Practice of Epidemiology

The Impact of Saturable Metabolism on Exposure-Response Relations in 2 Studies of Benzene-induced Leukemia

Jelle Vlaanderen, Lützen Portengen, Stephen M. Rappaport, Deborah C. Glass, Hans Kromhout, and Roel Vermeulen

* Correspondence to Dr. Roel Vermeulen, Environmental Epidemiology Division, Institute for Risk Assessment Sciences, P.O. Box 80.178, NL-3508 TD Utrecht, the Netherlands (e-mail: r.c.h.vermeulen@uu.nl).

Initially submitted December 14, 2010; accepted for publication March 16, 2011.

Enzymatic saturation of metabolic pathways is one factor that potentially contributes to the nonlinear exposure-response relations that are frequently reported in occupational epidemiologic studies. The authors propose an approach to explore the contribution of saturable metabolism to previously reported exposure-response relations by integrating predictive models of relevant biomarkers of exposure into the epidemiologic analysis. The approach is demonstrated with 2 studies of leukemia in benzene-exposed workers, one conducted in the Australian petroleum industry (1981–1999) and one conducted in a US rubber hydrochloride production factory in Ohio (1940–1996). The studies differed greatly in their magnitudes and durations of exposure. Substitution of biomarker levels for external estimates of benzene exposure reduced the fold difference of the log relative risk of leukemia per unit of cumulative exposure between the 2 studies by 11%–44%. Nevertheless, a considerable difference in the log relative risk per unit of cumulative exposure remained between the 2 studies, suggesting that exposure misclassification, differences in study design, and potential confounding factors also contributed to the heterogeneity in risk estimates.

benzene; case-control studies; cohort studies; leukemia; metabolism; occupational exposure; pharmacokinetics

Abbreviation: AIC, Akaike Information Criterion.

Cumulative exposure (the product of intensity and duration of exposure) is a standard exposure metric used in many occupational studies of chronic health effects (1). The frequent use of cumulative exposure as index of the target tissue dose is based on 3 broad assumptions: 1) The cumulative probability of developing a disease is proportional to the sum of the daily probabilities of developing a disease; 2) the daily probability of developing a disease increases monotonically with the concentration in the target tissue; and 3) the concentration in the target tissue is linearly related to the external exposure (2). Several authors have demonstrated that these assumptions are not always valid (3–5). For example, Doll and Peto (5) demonstrated that the cumulative probability of developing a disease is proportional to the sum of the daily probabilities of developing a disease; 2) the daily probability of developing a disease increases monotonically with the concentration in the target tissue; and 3) the concentration in the target tissue is linearly related to the external exposure (2). Several authors have demonstrated that these assumptions are not always valid (3–5). For example, Doll and Peto (5) demonstrated that the cumulative probability of developing a disease is proportional to the sum of the daily probabilities of developing a disease; 2) the daily probability of developing a disease increases monotonically with the concentration in the target tissue; and 3) the concentration in the target tissue is linearly related to the external exposure (2).

A higher relative risk was predicted for cumulative exposure delivered at high intensity for shorter duration than for cumulative exposure delivered at low intensity for longer duration. In attempting to explain their findings, Lubin et al. (3) speculated that the impact of “exposure delivery” might be explained by a nonlinear relation between exposure intensity (air concentration of arsenic) and the concentration of arsenic metabolites in the target tissue, due to saturation of arsenic metabolism. The effects of saturable metabolism on delivered dose have been shown to affect outcomes of toxicology studies (6) and have been suspected of being important in occupational epidemiologic studies (7).

Alteration of exposure-response relations due to saturable metabolism could affect the interpretation of occupational epidemiologic studies in risk assessment because the common practice of linear extrapolation of exposure-response curves derived at high exposure levels to lower
levels relevant to the general public could result in highly biased relative risks. The direction of such biases could be either positive or negative depending on whether the saturating pathway was activating or deactivating in terms of the ultimate toxicant(s). In either case, estimates of the relative risk per unit of exposure would no longer be comparable at different ends of the exposure spectrum.

