EVALUATING EXPOSURE-RESPONSE RELATIONSHIP IN 1,3-BUTADIENE AND LEUKEMIA STUDIES

EVANGELIA E. ANTONIOU1 and CHRIS KIRMAN2

1 Epicurus-Reviews, MetaAnalyses.com, Bilzen, Belgium
2 SciPinion, Bozeman, Montana, USA

Abstract

Objectives: 1,3-Butadiene (BD) exposure's link to leukemia is under regulatory scrutiny. The assessment methods for BD exposure risks have evolved from early animal and limited human studies to advanced exposure-response modelling with comprehensive quantitative data. The objective of this study is to explore the nuances of exposure-response modelling, investigating how various statistical methods have influenced the quantification of exposure-response relationships. Material and Methods: Although this study was not conducted as a formal systematic review, a search was performed in Medline/Pubmed to identify all human studies on leukemia risk assessment for BD exposure. This search included articles written in English. The electronic search spanned from inception of records until July 23, 2023, using the search term: "butadiene AND (leukaemia OR leukemia OR myeloid OR lymphoid)" and was restricted to human species. Focusing on the synthetic styrene-butadiene rubber (SBR) industry cohort study conducted by the University of Alabama at Birmingham, USA, this review evaluates various statistical models and factors influencing exposure-response modelling. Results: Peak exposures to BD may be more influential in the dose-response relationship than cumulative or long-term exposure. The authors recommend utilizing β-coefficients derived from the latest SBR study update, employing Cox proportional hazard modelling, non-lagged and non-transformed cumulative BD exposure, and adjusting for age and peak BD exposure. The study reveals that statistical model selection has a limited impact on the calculated dose-response effects. The significant variation in estimated cancer mortality values arises from additional assumptions needed for metrics like the excess leukemia risk or the occupational BD effective concentration. Conclusions: In conclusion, this study provides insights into exposure-response modelling for BD exposure and leukemia mortality, highlighting the importance of peak exposures. The recommended statistical approach offers a reliable basis for regulatory risk assessment and public health population metrics.

Key words: risk assessment, leukemia, exposure-response modelling, occupational cohort study, regulatory toxicology, 3-butadiene

INTRODUCTION

Over the years, investigations of the association between 1,3-butadiene (BD) exposure and the risk on leukemia and other cancers have drawn significant attention from various regulatory bodies [1–3]. The methodology for evaluating the risk of such health outcomes of BD exposure has undergone substantial development in recent decades. Initially, these models were based exclusively on animal studies or human studies with limited and generally qualitative exposure information [4]. In more recent years, progress in exposure-response modelling and the availability of comprehensive quantitative exposure data in humans has facilitated the formulation of risk assessments that account for potential health impacts on human populations exposed to BD. Four occupational cohort studies, conducted in the USA, have reported on the association between BD exposure and cancer mortality risk. Nevertheless, all regulatory bodies have based their cancer mortality risk assessment conclusions on the same occupational cohort study on workers.
exposed to BD as conducted by the University of Alabama at Birmingham (UAB), USA, in their epidemiological study of North American workers in the synthetic styrene-butadiene rubber (SBR) industry (hereafter called the SBR study). Several studies, which are described and cited below in the study selection, have been conducted using the SBR cohort data. The other 3 cohorts [4–8] have not been used often mainly due to lack of historical industrial hygiene sampling data [4–6,8], no available quantitative exposure data [4–6], smaller cohort sizes [4–6] and lack of long-term follow-up [7,8]. The SBR study, however, has been frequently updated throughout the years adding to the robustness of the dataset.

The SBR study data has been used multiple times by various researchers using diverse statistical approaches to quantify the dose-response relationship between BD exposure and leukemia mortality. The objective of this article is to explore the nuances of exposure-response modelling, with a particular focus on assessing how various statistical methods have influenced the quantification of exposure-response relationships differently.

