified BMR

Table 5. Comparison of Biomarker-based Dose-response Approaches for Anemia (RBC)

Approach	BMR[1]	EC[2] (ppm)
Benzene in Air (ppm) vs. RBC	1 sd	100
Urinary ttMA (mg/g Cre) vs. RBC	1 sd	100
Benzene in Air (ppm) vs. BFU-E	1 sd	20
Network Analysis of RBC [3]	1 sd	-

[1] Benchmark response, 1 control standard deviation from control mean

[2] Effective concentration of benzene in air leading to the specified BMR

[3] No consistent network dose-response relationship for anemia

Table 6. Comparison of Biomarker-based Dose-response Approaches for Leukopenia (WBC)

Approach	BMR[1]	EC[2] (ppm)
Benzene in Air (ppm) vs. WBC	1 sd	300
Urinary ttMA (mg/g Cre) vs. WBC	1 sd	300
Benzene in Air (ppm) vs. CFU-GM	1 sd	8
Network Analysis of WBC [3]	1 sd	-

[1] Benchmark response, 1 control standard deviation from control mean

[2] Effective concentration of benzene leading to the specified BMR

[3] No consistent network dose-response relationship for leukopenia

APPENDIX A

Bayesian Network[1]

A Bayesian network is a set of conditional equations which can be represented in graphic form as a collection of nodes and links between the nodes (Pearl, 2000). Bayesian networks are directed acyclic graphs, meaning the links between nodes are arrows that indicate the direction of influence, and there are no feedback loops or cycles. The nodes in the graph represent variables, such as biomarkers and disease responses, and the links indicate a certain relationship between the nodes, such as a dose-response function.

The graph is a convenient model for representing the probabilistic relationships between the nodes or variables. Suppose the graph consists of n nodes. The goal of the analysis is to find the probability density of the variables in the graph, $f(x_1, \ldots, x_n)$. It is not convenient to work with this full, multivariate distribution, particularly as n gets large. However, using the multiplication rule (Tanis and Hogg, 1993), this can be transformed into the product of the probabilities of the variables (e.g., disease response) conditional on knowledge of the other variables in the graph (e.g., biomarkers):

$$f(x_1,\mathrm{K}\ ,x_n) = \prod_{i=1}^{n} f(x_i \mid x_1,\mathrm{K}\ ,x_{i-1})\,.$$

The dependencies among the variables in the network are easily interpreted from inspection of the graph. Rather than being dependent on all other variables in the network, node X_i is only directly dependent upon those nodes from which an arrow points to X_i, called the parents of X_i (P_i):

$$f(x_1,\mathrm{K}\ ,x_n) = \prod_{i=1}^{n} f(x_i \mid P_i)\,.$$

Variables that are not connected by some pathway along the arrows in the graph are independent. Variables that are connected by some pathway, but where the pathway is intercepted by one or more other nodes, are conditionally independent given knowledge of the intervening nodes (Pearl, 2000).

Figure A1 illustrates a simple biomarker network. The graph consists of the exposure concentration entry point (C), and nodes representing a biomarker of exposure (X_1), a biomarker of effect (X_2), and the disease response (X_3). The arrows indicate a dose-response linkage between pairs of nodes. Since the biomarker of effect intercepts the path from exposure to disease, this graph indicates that the disease response is conditionally independent of exposure concentration given knowledge of the biomarker of effect. Likewise, knowledge of the exposure concentration does not help inform the value of the biomarker of effect once we know the value

[1] This appendix gives only a brief description of Bayesian networks, mostly gleaned from the work of Pearl (2000).

of the biomarker of exposure. This means that we can predict the toxic response without knowing the exposure concentration if we have information on the biomarker of exposure or effect. This model is correct if the biomarker of exposure is a marker that lies in the mode of action pathway for the toxic response, such as a toxic metabolite in the blood or target tissue, or the exposure marker is directly proportional to exposure, at least at the doses relevant to human exposure scenarios. This approach provides a method to validate biomarker linkage and quantify the exposure-dose-biomarker-response relationships.

Each arrow in the network represents a dose-response function that must be calibrated with the experimental data. The response of node X_i is modeled as a function of the values of P_i, the parents of X_i, and θ_i, the model parameters. For example, the Bayesian network shown in Figure A1 would be coded in the following set of three equations representing the expected values of the three nodes.

(1) $g(x_1 \mid c, \theta_1) =$ Probability the biomarker of exposure is x_1 given exposure concentration c

(2) $g(x_2 \mid x_1, \theta_2) =$ Probability the biomarker of effect is x_2 given exposure marker x_1

(3) $g(x_3 \mid x_1, x_2, \theta_3) = g(x_3 \mid x_2, \theta_3) =$ Probability of disease is x_3 given effect marker x_2

The form of the right-hand side of the equations depends on the particular choice of mathematical function for the dose-response. Various dose-response functions can be accommodated based on the nature of the dose-response information. For example, logistic functions were used in a case study with benzene to model the expected responses for all biomarkers of effect and toxicity endpoints:

$$g(X_i \mid P_i, \theta_i) = \gamma_i \Big/ 1 + e^{-\alpha_i - \sum_j \beta_j p_{i,j}},$$

where $p_{i,j}$ is the j^{th} element of the parents of X_i, and γ_i, α_i,and β_i are parameters of the logistic model (θ_i) to be fitted. For dichotomous endpoints, such as leukemia incidence, γ was fixed at 1. The logistic function was chosen because of its shape (i.e., S-shaped with a plateau) and relatively small number of parameters; however other models could be chosen instead. The shapes of the dose-exposure marker curves could not be described with the logistic function, so a power function was used:

$$g(X_i \mid P_i, \theta_i) = (\sum_j \beta_j p_{i,j})^{\gamma_i}$$

Bayesian analysis was used to calibrate the network, that is, find values for the parameter that produce predictions matching the experimental data. This approach takes prior distributions for the parameter values and updates them based on the observed experimental data to yield posterior parameter distributions. Markov chain Monte Carlo (MCMC) simulation was used to compute the posterior parameter distribution using MCSim software (Bois et al., 2005), which implements the Metropolis-Hastings algorithm (Gilks et al., 1996). Several examples of similar applications of MCMC simulation for calibrating complex, nonlinear models, (i.e., physiologically-based pharmacokinetic models) have been published recently (Bernillon and Bois, 2000; Bois, 2000 a,b; Gelman et al., 1996; Hack et al., 2006; Marino et al., 2006). A large

number of parameters were used in the network model (e.g., three parameters for each arrow between nodes for a logistic dose-response model), and the available data on some biomarkers of effects were sparse, but the MCMC simulation approach is a robust calibration which can be successful despite these limitations.

Prior parameter distributions for the dose-response model for each node are defined based on visual inspection of plots of the experimental data. Standard deviations for the model parameter distributions are typically set to indicate significant uncertainty about the value. The model is calibrated using both normal and log-normal prior distributions for the slope parameters of the dose-response models. The log-normal distributions are used to restrict the parameters to be greater than zero, which was consistent with prior knowledge of the biology. Geometric standard deviations corresponding to a 100% coefficient of variation are typically used to indicate significant uncertainty about the value of the slope parameter before analyzing the data. Normal distributions for the slope parameters are also used to allow the evaluation of certain validation diagnostics (i.e., probability that the slope is greater than zero). The normal distributions for the slope parameters are assigned uninformative distributions, since there is usually little prior information about the functional dependencies between the biomarkers.

MCMC simulation is an iterative computational method in which the parameter distribution begins at the prior distribution and eventually converges to the posterior distribution. Convergence to the posterior distribution is monitored by propagating 3 independent Markov chains (i.e., with different starting values for the simulations) and evaluating whether they converge to the same distribution by visual inspection of the graph of the chains and by computing the potential scale reduction (r) convergence statistic (Gelman, 1996). The potential scale reduction is the ratio of an upper bound estimate and a lower bound estimate of the variance of the parameter distribution in two or more chains. As the independent chains converge to the same distribution, r declines to unity. In practice, Gelman (1996) recommends an r of 1.2 or less as a criterion for convergence.