Here, we describe applications of a physiologically based pharmacokinetic model (8) and empirical models based on biomarker data (9, 10) to quantitatively relate levels of airborne exposure to internal levels of toxic metabolites in occupational epidemiologic studies. Our approach is based on the conjecture that if a health outcome is not proportional to the air concentration of an agent because of saturable metabolism, then the exposure-response relation would become linear (i.e., outcome proportional to exposure) when air concentrations are substituted by predictions of relevant biomarker levels from physiologically based pharmacokinetic or empirical models. As a result, relative risks that are expressed per unit of metabolite or biomarker level for a given toxicant should be more comparable across studies than relative risks expressed per unit of airborne exposure.

In a recent meta-regression analysis, we observed that relative risk estimates from occupational epidemiologic studies suggested a supralinear shape of the benzene-leukemia exposure-response curve (i.e., greater than proportional relative risk at lower exposure levels) (11). Such nonlinear behavior is supported by several lines of evidence suggesting that pathways leading to production of benzene metabolites saturate with increasing exposure to airborne benzene (9, 10, 12–18). The major metabolic pathways for benzene are shown in Figure 1. Benzene is metabolized by cytochrome P450 enzymes (primarily CYP2E1) to benzene oxide, which is in equilibrium with its valence tautomer oxepin, and is the source of all other metabolites. For airborne exposure levels between 0.1 and 10 ppm, phenol represents 70%–85% of the urinary metabolites; hydroquinone, t,t-muconic acid, and catechol each represent 5%–10%; and \( S \)-phenylmercapturic acid represents <1% (9).

To illustrate our approach, we used data from 2 prominent occupational studies of benzene-induced leukemia, namely, the Health Watch Study (19) and the Pliofilm Study (20). The more recent Health Watch Study was performed in the petroleum industry where workers were generally exposed
for many years to relatively low air concentrations of benzene (seldom above 5 ppm) (19). In contrast, the older Pliofilm Study was performed in a chemical factory where workers experienced much higher benzene exposures (up to 60 ppm) but for short periods of time (20). Interestingly, the Health Watch Study reported much higher relative risks for leukemia by cumulative exposure category than the Pliofilm Study, even though the Health Watch Study included lower cumulative exposures than the Pliofilm Study (19, 20). Here, we examine the difference in log relative risk of leukemia per unit of cumulative benzene exposure in the Health Watch and Pliofilm studies and explore the extent to which this difference might be explained by saturable benzene metabolism.

**MATERIALS AND METHODS**

In Figure 2, we present a scheme of our approach for integrating the predictions of benzene metabolite levels into the epidemiologic analyses of benzene exposure and leukemia risk. The scheme consists of the following steps.

**Extraction of unique benzene exposure intensities from the study data**

The Pliofilm and Health Watch studies were selected on the basis of their quality (21) and the large contrast between them in terms of the duration and intensity of benzene exposure. All unique intensities of benzene exposure (I\textsubscript{ben}unique) assigned to individual subjects in the original epidemiologic analysis were extracted. We interpret I\textsubscript{ben}unique as the arithmetic mean of an underlying distribution of 8-hour average exposure intensity levels experienced by a subject.

**Simulation of the underlying distribution for each unique exposure intensity**

To reconstruct the full distribution of exposure intensities that each study subject is likely to have experienced, we simulated the underlying exposure distribution of each I\textsubscript{ben}unique as follows. First, within-worker variance components (δ\textsubscript{ww}) and between-worker variance components (δ\textsubscript{bw}) were assumed for the Health Watch and Pliofilm studies by using the estimated median values of δ\textsubscript{ww} and δ\textsubscript{bw} reported by Kromhout et al. (22) for the chemical industry (Pliofilm Study: δ\textsubscript{ww} = 2.05, δ\textsubscript{bw} = 1.49) and the petroleum-refining industry (Health Watch Study: δ\textsubscript{ww} = 3.35, δ\textsubscript{bw} = 1.43).