MATERIAL AND METHODS

Sources

Although this analysis was not conducted as a formal systematic review, a search was performed in Medline/Pubmed to identify all human studies with an outcome on leukemia mortality and BD exposure. This search included articles written in English. The electronic search spanned from inception of records until July 23, 2023, using the search term: “butadiene AND (leukaemia OR leukemia OR myeloid OR lymphoid)” and was restricted to human species. Only occupational cohort studies were selected. Conversely, case series, cross-sectional studies, reviews, and animal studies were excluded. In a second phase, the authors further excluded studies that provided no quantitative exposure-response metrics, no leukemia mortality endpoint, and cohort studies in which no Cox regression survival analysis was used.

Methods for characterizing the exposure-response relationship for BD and leukemia from epidemiology data have included Poisson and Cox proportional hazards regression. Unlike Poisson regression, Cox regression optimally controls for age by considering each specific age rather than age categories. This also allows cumulative exposure (e.g., BD ppm-years) to be treated as a continuous variable instead of necessitating categorization. Since the SBR study adopted Cox proportional hazard models in its most recent updates, particularly post-1998, this study explores data from these hazard models derived from the years 1998 and 2009 of the SBR follow up.

Data collection

The authors examined the statistical analysis results for each publication, specifically focusing on the regression terms related to BD’s cancer potency, which is expressed through the metric of cancer slope factor or slope of the exposure-response relationship (i.e., β-coefficients). These analyses involved regressing cumulative BD (ppm-years) exposure against the time to die from leukemia. Cumulative exposures emphasize both exposure magnitude and duration. While alternative exposure metrics have been considered, no straightforward alternative models have emerged thus far. The impact of each statistical model was evaluated by:

- the cohort size and number of patients,
- whether lag years were included in the analysis,
- the leukemia type(s) assessed,
- if other BD exposure variables were included in the model,
- whether the exposure variables were transformed,
- which non-exposure covariate variables were factored into the regression model.

For each regression analyses, the authors noted the resulting β-coefficients of the BD exposure variable.
along with their standard errors (SEs). Unless otherwise specified, these β-coefficients represent an increased leukemia mortality risk per cumulative BD ppm-year after Cox regression analyses. Where relevant, the goodness-of-fit of the models (as expressed by the Akaike information criterion (AIC) [9] or the –2 log likelihood (–2LL) of each statistical model) was also assessed.

**Presentation of results**

After a brief overview of the SBR study, the authors will summarize the different statistical models used and their impact on the exposure-response relationship. This is achieved by discussing 6 crucial components in the exposure-response modelling procedure:

- the influence of time (with a focus on all-type leukemia),
- the selection of the exposure variable (concentrating on cumulative BD exposure and its transformations),
- the potential role of lag years (that is, whether or not recent exposures should be included),
- the influence of non-exposure and exposure covariates (with emphasis on the cumulative number of BD high-intensity tasks),
- whether all-type leukemia should be the preferred endpoint in BD exposure-dose modelling.

**RESULTS**

**Study selection**

The initial database search yielded a total of 187 publications. Among these studies, 28 publications were further screened for eligibility in the review. From this subset, 13 publications [10–22] with primary data based on the SBR updates after 1998 were considered for inclusion in the review. Ultimately, eight publications, which presented exposure-response data utilizing Cox proportional hazard model data derived from the SBR study of 1998 and 2009 were included in this impact analysis [10–17].

**SBR study description**

Among the first studies initially conducted within SBR production, one study consisted of 2756 workers from a 2-plant complex in Texas, USA [23] and a second group of studies included about 13 000 workers from 8 facilities across the USA and Canada studied by researchers at John Hopkin’s University (JHU), USA [24–27]. Later, a team of researchers from the UAB expanded on the number of workers, by investigating the same 2-plant complex in Texas plus 7 of the 8 plants studied by the JHU extending the workers’ follow up until 1991 with 15 649 workers [28]. The UAB research group continued to update the cohort’s follow up, enhancing their assessment of BD exposures and addressing potential confounding co-exposures in their analyses [29–31]. Subsequently, the refinement of exposure assessments was coupled with an expansion in the cohort’s size including follow up extensions until 1998 [13–20], 2002 [21] and the latest follow up update extending to 2009 [10–12,22]. In the 1998 update the SBR study comprised of 17 924 workers measuring 120 leukemia decedents. Starting from 2002, 4861 female workers were added to the cohort [21]. In its most recent update (2009), the SBR study comprises 22 785 (17 924 men and 4861 women) cohort members with a follow up of 65 years and 132 leukemia decedents (Table 1).