The model file defining the Bayesian network model is shown in Appendix B. The model is in MCSim (Bois et al., 2005) format. It includes a legend for the nodes named in the model code, the specification of output parameters, an initialization section, and a section for calculating the results of the dose-response network model. The prior distribution specifications and experimental data are defined in a separate simulation file that is used as input to the model. An example simulation file in MCSim format is shown in Appendix C.

Biomarker network-based Dose-response Model

If a Bayesian network model is implemented, the network of biomarkers may be used to analyze the dose-response using Monte Carlo simulation with the posterior parameter distributions of the validated network. A uniform distribution of exposure concentrations is defined, and concentrations and dose-response parameter values are iteratively sampled from the distributions and used as input to the network model to generate a distribution of disease responses. The output of this analysis is examined to identify the range of exposure concentrations leading to the target disease response.

Figure A1.
Simple Bayesian network with concentration (C), a biomarker of exposure (X_1), a biomarker of effect or precursor (X_2), and disease response (X_3).

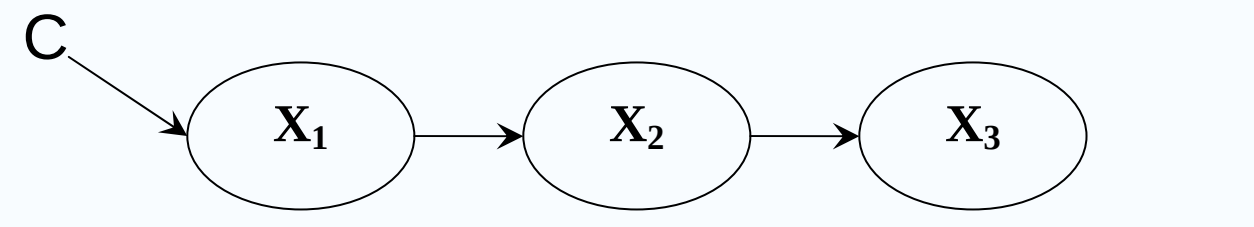

Table A-1. Bayesian slope parameters, starting with uninformative normal distributions.

From	To	Average	Standard deviation	Pr > 0
Conc	Blood benzene	3.8	4	0.83
Conc	ttMA	1.6	0.86	0.97
Blood benzene	ttMA	0.5	0.22	0.99
Conc	8-OHdG	**0.8**	**0.68**	**0.88**
Blood benzene	8-OHdG	-0.28	0.92	0.38
ttMA	8-OHdG	0.29	0.72	0.66
Conc	Micronuclei	**0.68**	**0.73**	**0.82**
Blood benzene	Micronuclei	-0.017	0.97	0.49
ttMA	Micronuclei	0.35	0.91	0.65
8-OHdG	Micronuclei	0.53	0.67	0.79
Conc	Hyperdiploidy	**0.66**	**0.74**	**0.81**
Blood benzene	Hyperdiploidy	0.058	1	0.52
ttMA	Hyperdiploidy	0.38	0.95	0.66
8-OHdG	Hyperdiploidy	0.61	0.75	0.79
Conc	CFU-GEMM	-0.26	0.86	0.38
Blood benzene	CFU-GEMM	**-0.11**	**0.99**	**0.46**
ttMA	CFU-GEMM	-0.23	0.96	0.41
8-OHdG	CFU-GEMM	-0.62	0.81	0.22
Micronuclei	CFU-GEMM	-0.4	0.92	0.33
Hyperdiploidy	CFU-GEMM	**-0.089**	**0.99**	**0.46**
CFU-GEMM	BFU-E	**0.62**	**0.91**	**0.75**
CFU-GEMM	CFU-GM	**0.43**	**0.91**	**0.68**
Conc	ALA	0.02	0.99	0.51
Blood benzene	ALA	**0.18**	**0.95**	**0.58**
ttMA	ALA	0.12	0.99	0.55
ALA	Hemoglobin	**-0.63**	**0.69**	**0.18**
CFU-GEMM	RBC	**0.68**	**0.8**	**0.80**
Hemoglobin	RBC	0.64	0.84	0.78
CFU-GM	WBC	**0.72**	**0.82**	**0.81**
Conc	JUN	-0.51	0.8	0.26
Blood benzene	JUN	**-0.065**	**1**	**0.47**
ttMA	JUN	-0.33	0.95	0.36
8-OHdG	JUN	-0.73	0.76	0.17
8-OHdG	Leukemia	**0.18**	**0.24**	**0.77**
Micronuclei	Leukemia	0.2	0.86	0.59
Hyperdiploidy	Leukemia	-0.0066	0.98	0.50
JUN	Leukemia	0.51	0.78	0.74

Bold text: Slope parameter that appeared to have the largest and most precise influence on the subsequent marker or disease endpoint in the pathway. Larger Pr>0 is better.

Table A-2. Bayesian slope parameters, starting with uninformative lognormal distributions.

From	To	Geometric Mean	Geometric standard deviation	Coefficient of variation
Conc	Blood benzene	2.8	2.3	1.01
Conc	ttMA	1.6	1.6	0.53
Blood benzene	ttMA	0.47	1.5	0.46
Conc	8-OHdG	0.15	2.2	0.96
Blood benzene	8-OHdG	0.11	2.3	1.00
ttMA	8-OHdG	**0.13**	**2.2**	**0.91**
Conc	Micronuclei	0.12	2.3	0.97
Blood benzene	Micronuclei	0.11	2.3	1.04
ttMA	Micronuclei	**0.12**	**2.3**	**1.00**
8-OHdG	Micronuclei	0.11	2.4	1.04
Conc	Hyperdiploidy	0.11	2.2	0.93
Blood benzene	Hyperdiploidy	**0.11**	**2.3**	**0.97**
ttMA	Hyperdiploidy	0.11	2.3	1.00
8-OHdG	Hyperdiploidy	0.12	2.3	1.02
Conc	CFU-GEMM	0.071	1.8	0.65
Blood benzene	CFU-GEMM	0.11	2.2	0.95
ttMA	CFU-GEMM	**0.085**	**1.9**	**0.72**
8-OHdG	CFU-GEMM	**0.079**	**1.9**	**0.68**
Micronuclei	CFU-GEMM	0.099	2	0.81
Hyperdiploidy	CFU-GEMM	0.11	2.3	0.97
CFU-GEMM	BFU-E	0.13	2.4	1.07
CFU-GEMM	CFU-GM	0.13	2.2	0.96
Conc	ALA	0.11	2.5	1.11
Blood benzene	ALA	**0.12**	**2.3**	**1.01**
ttMA	ALA	0.11	2.4	1.09
ALA	Hemoglobin	0.084	2	0.76
CFU-GEMM	RBC	**0.15**	**2.3**	**0.98**
Hemoglobin	RBC	**0.14**	**2.2**	**0.94**
CFU-GM	WBC	0.13	2.4	1.07
Conc	JUN	0.066	1.8	0.65
Blood benzene	JUN	0.11	2.2	0.95
ttMA	JUN	**0.079**	**1.9**	**0.72**
8-OHdG	JUN	0.073	1.9	0.72
8-OHdG	Leukemia	**0.061**	**1.8**	**0.65**
Micronuclei	Leukemia	0.11	2	0.78
Hyperdiploidy	Leukemia	0.11	2.3	1.03
JUN	Leukemia	**0.098**	**1.9**	**0.70**

Bold text: Slope parameter that appeared to have the largest and most precise influence on the subsequent marker or disease endpoint in the pathway.
Smaller coefficient of variation is better.

References

Bois F, Maszle D, Revzan K, Tillier S, and Yuan Z. 2005. MCSIM: logiciel de simulation **[**MCSim Simulation Package]. Experimental Toxicology Unit, INERIS, Verneuil-en-Halatte, France. http://toxi.ineris.fr/en/activites/toxicologie_quantitative/mcsim/mcsim.php

Gelman A. 1996. Inference and monitoring convergence. In : Gilks WR, Richardson S, Spiegelhalter DJ (Eds), Markov chain monte carlo in practice. Chapman & Hall, CRC Press Inc., Boca Raton, FL, pp 131-143.