Then, the natural logarithm of the geometric mean (I\textsubscript{m}unique-gm) corresponding to each I\textsubscript{ben}unique was calculated as \(\ln(I_{\text{m\_unique-gm}}) = \ln(I_{\text{unique}}) - \ln(\delta_{\text{ww}} + \delta_{\text{bw}})/2\) (23). Next, a geometric mean exposure intensity (I\textsubscript{m\_worker-gm}) was calculated for each simulated worker (I\textsubscript{m\_worker-gm}) = \(\ln(I_{\text{m\_unique-gm}})\). The geometric mean of I\textsubscript{m\_worker-gm} was calculated for each unique exposure level (I\textsubscript{m\_unique-gm}), and the arithmetic mean (I\textsubscript{m\_unique}) was estimated on the basis of I\textsubscript{m\_unique-gm} and its variance.

5) Substitution of I\textsubscript{ben\_unique} with I\textsubscript{m\_unique} in the calculation of cumulative exposure.

**Figure 2.** Approach to substitute benzene with measures of relevant benzene metabolite biomarkers. 1) Extraction of unique benzene exposure intensities (I\textsubscript{ben\_unique}) from the occupational study data. 2) Simulation of the underlying distribution for each unique exposure intensity. For each I\textsubscript{ben\_unique}, an underlying distribution consisting of 1,000 measurements was simulated. The simulation was based on 50 (hypothetical) workers per I\textsubscript{ben\_unique} and 20 (hypothetical) exposure measurements per worker. 3) Prediction of relevant metabolite levels (I\textsubscript{m\_sim}), sum of catechol, hydroquinone, t\textsuperscript{-}muconic acid, S-phenyl mercapturic acid, and phenol, based on I\textsubscript{m\_unique}. 4) Calculation of the arithmetic mean of I\textsubscript{m\_unique} (I\textsubscript{m\_unique}). The geometric mean of I\textsubscript{m\_worker-gm} was calculated for each simulated worker (I\textsubscript{m\_worker-gm}), the median of I\textsubscript{m\_worker-gm} was calculated for each unique exposure level (I\textsubscript{m\_unique-gm}), and the arithmetic mean (I\textsubscript{m\_unique}) was estimated on the basis of I\textsubscript{m\_unique-gm} and its variance. 5) Substitution of I\textsubscript{ben\_unique} with I\textsubscript{m\_unique} in the calculation of cumulative exposure.
total, 1,000 values of $I_{bsim}$ exposure were simulated for each value of $I_{b_{unique}}$.

**Prediction of urinary metabolite levels**

The sum of urinary levels of catechol, hydroquinone, $t,t$-muconic acid, phenol, and $S$-phenylmercapturic acid (hereafter “sum of metabolites”) was used as the indicator of benzene metabolism. These sums of metabolites were estimated from the corresponding values of $I_{bsim}$ by 3 different methods.

**Physiologically based pharmacokinetic model.** The physiologically based pharmacokinetic model reported by Yokley et al. (16) was used to predict the amount of catechol, hydroquinone conjugates, $t,t$-muconic acid, phenol conjugates, and $S$-phenylmercapturic acid excreted in urine after 8 hours of continuous exposure to benzene. We constructed the physiologically based pharmacokinetic model using Berkeley Madonna, version 8.3.14, software (University of California, Berkeley, California). We assumed an 8-hour urinary output of 0.4 L (24) to convert the amount of the sum of metabolites that was predicted by the physiologically based pharmacokinetic model into urinary concentrations of the sum of metabolites.

**Michaelis-Menten–like model.** The Michaelis-Menten–like model reported by Rappaport et al. (10) was used to predict the sum of metabolites. This Michaelis-Menten–like model assumes 2 metabolic pathways that compete for access to benzene plus an intercept that accounts for a background level of benzene metabolites from dietary and endogenous sources (25). Parameters for the Michaelis-Menten–like model had been estimated by fitting the model to data from 263 nonsmoking females from 2 studies of Chinese workers for whom individual levels of airborne benzene (0.001–299 ppm) and urinary metabolite levels had been documented (10, 26, 27).

**Regression splines model.** The 5 regression splines reported by Kim et al. (9) were used to relate airborne benzene exposure to urinary levels of catechol, hydroquinone, $t,t$-muconic acid, phenol, and $S$-phenylmercapturic acid. The predictions of the individual regression splines were summed to calculate the sum of metabolites. The splines had been derived from a cross-sectional study that included airborne benzene (0.03–88.9 ppm) and urinary metabolite measurements of 326 exposed workers in Tianjin, China (9).