**The exposure-response effect of cumulative BD exposure in different follow-up rounds**

Two statistical models were tested across different follow-up periods of the SBR study (Table 2). These were Cox regression models in which the relationship between cumulative BD exposure (in ppm-years) was related to the mortality risk of leukemia. One model adjusted only for age and the other model adjusted for age and the cumulative number of BD high intensity tasks (HITs), i.e., tasks with exposures ≥100 ppm BD, also called peak exposures. When correcting for age only, the increase leukemia mortality risk per ppm-year has remained remark-
Interestingly, in this latest dataset, this exposure-response relationship for cumulative exposures was also no longer statistically significant. The SEs estimates remained stable too, regardless of follow-up or covariate adjustment. This outcome highlights the importance of peak exposures versus chronic cumulative exposures and for considering adjustment for peak exposures in dose-response analyses of this kind.

Alternative transformations of BD exposure

Two studies [11,17] examined the influence of transforming cumulative BD exposure through logarithmic or square root conversions (Table 3). In both cases, the findings indicate that the model’s fit remained relatively consistent, regardless of the specific type of transformation applied. Furthermore, Cheng et al. [17] also showed that the categorization of continuous BD exposure into deciles did not notably enhance the statistical fit, concluding that the untransformed BD variable in its full range and continuous form was preferable. This observation implies that for the use of untransformed cumulative exposure to BD in ppm-years is advisable, for the sake of model accuracy and ease of interpretation.
EXPOSURE-RESPONSE RELATIONSHIP IN BD AND LEUKEMIA

ORIGINAL PAPER

Table 3. Model fit of models on increased leukemia mortality risk per cumulative 1,3-butadiene (BD) ppm-year based on Cox regression analyses, adjusted for age, in different styrene-butadiene rubber publications

<table>
<thead>
<tr>
<th>Transformation</th>
<th>Model fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>1384.12 (−2LL)</td>
</tr>
<tr>
<td>logarithmic</td>
<td>1378.77 (−2LL)</td>
</tr>
<tr>
<td>square root</td>
<td>1346.64 (−2LL)</td>
</tr>
<tr>
<td>none</td>
<td>2007.44 (AIC)</td>
</tr>
<tr>
<td>logarithmic</td>
<td>1998.24 (AIC)</td>
</tr>
<tr>
<td>square root</td>
<td>1998.89 (AIC)</td>
</tr>
</tbody>
</table>

−2LL = −2 log likelihood (a lower value indicates a better model fit);
AIC = Akaike information criterion (a lower value indicates a better balance between model complexity and model fit).
Exposure may or may not be converted to square root or logarithmic scales.

Differences in outcome based on ceiling levels of BD exposure

Sielken and Valdez-Flores [14] investigated the effect of ceiling levels for cumulative BD ppm-years. A ceiling of e.g., ≤100 ppm-years means that any cumulative exposure that is >100 ppm-years is disregarded (Table 4). The β-coefficients in this study seem to get larger as the restriction becomes tighter. Although approached differently, this finding also suggests that peak exposure to BD may have a clearer dose-response association with leukemia mortality than long duration exposures of low intensity. However, the SEs increased and consequently the statistical significance for leukemia diminished when the modeling was restricted to person years with an accumulation of ≤300, ≤200 or ≤100 cumulative BD ppm-years.