Gelman A, Bois F, and Jiang J. 1996. Physiological pharmacokinetic analysis using population modeling and informative prior distributions. J. Am. Stat. Assoc. 91, 1400-1412.

Gilks WR, Richardson S, Spiegelhalter DJ. 1996. Markov chain monte carlo in practice. Chapman & Hall, CRC Press Inc., Boca Raton, FL.

Hack, C. Eric. 2006. Bayesian analysis of physiologically based toxicokinetic and toxicodynamic models. Toxicology 221(2-3):241-248.

Hogg RV and Tanis EA. 1993. Probability and statistical inference, 4th ed. MacMillan Publishing Company, New York.

APPENDIX B. Model Code

```
#Model file for benzene biomarker Bayesian network

# Legend: parameter name = biomarker name        parents = [list of parent nodes]

# External concentration in air
# y1 = Conc

# Internal dose metrics
# y2 = Blood BZ              p2 = y1
# y3 = Urine ttMA            p3 = y1-2
# y4 = BO-Alb                p4 = y1-2       (omitted)
# y5 = 1,4BQ-Alb             p5 = y1-2       (omitted)

# Oxidative stress
# y6 = 8OHdG                 p6 = y1-5
# y7 = Urine 8OHdG           p7 = y1-6       (omitted)

# Genotoxicity
# y8 = Micronuclei           p8 = y1-7
# y9 = SSB                   p9 = y1-7       (omitted)
# y10 = Monosomy             p10 = y1-7      (omitted)
# y11 = Trisomy              p11 = y1-7      (omitted)
# y12 = Breakage             p12 = y1-7      (omitted)
# y13 = Hyperdiploidy        p13 = y1-7
# y14 = SCA                  p14 = y1-7      (omitted)

# Stem cell cytotoxicity
# y15 = CFUGEMM              p15 = y1-14
# y16 = BFUE                 p16 = y15
# y17 = CFUGM                p17 = y15

# Heme synthesis
# y18 = ALAs in lymphocytes      p18 = y1-5      (omitted)
# y19 = ALA in erythrocytes      p19 = y18       since y18 was omitted, y19 now inherits the
parents of 18, y1-5
# y20 = Hemes in lymphocytes     p20 = y1-5      (omitted)
# y21 = PP in lymphocytes        p21 = y20       (omitted)
# y22 = PPDNA in lymphocytes     p22 = y21       (omitted)
# y23 = Hemoglobin               p23 = y20       since y20 was omitted, y23 now get parent
ALA, y19

#Anemia
# y24 = RBC                  p24 = y16,23
# y25 = RBCa                 (omitted)
```

```
#Leukopenia
# y26 = WBC                 p26 = y17
# y27 = WBCa                (omitted)

#Gene expression changes
# y28 = JUN                 p28 = y1-7

#Leukemia
# y29 = Leukemia            p29 = y6-14, 28


#Specify outputs
Outputs = {ln02,ln03,ln06,ln08,ln13,ln15,ln16,ln17,ln19,ln23,ln24,ln26,ln28,p29,n29,y29};

# Initialize parameters
Conc = 0;
            b02 = 1; b0102=0;
            b03 = 1; b0103=0; b0203=0;
max04 = 0; b04 = 0; b0104=0; b0204=0;
max05 = 0; b05 = 0; b0105=0; b0205=0;
max06 = 0; b06 = 0; b0106=0; b0206=0; b0306=0; b0406=0; b0506=0;
max07 = 0; b07 = 0; b0107=0; b0207=0; b0307=0; b0407=0; b0507=0; b0607=0;
max08 = 0; b08 = 0; b0108=0; b0208=0; b0308=0; b0408=0; b0508=0; b0608=0; b0708=0;
max09 = 0; b09 = 0; b0109=0; b0209=0; b0309=0; b0409=0; b0509=0; b0609=0; b0709=0;
max10 = 0; b10 = 0; b0110=0; b0210=0; b0310=0; b0410=0; b0510=0; b0610=0; b0710=0;
max11 = 0; b11 = 0; b0111=0; b0211=0; b0311=0; b0411=0; b0511=0; b0611=0; b0711=0;
max12 = 0; b12 = 0; b0112=0; b0212=0; b0312=0; b0412=0; b0512=0; b0612=0; b0712=0;
max13 = 0; b13 = 0; b0113=0; b0213=0; b0313=0; b0413=0; b0513=0; b0613=0; b0713=0;
max14 = 0; b14 = 0; b0114=0; b0214=0; b0314=0; b0414=0; b0514=0; b0614=0; b0714=0;
max15 = 0; b15 = 0; b0115=0; b0215=0; b0315=0; b0415=0; b0515=0; b0615=0; b0715=0;
                    b0815=0; b0915=0; b1015=0; b1115=0; b1215=0; b1315=0; b1415=0;
max16 = 0; b16 = 0; b1516=0;
max17 = 0; b17 = 0; b1517=0;
max18 = 0; b18 = 0; b0118=0; b0218=0; b0318=0; b0418=0; b0518=0;
max19 = 0; b19 = 0; b0119=0; b0219=0; b0319=0; b0419=0; b0519=0;
max20 = 0; b20 = 0; b0120=0; b0220=0; b0320=0; b0420=0; b0520=0;
max21 = 0; b21 = 0; b2021=0;
max22 = 0; b22 = 0; b2122=0;
max23 = 0; b23 = 0; b1923=0;
max24 = 0; b24 = 0; b1624=0; b2324=0;
            b25 = 0;
max26 = 0; b26 = 0; b1726=0;
            b27 = 0;
max28 = 0; b28 = 0; b0128=0; b0228=0; b0328=0; b0428=0; b0528=0; b0628=0; b0728=0;
            b29 = 0; b0629=0; b0729=0; b0829=0; b0929=0; b1029=0; b1129=0; b1229=0;
```

```
                    b1329=0; b1429=0; b2829=0;

CalcOutputs { #Calculate outputs