Note that there is a considerable overlap (240 nonsmoking females) in the data that were used to derive the regression splines and the parameters for the Michaelis-Menten–like model. For the empirical models (the Michaelis-Menten–like model and the regression splines model), the urinary level of metabolites predicted for a benzene exposure of 0 ppm was subtracted from the model prediction to isolate the contribution of benzene exposure to the sum of metabolites. Because the regression splines were fitted by using log-transformed airborne benzene exposure levels, we used the second derivative of the splines at the first knot (0.004 ppm for hydroquinone, $t,t$-muconic acid, phenol, and $S$-phenylmercapturic acid, as well as 0.04 ppm for catechol) to predict the shape of the curve for exposures below the first knot. The intercepts of the second derivatives at 0 ppm occupational benzene exposure (catechol = 11.93 $\mu$mol/L, hydroquinone = 5.79 $\mu$mol/L, $t,t$-muconic acid = 0.79 $\mu$mol/L, phenol = 63.23 $\mu$mol/L, $S$-phenylmercapturic acid = 0.003 $\mu$mol/L) were summed to estimate the

[Figure 3. Cumulative distribution of cumulative exposure, duration of exposure, and average intensity of exposure to benzene in the Pliofilm (1940–1996) and Health Watch (1981–1999) studies. Plots are based on occupationally benzene-exposed individuals only; non-exposed individuals in the Pliofilm Study were excluded (30% of the study population). A, cumulative exposure; B, duration of exposure; C, average intensity of exposure. Solid curve, Pliofilm Study; long dashed curve, Health Watch Study.]

The background level of the sum of metabolites for the regression splines model.

The 3 models were used to predict Im_{sim} for each level of Ib_{sim}. Subsequently, the geometric mean of Im_{sim} was calculated for each simulated worker (Im_{worker}), the median of Im_{worker} was calculated for each unique exposure level (Im_{unique-gm}), and the arithmetic mean (Im_{unique}) was estimated on the basis of Im_{unique-gm} and its variance (Figure 2).

**Epidemiologic analyses**

Epidemiologic analyses were conducted by using 2 measures of exposure intensity, namely, each subject’s arithmetic mean air concentration (Ib_{unique}) and the corresponding arithmetic mean metabolite level (Im_{unique}), based on the physiologically based pharmacokinetic, Michaelis-Menten-like, and regression splines models. Ib_{unique} and Im_{unique} levels were combined with the duration of exposure data from the original studies to calculate cumulative and average intensity of exposure to benzene or the sum of metabolites. The log relative risk per unit of exposure to benzene or to the sum of metabolites was estimated with conditional logistic regression for the Health Watch Study data (similar to the analysis of Glass et al. (19)) and with Cox proportional hazards regression for the Pliofilm Study data (similar to the analysis of Rinsky et al. (20)). The Akaike Information Criterion (AIC) (28) was used to compare the fit of the epidemiologic models to the data. Both types of models were fitted with the PHREG procedure in SAS, version 9.2, software (SAS Institute, Inc., Cary, North Carolina). All models were run 5 times to account for the potential influence of random variation in the simulation of the underlying distribution of unique exposure intensities. The MIANALYZE procedure in SAS software was used to combine the results of the 5 runs. To explore whether there was evidence for a nonlinear effect of exposure to benzene or to the sum of metabolites, we repeated the epidemiologic analyses with the clogit (Health Watch Study) and coxph (Pliofilm Study) procedures from the survival package for R, version 2.9.1, program (R Foundation for Statistical Computing, Vienna, Austria) and included a penalized spline (degrees of freedom based on the AIC) for exposure. A likelihood ratio test demonstrated that allowing more flexibility for the effect of exposure within studies did not