# Logistic model, Yi = maxi / (1 + exp[-(b0 + bji*pij)]),
# where maxi is the max yi, b0 is the background, bji are j slope parameters,
# and pij is the set of j parents of node i.
# A power model is used for urinary ttMA, ttMA = (B1*Conc + B2*BB)**B3
# Nodes are removed by removing their distributions in the simulation file.
# If a node's only parent is removed from the network, the child's "grandparents" must be
rerouted
# to the child and bypass the removed parent.
y02 = pow((b0102*Conc),b02);
        # Blood benzene
y03 = pow((b0103*Conc + b0203*y02),b03);
        # Urine ttMA (power model)
y04 = max04/(1.0+exp(-(b04 + b0104*Conc + b0204*y02)));
        # BO-Albumin adducts
y05 = max05/(1.0+exp(-(b05 + b0105*Conc + b0205*y02)));
        # 1,4BQ-Albumin adducts
y06 = max06/(1.0+exp(-(b06 + b0106*Conc + b0206*y02 + b0306*y03 + b0406*y04 +
b0506*y05))); # 8-OHdG in lymphocytes
y07 = max07/(1.0+exp(-(b07 + b0107*Conc + b0207*y02 + b0307*y03 + b0407*y04 +
b0507*y05 +
                        b0607*y06)));
        # 8-OHdG in urine
y08 = max08/(1.0+exp(-(b08 + b0108*Conc + b0208*y02 + b0308*y03 + b0408*y04 +
b0508*y05 +
                        b0608*y06 + b0708*y07)));
        # MN in lymphocytes
y09 = max09/(1.0+exp(-(b09 + b0109*Conc + b0209*y02 + b0309*y03 + b0409*y04 +
b0509*y05 +
                        b0609*y06 + b0709*y07)));
        # SSB
y10 = max10/(1.0+exp(-(b10 + b0110*Conc + b0210*y02 + b0310*y03 + b0410*y04 +
b0510*y05 +
                        b0610*y06 + b0710*y07)));
        # Monosomy
y11 = max11/(1.0+exp(-(b11 + b0111*Conc + b0211*y02 + b0311*y03 + b0411*y04 +
b0511*y05 +
                        b0611*y06 + b0711*y07)));
        # Trisomy
y12 = max12/(1.0+exp(-(b12 + b0112*Conc + b0212*y02 + b0312*y03 + b0412*y04 +
b0512*y05 +
                        b0612*y06 + b0712*y07)));
        # Breakage
```

```
y13 = max13/(1.0+exp(-(b13 + b0113*Conc + b0213*y02 + b0313*y03 + b0413*y04 +
b0513*y05 +
                        b0613*y06 + b0713*y07)));
        # Hyperdiploidy
y14 = max14/(1.0+exp(-(b14 + b0114*Conc + b0214*y02 + b0314*y03 + b0414*y04 +
b0514*y05 +
                        b0614*y06 + b0714*y07)));
        # SCA
y15 = max15/(1.0+exp(-(b15 - b0115*Conc - b0215*y02 - b0315*y03 - b0415*y04 - b0515*y05
-
                        b0615*y06 - b0715*y07 - b0815*y08 - b0915*y09 - b1015*y10 -
                        b1115*y11 - b1215*y12 - b1315*y13 - b1415*y14)));
        # CFU-GEMM
y16 = max16/(1.0+exp(-(b16 + b1516*y15)));
        # BFU-E
y17 = max17/(1.0+exp(-(b17 + b1517*y15)));
        # CFU-GM
y18 = max18/(1.0+exp(-(b18 + b0118*Conc + b0218*y02 + b0318*y03 + b0418*y04 +
b0518*y05))); # ALAs
#y19 = max19/(1.0+exp(-(b19 + b1918*y18)));
        # ALA
y19 = max19/(1.0+exp(-(b19 + b0119*Conc + b0219*y02 + b0319*y03 + b0419*y04 +
b0519*y05))); # ALAs was removed
y20 = max20/(1.0+exp(-(b20 + b0120*Conc + b0220*y02 + b0320*y03 + b0420*y04 +
b0520*y05))); # Hemes
y21 = max21/(1.0+exp(-(b21 + b2021*y20)));
        # PP
y22 = max22/(1.0+exp(-(b22 + b2122*y21)));
        # PP-DNA adducts
#y23 = max23/(1.0+exp(-(b23 + b2023*y20)));
        # Hemoglobin
y23 = max23/(1.0+exp(-(b23 - b1923*y19)));                      # omitted PP so ALA
not points to hemoglobin
y24 = max24/(1.0+exp(-(b24 + b1624*y16 + b2324*y23)));
        # RBC count
p25 =   1.0/(1.0+exp(-(b25)));
        # RBC abnormality
y26 = max26/(1.0+exp(-(b26 + b1726*y17)));
        # WBC count
p27 =   1.0/(1.0+exp(-(b27)));
        # WBC abnormality
y28 = max28/(1.0+exp(-(b28 - b0128*Conc - b0228*y02 - b0328*y03 - b0428*y04 - b0528*y05
-
                        b0628*y06 - b0728*y07)));
        # JUN expression
```

```
p29 =   1.0/(1.0+exp(-(b29 + b0629*y06 + b0729*y07 + b0829*y08 + b0929*y09 + b1029*y10
+
                                b1129*y11 + b1229*y12 + b1329*y13 + b1429*y14 - b2829*y28)));
        # Leukemia

# Log-transformed outputs
# Note log(x) is the natural log (common log is log10(x))
ln02 = log(y02); ln03 = log(y03); ln06 = log(y06); ln08 = log(y08); ln13 = log(y13);
ln15 = log(y15); ln16 = log(y16); ln17 = log(y17); ln19 = log(y19);
ln23 = log(y23); ln24 = log(y24); ln26 = log(y26); ln28 = log(y28);

# Dummy statements for binomially distributed data. Overwritten by data in simulation file.
y29 = 0;
n29 = 0;
}
```