<table>
<thead>
<tr>
<th>Study</th>
<th>Estimate (SE)</th>
<th>AIC</th>
<th>1 ppm-year^b</th>
<th>10 ppm-years^b</th>
<th>20 ppm-years^b</th>
<th>40 ppm-years^b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>RR 95% CI</td>
<td>RR 95% CI</td>
<td>RR 95% CI</td>
<td>RR 95% CI</td>
</tr>
<tr>
<td>Health Watch</td>
<td>9.40 \times 10^{-2} (2.71 \times 10^{-2})</td>
<td>103.5</td>
<td>1.10, 1.04, 1.16</td>
<td>2.57, 1.51, 4.37</td>
<td>6.58, 2.27, 19.07</td>
<td>43.33, 5.16, 363.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P = 0.0005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pliofilm</td>
<td>5.10 \times 10^{-3} (1.11 \times 10^{-3})</td>
<td>206.9</td>
<td>1.01, 1.00, 1.01</td>
<td>1.05, 1.03, 1.08</td>
<td>1.11, 1.06, 1.16</td>
<td>1.23, 1.12, 1.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P &lt; 0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>Estimate (SE)</th>
<th>AIC</th>
<th>1 year^c</th>
<th>5 years^c</th>
<th>10 years^c</th>
<th>20 years^c</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>RR 95% CI</td>
<td>RR 95% CI</td>
<td>RR 95% CI</td>
<td>RR 95% CI</td>
</tr>
<tr>
<td>Health Watch</td>
<td>3.39 \times 10^{-3} (3.02 \times 10^{-2})</td>
<td>119.0</td>
<td>1.03, 0.62, 1.72</td>
<td>1.18, 0.69, 15.22</td>
<td>1.40, 0.01, 231.60</td>
<td>1.97, 0, 53,637</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P = 0.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pliofilm</td>
<td>7.17 \times 10^{-2} (2.15 \times 10^{-2})</td>
<td>214.3</td>
<td>1.07, 1.03, 1.12</td>
<td>1.43, 1.16, 1.77</td>
<td>2.05, 1.34, 3.12</td>
<td>4.20, 1.81, 9.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P = 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>Estimate (SE)</th>
<th>AIC</th>
<th>1 ppm^c</th>
<th>2 ppm^c</th>
<th>5 ppm^c</th>
<th>10 ppm^c</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>RR 95% CI</td>
<td>RR 95% CI</td>
<td>RR 95% CI</td>
<td>RR 95% CI</td>
</tr>
<tr>
<td>Health Watch</td>
<td>2.06 (5.86 \times 10^{-1})</td>
<td>103.5</td>
<td>7.81, 2.48, 24.66</td>
<td>61.04, 6.13, 608</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P = 0.0005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pliofilm</td>
<td>2.50 \times 10^{-2} (1.27 \times 10^{-2})</td>
<td>216.3</td>
<td>1.03, 0.94, 1.12</td>
<td>1.05, 0.88, 1.25</td>
<td>1.14, 0.74, 1.75</td>
<td>1.29, 0.54, 3.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P = 0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: AIC, Akaike Information Criterion; CI, confidence interval; NC, not calculated (outside the range of study exposure levels); RR, relative risk; SE, standard error.

^a Estimate of the log relative risk of leukemia per unit of exposure. Estimate and associated standard error are from conditional logistic regression for the Health Watch data and from Cox proportional hazards regression for the Pliofilm data.

^b Relative risks and 95% confidence intervals were predicted with the log relative risk per unit of exposure and the associated standard error.

^c Calculated as cumulative exposure divided by duration of exposure.
significantly increase model fit in any of the analyses (results not shown). Consequently, only the linear estimate of the log relative risk per unit of exposure was used in the current analysis.

RESULTS

Health Watch and Pliofilm studies

In Figure 3, the characteristics of the exposure metrics used in the Health Watch Study and Pliofilm Study are compared by plotting the cumulative distribution of cumulative exposure (part A), duration of exposure (part B), and average intensity of exposure to benzene (part C). The plots for the Pliofilm Study are based on exposed workers only (70% of the study population). Figure 3A illustrates that the Health Watch Study population had consistently lower cumulative exposures than the Pliofilm Study population. More extreme differences between the Health Watch Study and the Pliofilm Study were observed for duration of exposure (the full population of the Health Watch Study was exposed for 5–10 years, while the majority of the Pliofilm Study population was exposed for less than 1 year) and average intensity of exposure (calculated as cumulative exposure divided by duration of exposure). The average intensity of exposure in the Health Watch Study ranged from 0.001 to 1 ppm, while in the Pliofilm Study the average intensity of exposure ranged from 1 to 40 ppm. The median assigned exposure intensity was 0.1 ppm in the Health Watch Study and 23 ppm in the Pliofilm Study (not shown in Figure 3).