APPENDIX C
Simulation File

```
# Verified network including only the best biomarkers

# Run 3 independent chains, the last argument is the random seed
#MCMC("Chain_1.out", "", "", 30000, 0, 5, 30000, 3245211);
#MCMC("Chain_2.out", "", "", 30000, 0, 5, 30000, 2394879);
MCMC("Chain_3.out", "", "", 30000, 0, 5, 30000, 3457698);

# Syntax for distributions:
# Lognormal: Distrib(parameter, Lognormal, geometric mean, geometric sd)
# Normal   : Distrib(parameter, Normal, mean, sd)
# Binomial : Likelihood(Data, Binomial, Probability, Number of trials)
# A geometric standard deviation of 2.3 corresponds to a CV of 100%

Level{
# Internal dose metrics
# y03= Urine ttMA     p3 = y1-2
Distrib(b03  , LogNormal, 0.5, 2.3);          # This is the power term
Distrib(b0103, LogNormal, 1, 2.3);

# Oxidative stress
# y06= 8-OHdG                 p6 = y1-5
Distrib(max06, LogNormal, 30, 1.6);          #Almost plateau, so use 50% CV
Distrib(b06  , Normal, -2, 5);
Distrib(b0306, LogNormal, 0.5, 2.3);

# Stem cell cytotoxicity
# y15 = CFUGEMM            p15 = y1-14
Distrib(max15, LogNormal, 3, 1.28);           #Plateau, use only 25% CV
Distrib(b15  , Normal, 2, 5);
Distrib(b0315, LogNormal, 0.5, 2.3);
Distrib(b0615, LogNormal, 0.5, 2.3);

# y16 = BFUE          p16 = y15
Distrib(max16, LogNormal,  20, 1.28);
Distrib(b16  , Normal, -5, 5);
Distrib(b1516, LogNormal, 0.5, 2.3);

# y17 = CFUGM                 p17 = y15
Distrib(max17, LogNormal, 7, 1.6);
Distrib(b17  , Normal,-2, 5);
Distrib(b1517, LogNormal, 0.5, 2.3);

#Anemia
```

```
# y24 = RBC             p24 = y16,23
Distrib(max24, LogNormal, 5, 1.28);
Distrib(b24  , Normal,-5, 5);
Distrib(b1624, LogNormal, 0.5, 2.3);

#Leukopenia
# y26 = WBC            p26 = y17
Distrib(max26, LogNormal, 7, 1.28);          # More certain of max
Distrib(b26  , Normal,-3, 5);
Distrib(b1726, LogNormal, 0.5, 2.3);

#Leukemia
# y29 = Leukemia      p29 = y6-14, 28
Distrib(b29  , Normal, -5, 5);
Distrib(b0629, LogNormal, 0.5, 2.3);

# 0.47 corresponds to a 50% prior CV for data and model error
Likelihood(Data(ln03), Normal, Prediction(ln03), 0.47);
Likelihood(Data(ln06), Normal, Prediction(ln06), 0.47);
Likelihood(Data(ln15), Normal, Prediction(ln15), 0.47);
Likelihood(Data(ln16), Normal, Prediction(ln16), 0.47);
Likelihood(Data(ln17), Normal, Prediction(ln17), 0.47);
Likelihood(Data(ln24), Normal, Prediction(ln24), 0.47);
Likelihood(Data(ln26), Normal, Prediction(ln26), 0.47);
Likelihood(Data(y29), Binomial, Prediction(p29), Data(n29));

Level{
# Use -1 for missing data

 Simulation { #1
 Conc = 3.13776E-06;
 Print(ln03,ln06,1);
 Print(ln15,ln16,ln17,ln24,1);
 Print(ln26,p29,1);
 Data ( ln03 , -1.322005871 );        # Urinary TTMA
 Data ( ln06 , 1.319085611 );         # lymphocyte 8-OHdG
 Data ( ln15 , -1 );                  # CFU-GEMM +EPO
 Data ( ln16 , -1 );                  # BFU-E +EPO
 Data ( ln17 , -1 );                  # CFU-GM +EPO
 Data ( ln24 , -1 );                  # RBC
 Data ( ln26 , 1.943048917 );         # WBC
 Data ( y29 , 0 );                    # Leukemia
 Data ( n29 , 13 ); #
 }

 Simulation { #2
```

```
Conc = 0.044858635;
Print(ln03,ln06,1);
Print(ln15,ln16,ln17,ln24,1);
Print(ln26,p29,1);
Data ( ln03 , -0.985793938 );        # Urinary TTMA
Data ( ln06 , -1 );                  # lymphocyte 8-OHdG
Data ( ln15 , 1.252762968 );         # CFU-GEMM +EPO
Data ( ln16 , 2.995732274 );         # BFU-E +EPO
Data ( ln17 , 1.791759469 );         # CFU-GM +EPO
Data ( ln24 , 1.547562509 );         # RBC
Data ( ln26 , 1.893111963 );         # WBC
Data ( y29 , 0 );                    # Leukemia
Data ( n29 , 27 ); #
}

Simulation { #3
Conc = 0.79550475;
Print(ln03,ln06,1);
Print(ln15,ln16,ln17,ln24,1);
Print(ln26,p29,1);
Data ( ln03 , -0.212780764 );        # Urinary ttMA
Data ( ln06 , 1.541159072 );         # lymphocyte 8-OHdG
Data ( ln15 , -1 );                  # CFU-GEMM +EPO
Data ( ln16 , -1 );                  # BFU-E +EPO
Data ( ln17 , -1 );                  # CFU-GM +EPO
Data ( ln24 , -1 );                  # RBC
Data ( ln26 , 1.834180185 );         # WBC
Data ( y29 , 0 );                    # Leukemia
Data ( n29 , 137 ); #
}

Simulation { #4
Conc = 2.0946499;
Print(ln03,ln06,1);
Print(ln15,ln16,ln17,ln24,1);
Print(ln26,p29,1);
Data ( ln03 , 1.523880024 );         # Urinary TTMA
Data ( ln06 , -1 );                  # lymphocyte 8-OHdG
Data ( ln15 , 0.916290732 );         # CFU-GEMM +EPO
Data ( ln16 , 2.995732274 );         # BFU-E +EPO
Data ( ln17 , 1.609437912 );         # CFU-GM +EPO
Data ( ln24 , -1 );                  # RBC
Data ( ln26 , 1.733423892 );         # WBC
Data ( y29 , 1 );                    # Leukemia
Data ( n29 , 76 ); #
}
```

```
Simulation { #5
Conc = 5.604657396;
Print(ln03,ln06,1);
Print(ln15,ln16,ln17,ln24,1);
Print(ln26,p29,1);
Data ( ln03 , 2.200552367 );        # Urinary ttMA
Data ( ln06 , -1 );                 # lymphocyte 8-OHdG
Data ( ln15 , -1 );                 # CFU-GEMM +EPO
Data ( ln16 , -1 );                 # BFU-E +EPO
Data ( ln17 , -1 );                 # CFU-GM +EPO
Data ( ln24 , -1 );                 # RBC
Data ( ln26 , -1 );                 # WBC
Data ( y29 , 3 );                   # Leukemia
Data ( n29 , 186 ); #
}

Simulation { #6
Conc = 14.17311154;
Print(ln03,ln06,1);
Print(ln15,ln16,ln17,ln24,1);
Print(ln26,p29,1);
Data ( ln03 , 2.344206768 );        # Urinary ttMA
Data ( ln06 , -1 );                 # lymphocyte 8-OHdG
Data ( ln15 , -1 );                 # CFU-GEMM +EPO
Data ( ln16 , -1 );                 # BFU-E +EPO
Data ( ln17 , -1 );                 # CFU-GM +EPO
Data ( ln24 , 1.526056303 );        # RBC
Data ( ln26 , 1.85629799 );         # WBC
Data ( y29 , 3 );                   # Leukemia
Data ( n29 , 182 ); #
}

Simulation { #7
Conc = 32.43104135;
Print(ln03,ln06,1);
Print(ln15,ln16,ln17,ln24,1);
Print(ln26,p29,1);
Data ( ln03 , 2.7047113 );          # Urinary TTMA
Data ( ln06 , 3.262701304 );        # lymphocyte 8-OHdG
Data ( ln15 , 0 );                  # CFU-GEMM +EPO
Data ( ln16 , 2.302585093 );        # BFU-E +EPO
Data ( ln17 , 1.098612289 );        # CFU-GM +EPO
Data ( ln24 , -1 );                 # RBC
Data ( ln26 , 1.665818246 );        # WBC
Data ( y29 , 1 );                   # Leukemia
```

```
 Data ( n29 , 115 ); #
 }

 Simulation { #8
 Conc = 109.3264178;
 Print(ln03,ln06,1);
 Print(ln15,ln16,ln17,ln24,1);
 Print(ln26,p29,1);
 Data ( ln03 , 3.973400285 );        # Urinary TTMA
 Data ( ln06 , 3.397523976 );        # lymphocyte 8-OHdG
 Data ( ln15 , -1 );                 # CFU-GEMM +EPO
 Data ( ln16 , -1 );                 # BFU-E +EPO
 Data ( ln17 , -1 );                 # CFU-GM +EPO
 Data ( ln24 , 1.435084525 );        # RBC
 Data ( ln26 , 1.776645831 );        # WBC
 Data ( y29 , 4 );                   # Leukemia
 Data ( n29 , 91 ); #
 }

 }
}
END.
```