In Table 1, the estimates for the log relative risk per unit of cumulative exposure, duration of exposure, and average intensity of exposure to benzene are reported for the Health Watch Study and the Pliofilm Study. Relative risks with 95% confidence intervals were calculated for 4 different exposure levels on the basis of the log relative risk per unit of exposure and the associated standard error. The estimate of the log relative risk per unit of cumulative exposure was significant in both the Health Watch Study (P = 0.0005) and the Pliofilm Study (P < 0.0001). However, the slope estimate for the Health Watch Study was ~20-fold higher than the estimate for the Pliofilm Study. Interestingly, no formal statistical significant effect for duration of exposure was observed in the Health Watch Study (P = 0.26), while the effect in the Pliofilm Study was significant (P = 0.001). Furthermore, the effect of average intensity of exposure observed in the Health Watch Study was ~100-fold higher than the effect of average intensity observed in the Pliofilm Study. The comparison of AIC values (28) across the different models in Table 1 showed that the models based on cumulative exposure and average intensity of exposure to benzene fit the data equally well in the Health Watch Study but not in the Pliofilm Study, where cumulative exposure to benzene provided a much better fit than average intensity of exposure.

Predicted levels of benzene metabolites

Figure 4A (exposure range, 0–150 ppm) and Figure 4B (exposure range, 0–10 ppm) show the sum of metabolites predicted by the physiologically based pharmacokinetic model, Michaelis-Menten–like model, and regression splines model. The 2 reference lines in Figure 4A indicate the maximum exposure intensity levels that were assigned in the Health Watch Study (~17 ppm) and the Pliofilm Study (~60 ppm). At benzene concentrations above about 50 ppm, the metabolite levels predicted from the regression splines were substantially higher than those from the other 2 models (Figure 4A). For the exposure range below 10 ppm, the Michaelis-Menten–like and regression splines models predicted greater-than-proportional production of the sum of metabolites, while the prediction of the physiologically based pharmacokinetic model was proportional to the airborne benzene concentration (Figure 4B).
Table 2. Log Relative Risk of Leukemia per Unit Cumulative Exposure to Benzene or Sum of Metabolitesa in the Health Watch (1981–1999) and Pliofilm (1940–1996) Studies

<table>
<thead>
<tr>
<th>Exposure Metric</th>
<th>Health Watch Study Estimate (SE)b</th>
<th>AICc</th>
<th>Pliofilm Study Estimate (SE)b</th>
<th>AICc</th>
<th>Fold Differencea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzene</td>
<td>98.9 × 10⁻³ (2.93 × 10⁻⁴)</td>
<td>103.7</td>
<td>5.35 × 10⁻³ (1.23 × 10⁻⁵)</td>
<td>207.0</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>P = 0.0007</td>
<td></td>
<td>P &lt; 0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum of metabolitesa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBPK model</td>
<td>2.09 × 10⁻³ (6.16 × 10⁻⁴)</td>
<td>104.4</td>
<td>0.13 × 10⁻³ (3.03 × 10⁻⁵)</td>
<td>207.5</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>P = 0.0007</td>
<td></td>
<td>P &lt; 0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MML model</td>
<td>1.22 × 10⁻³ (3.54 × 10⁻⁴)</td>
<td>103.9</td>
<td>0.12 × 10⁻³ (2.79 × 10⁻⁵)</td>
<td>207.6</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>P = 0.0006</td>
<td></td>
<td>P &lt; 0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression splines model</td>
<td>1.32 × 10⁻³ (3.78 × 10⁻⁴)</td>
<td>103.3</td>
<td>0.11 × 10⁻³ (2.45 × 10⁻⁵)</td>
<td>207.2</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>P = 0.0005</td>
<td></td>
<td>P &lt; 0.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: AIC, Akaike Information Criterion; MML, Michaelis-Menten like; PBPK, physiologically based pharmacokinetic; SE, standard error.

a Sum of catechol, hydroquinone, t-muconic acid, S-phenyl mercapturic acid, and phenol.
b Estimates and associated standard errors are from conditional logistic regression for the Health Watch data and from Cox proportional hazards regression for the Pliofilm data. All models included a simulation step to estimate the underlying distribution of unique exposure intensities used in the original epidemiologic analysis. Models were run 5 times to account for the potential influence of random variation in the simulation of the underlying distribution of unique exposure intensities. The MIANALYZE procedure in SAS, version 9.2, software (SAS Institute, Inc., Cary, North Carolina) was used to combine the results of the 5 runs.

c Calculated as the average of the AIC values reported for the 5 runs.
d "Fold difference" is defined as the log relative risk per unit of cumulative exposure derived in the Health Watch Study divided by the log relative risk per unit of cumulative exposure derived in the Pliofilm Study.

Effect of modeling saturable metabolism on risk estimates

In Table 2, the slope estimates are given for the log relative risk of leukemia per unit of either cumulative exposure to benzene or the sum of metabolites. Based on cumulative exposure to benzene, the unit risk estimate for the Pliofilm Study was 18-fold higher than that for the Health Watch Study. When cumulative sum of metabolites was used as the predictor variable instead of cumulative benzene exposure, the difference in unit risks between the 2 decreased to 16-fold (11% decrease) for the physiologically based pharmacokinetic model, 12-fold (33% decrease) for the regression splines model, and 10-fold (44% decrease) for the Michaelis-Menten–like model. Evaluation of AIC values indicated that modeling saturable metabolism did not improve the fits of models for either the Health Watch Study or the Pliofilm Study. Analyses based on average intensity of benzene exposure and average intensity of the sum of metabolites resulted in a similar (albeit stronger) impact of modeling saturable metabolism on the leukemia risk estimates in the Health Watch Study and the Pliofilm Study ranging from a 29% to a 51% decrease in fold difference.

DISCUSSION

By considering possible saturation of metabolism in 2 prominent epidemiologic studies of benzene-induced leukemia (Pliofilm Study and Health Watch Study), we were able to narrow the difference in estimates of the log relative risk per unit of cumulative exposure by about a factor of 2 (Pliofilm Study/Health Watch Study: decreased from 18 for external exposure to 10 for the Michaelis-Menten–like model of metabolite levels). We therefore provided some evidence that the heterogeneity in relative risk estimates between these 2 studies might be explained, in part, by saturable metabolism. The interpretation of our findings is complicated by the fact that saturable metabolism was probably not the only factor contributing to the heterogeneity of leukemia risk estimates reported in the Health Watch and Pliofilm studies. Exposure misclassification is a common problem in retrospective occupational cohort studies (29) that could have affected the estimated log relative risk per unit of cumulative exposure in the 2 studies. Systematic differences in exposure misclassification between the studies might have resulted from the considerable degree of extrapolation and expert judgment, particularly in the Pliofilm Study which had access to far fewer benzene measurements than did the Health Watch Study (~750 measurements in the Pliofilm Study (30) compared with >3,870 measurements in the Health Watch Study (31)). Indeed, the limited quality of the exposure data in the Pliofilm Study resulted in 3 groups of authors publishing 3 different sets of exposure estimates, based on the same data (20, 32–34). Other potential explanations for the heterogeneity in risk estimates between the Health Watch Study and the Pliofilm Study could be related to differences between the study populations in their susceptibility to benzene-induced health effects, differences in background rates for leukemia (it has been suggested that background rates were too low in the Health Watch Study (35)), differences in study design (nested case-control vs. cohort), or differences in confounding factors (7, 36). In addition, leukemia is not a single-disease entity. Stronger
associations have been shown for some leukemia subtypes (e.g., acute myeloid leukemia) than for other subtypes. Differences between the study populations in the distribution of subtypes that contributed to the overarching disease outcome “leukemia” might have contributed to heterogeneity in the risk estimates as well.