Figure 1. Conceptual Representation of the Biomarker Decision Support System

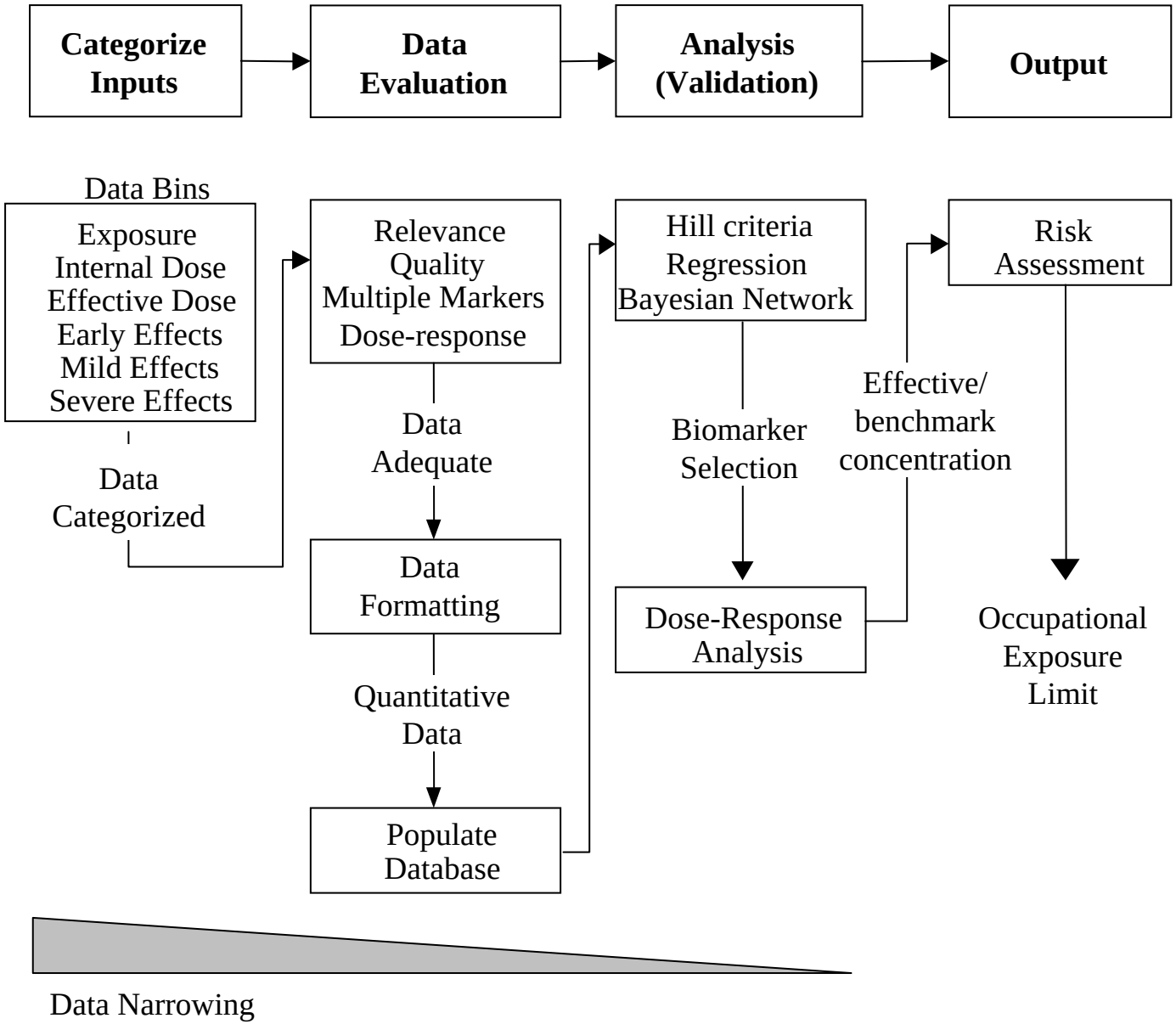

Figure 2. Graphical Analysis of Precursors of Leukemia Incidence

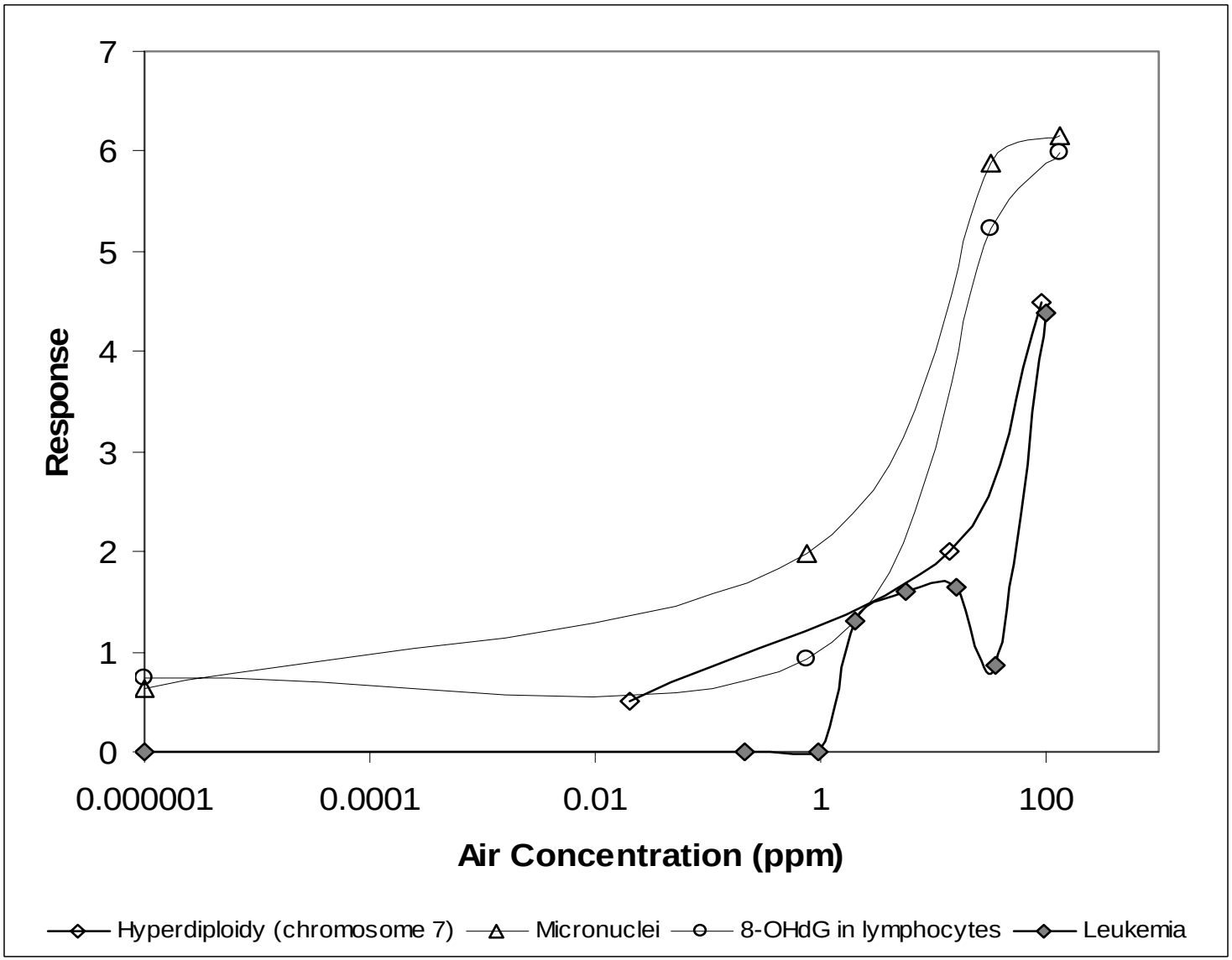

Figure 3. Graphical Comparison of Precursors of Anemia (RBC)

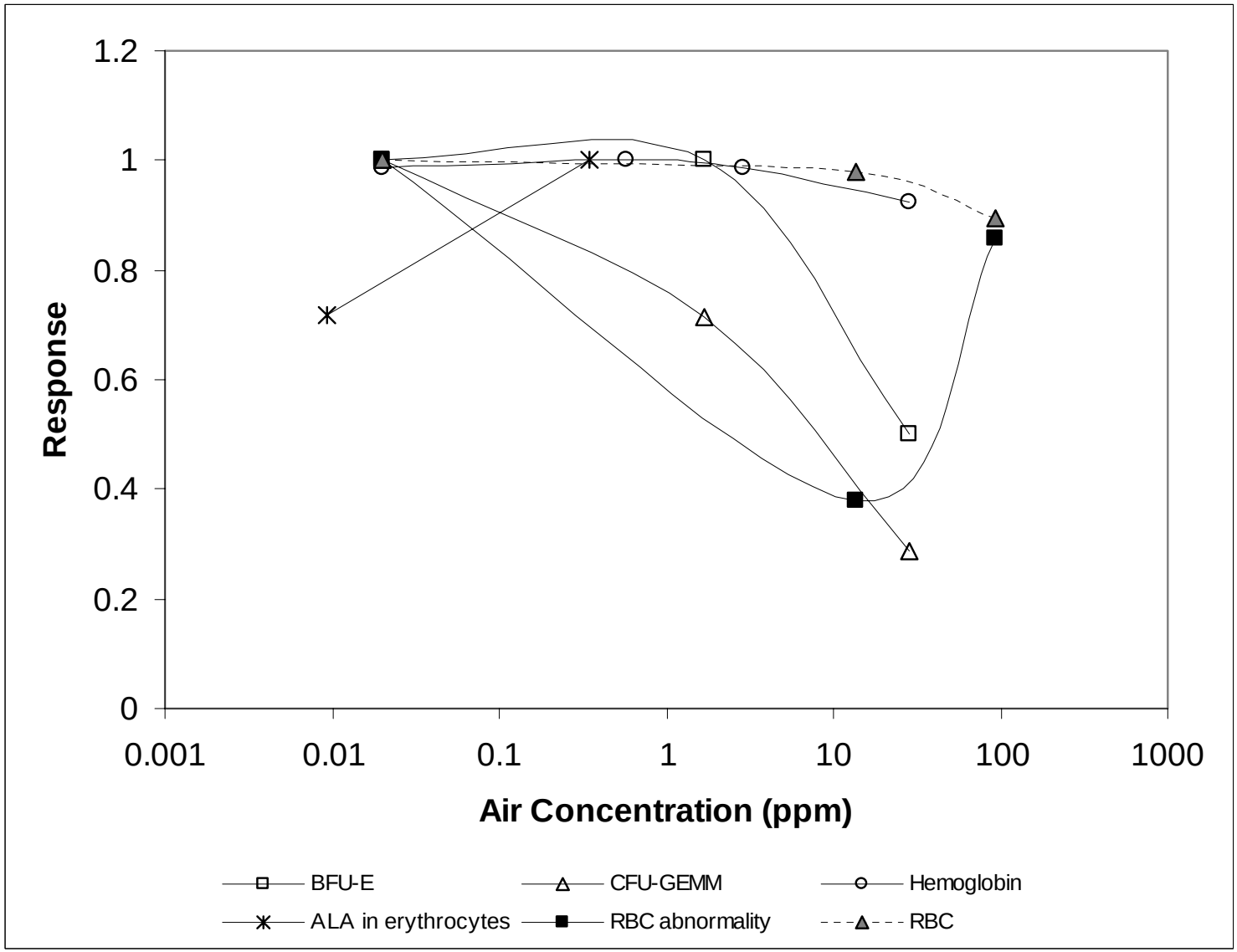

Figure 4. Graphical Comparison of Precursors of Leukopenia (WBC)