Our decision to use the sum of the major urinary benzene metabolites to reflect saturable metabolism was motivated by the readily available models that quantitatively described the relation between airborne exposure to benzene and these biomarkers of exposure and the fact that they collectively account for virtually all of the metabolized benzene dose (9, 10, 16). Furthermore, although some individual metabolites (i.e., hydroquinone and 1,4-benzoquinone) are frequently mentioned as being the most toxic benzene metabolites, there is still considerable uncertainty regarding exactly which metabolites are causally related to leukemia (15, 9, 25). When we applied our approach using the predictions of the physiologically based pharmacokinetic and regression splines models for only urinary hydroquinone concentrations, the results were essentially the same as for the sum of metabolites (refer to Web Table 1, which is posted on the Journal’s Web site (http://aje.oupjournals.org/)). Although the biomarkers were not measured in the target tissue (the bone marrow, where benzene-induced toxicologic effects are primarily thought to take place), they do reflect saturation of the cytochrome P450 enzymes that are primarily active in the liver (10, 15).

We should consider the possibility that the predictive models included in our approach did not fully reflect the extent to which saturable metabolism actually occurs in humans. In the physiologically based pharmacokinetic model, error might have been introduced by invalid specification of the model or its parameters, and measurement errors of benzene and its metabolites might have played a role in the empirical spline and Michaelis-Menten–like models. It is also important to mention that the empirical models were derived in populations of Chinese predominantly female workers that may not be directly comparable to the predominantly white male workers that were included in the study populations of the Health Watch Study and the Pliofilm Study.

Our approach is a form of dosimetric modeling that constructs measures of exposure by using explicit hypotheses about the exposure-dose and/or dose-risk relation (37). Extending the approach by incorporating explicit hypotheses about the interaction between duration of exposure and the concentration in the target tissue, clearance from the target tissue, repair processes, and so on could provide further insight into the exposure-disease relations. Conolly et al. (38) presented an example of such an extended model applied to formaldehyde and cancer of the respiratory tract. However, extension of our approach in the current example of benzene and leukemia would likely be limited by the statistical power and the levels of detail that are available in the exposure data in the Health Watch and Pliofilm studies.

As an alternative to constructing hypothesis-based dosimetric models, flexible approaches that allow independent weighting of duration, intensity, and timing of exposure can also be used to better predict cancer incidence in occupational studies and to evaluate heterogeneity in risk estimates across studies. Several approaches have been proposed for doing this (e.g., refer to references 3, 4, 39) but are generally difficult to apply in occupational studies that often lack sufficient statistical power for complex models. Knowledge about saturable metabolism can provide clues as to where to expect nonlinearity in flexible models and will therefore be helpful in the interpretation of the outcomes of flexible approaches.

Our approach provides a general framework for estimating internal exposure levels from external measurements and then using these internal levels to explore the effect of saturation on exposure-response relations. Our approach can be applied to studies of any toxicant, provided that substantial data are available for external exposure levels and that the metabolic pathway supports a hypothesis regarding the impact of saturable metabolism on the exposure-response relation.

ACKNOWLEDGMENTS

Author affiliations: Institute for Risk Assessment Sciences, Utrecht University, Utrecht, the Netherlands (Jelle Vlaanderen, Lützen Portengen, Hans Kromhout, Roel Vermeulen); School of Public Health, University of California, Berkeley, California (Stephen M. Rappaport); and Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia (Deborah C. Glass).

This work was performed as part of the work package “integrated risk assessment” of the ECNIS Network of Excellence (Environmental Cancer Risk, Nutrition, and Individual Susceptibility), operating within the European Union’s 6th Framework Program, Priority 5: “Food Quality and Safety” (FOOD-CT-2005-513943). S. M. R. has received research support from the American Petroleum Institute and the American Chemistry Council, as well as consulting and expert testimony fees from law firms representing plaintiffs’ cases involving exposure to benzene. D. C. G. receives funding from the Australian Institute of Petroleum for the ongoing maintenance and updating of the Health Watch Cohort.

Conflict of interest: none declared.

REFERENCES

4. Seixas NS, Robins TG, Becker M. A novel approach to the characterization of cumulative exposure for the study of