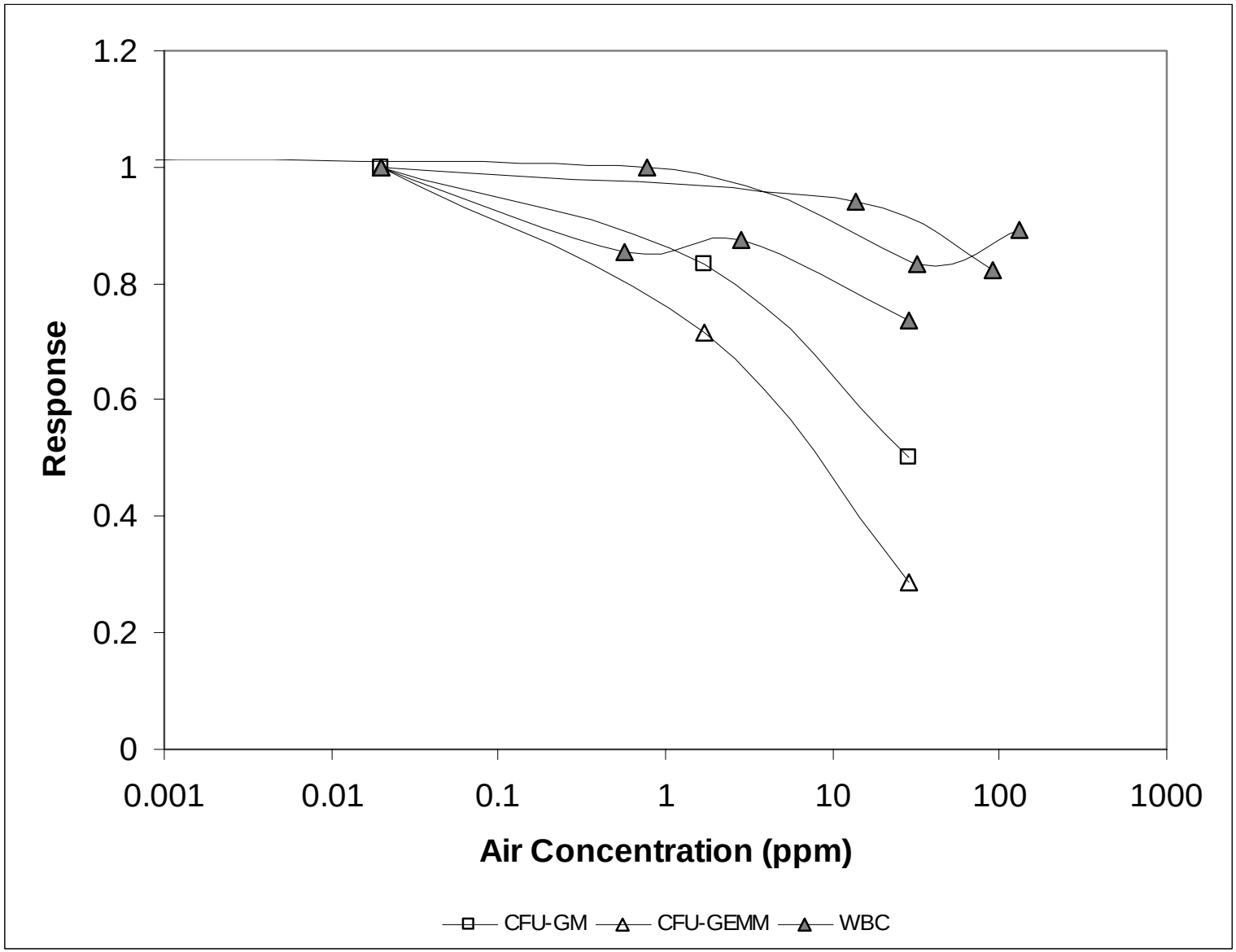

Figure 5. Network of Candidate Biomarkers of Benzene Exposure and Effect

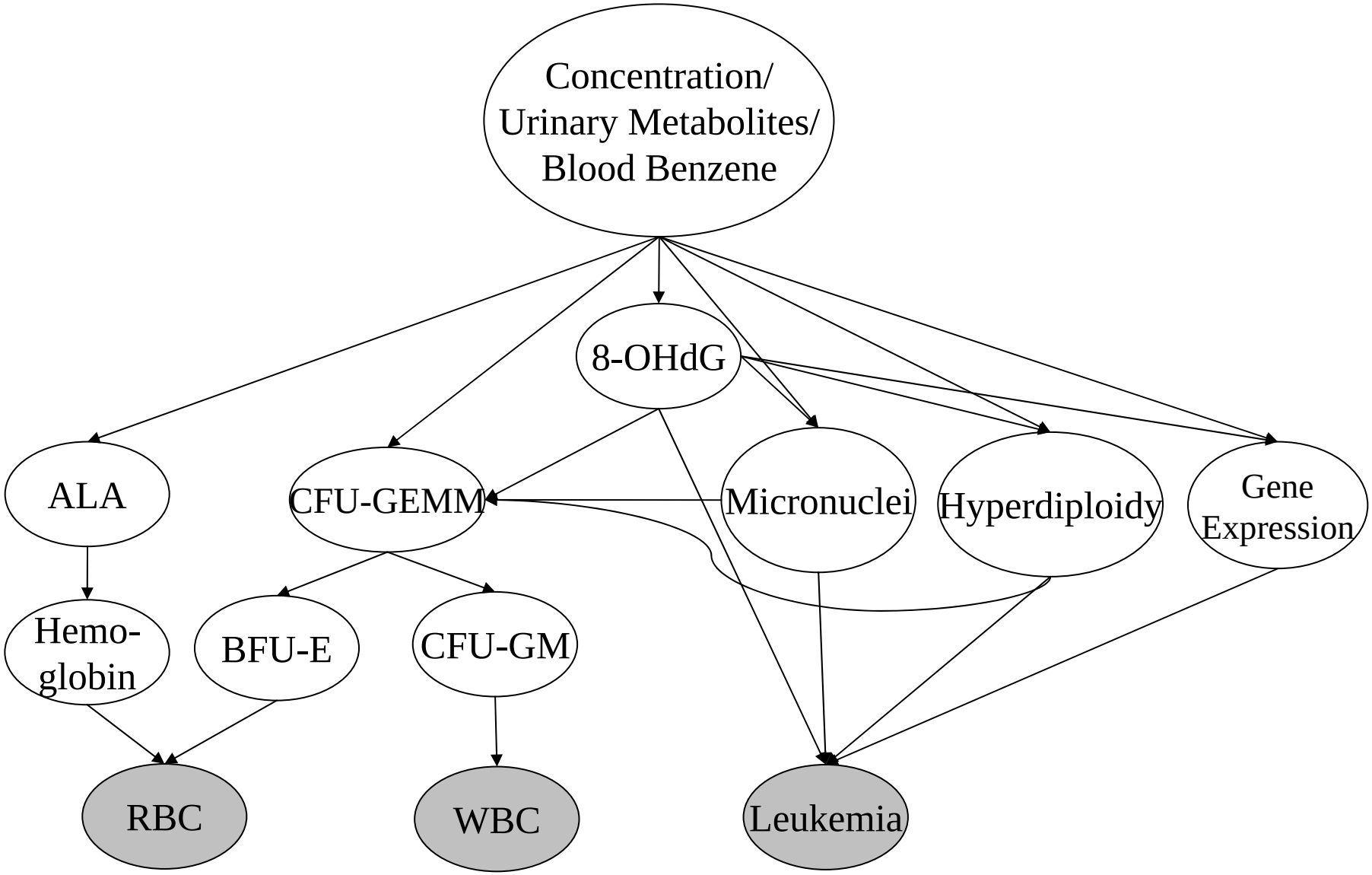

Figure 6. Validated Network of Biomarkers of Benzene Exposure and Effect

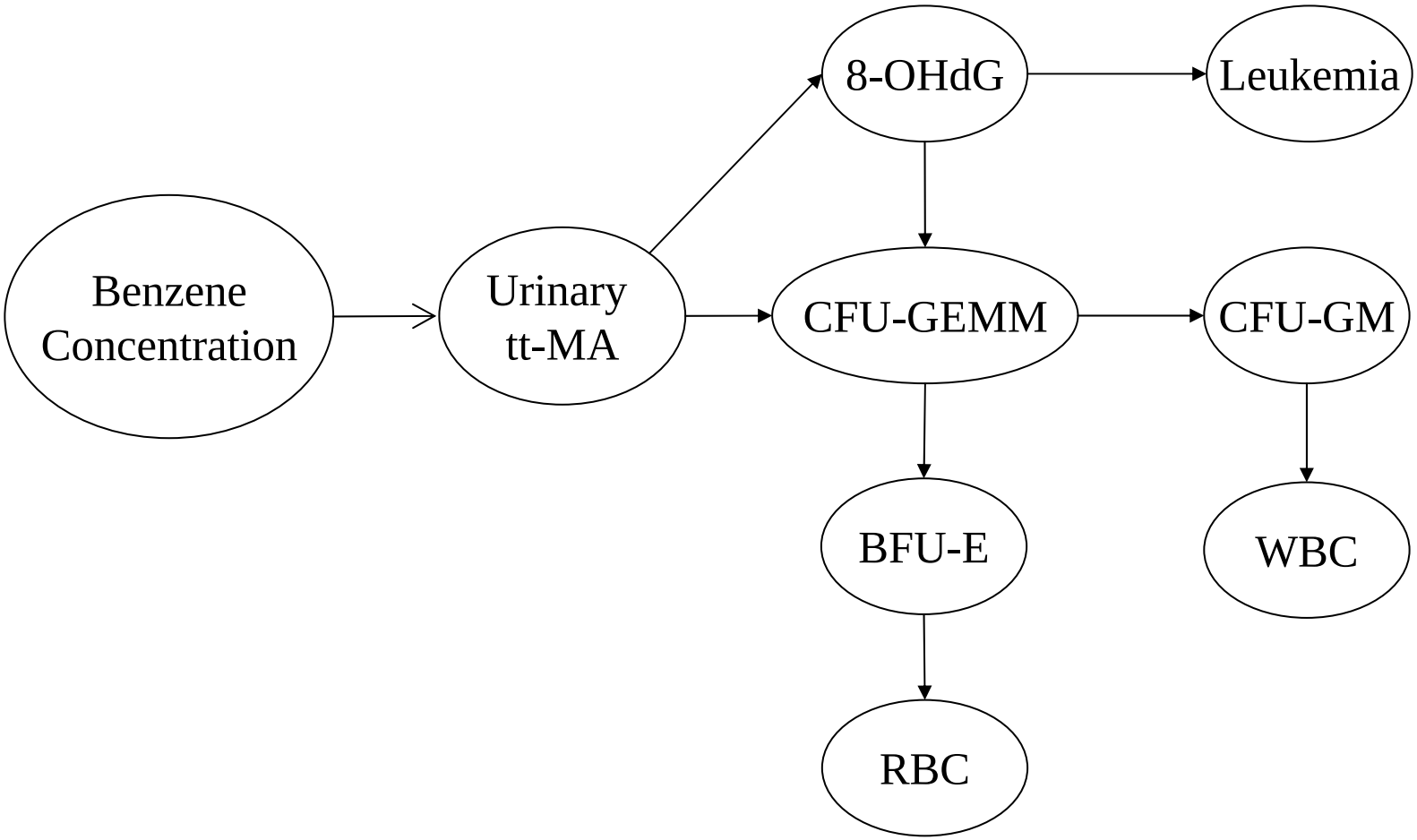

Figure 7. Uncertainty Distribution for the Slope from % Micronuclei to Leukemia

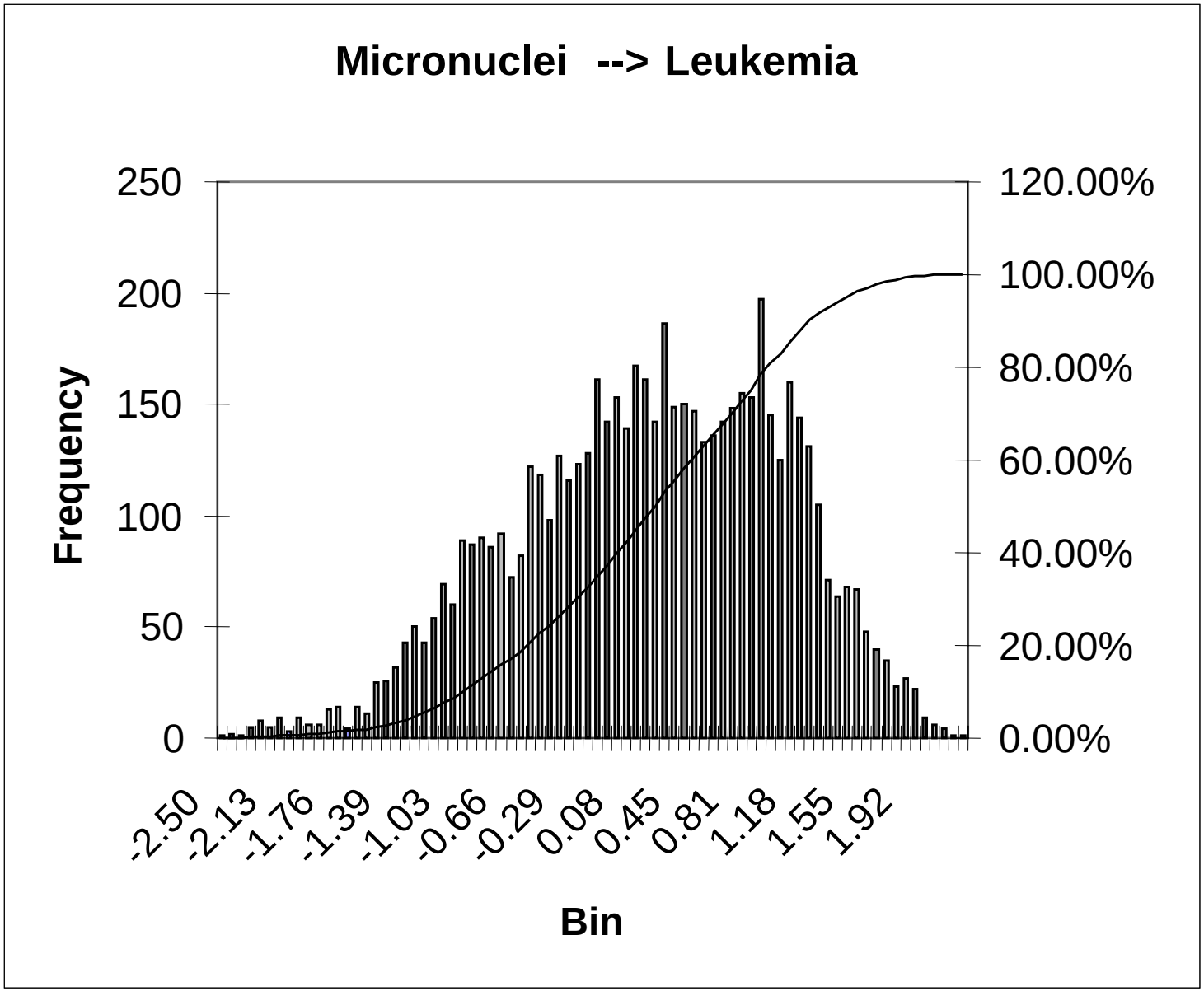

Figure 8. Uncertainty Distribution for the Slope from 8-OHdG to Leukemia Pr (Slope > 0) = 85%, based on the posterior frequency distribution for this parameter from the MCMC calibration of the Bayesian network of potential biomarkers.