Impact of lagged exposures

Sielken and Valdez-Flores [13] presented a Cox regression model for the dose-response association between cumulative BD exposure (ppm-years) and leukemia mortality, with adjustments for age and peak exposure across various lag years (Table 5). The cumulative BD exposure is lagged when both exposures that took place 40 years ago and the most recent exposures (such as those within the last 5 years, 10 years or 20 years) are not included in the analyses. The β-coefficient for these analyses, without taking lag years into account, was 0.00022 (Table 2). One would expect the β values to increase with increasing lag time, however, Table 5 shows that the modeling of lagged exposures did not influence this β-coefficient substantially and may therefore not be needed. Similarly, in the study by Valdez-Flores et al. [12] (that did not adjust for BD HITs) models with lagged BD ppm-years did not outperform the model with unlagged BD ppm-years at the 5% significance level.

Table 4. Beta-coefficients of an increased leukemia mortality risk per cumulative 1,3-butadiene (BD) ppm-year based on Cox regression analyses, adjusted for age [14]

<table>
<thead>
<tr>
<th>BD exposure</th>
<th>β</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤1338 ppm-years</td>
<td>0.00121</td>
<td>0.00036</td>
</tr>
<tr>
<td>≤1000 ppm-years</td>
<td>0.00145</td>
<td>0.00047</td>
</tr>
<tr>
<td>≤500 ppm-years</td>
<td>0.00296</td>
<td>0.00086</td>
</tr>
<tr>
<td>≤400 ppm-years</td>
<td>0.00283</td>
<td>0.00111</td>
</tr>
<tr>
<td>≤300 ppm-years</td>
<td>0.00305</td>
<td>0.00155</td>
</tr>
<tr>
<td>≤200 ppm-years</td>
<td>0.00089</td>
<td>0.00267</td>
</tr>
<tr>
<td>≤100 ppm-years</td>
<td>0.00224</td>
<td>0.00536</td>
</tr>
</tbody>
</table>

Table 5. Beta-coefficients of an increased leukemia mortality risk per cumulative 1,3-butadiene (BD) ppm-year by exposure lag time based on Cox regression analyses, adjusted for age and BD high intensity tasks (HITs) [13]

<table>
<thead>
<tr>
<th>Lag time</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.00023</td>
</tr>
<tr>
<td>10</td>
<td>0.00023</td>
</tr>
<tr>
<td>15</td>
<td>0.00026</td>
</tr>
<tr>
<td>20</td>
<td>0.00027</td>
</tr>
</tbody>
</table>
Leukemia as the best potential blood cancer endpoint

Leukemia is a broad term encompassing a group of blood cancers that can be classified into different subtypes based on various factors, including the type of blood cells affected (lymphoid or myeloid). Three studies presented Cox regression models that were adjusted for age and peak exposure to BD HITs across leukemia subtypes (Table 7). The findings indicate that the estimated exposure-response effects for chronic lymphoid leukemia (CLL), chronic myeloid leukemia (CML) and myeloid neoplasms in general are comparatively lower than those for all leukemia combined (Table 1). Therefore, leukemia appears to be the preferred endpoint for regulatory considerations due to its heightened sensitivity as an outcome as well as the larger number of cases available. That being said, the “all leukemias combined” category may not be the most biologically aligned endpoint given that the leukemia subtypes have unique histological features [37].

At present, the SBR study possesses information regarding leukemia mortality but lacks data on leukemia incidence. It would have been more advantageous to have

### Table 6. Beta-coefficients of an increased leukemia mortality risk per cumulative 1,3-butadiene (BD) ppm-year from Cox regression analyses, adjusted for age and other covariates, in different styrene-butadiene rubber publications

<table>
<thead>
<tr>
<th>Non-exposure covariate adjustment in addition to age</th>
<th>β</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>years since hire</td>
<td>0.00029</td>
<td>0.00011</td>
</tr>
<tr>
<td>plant</td>
<td>0.00026</td>
<td>0.00012</td>
</tr>
<tr>
<td>years since hire</td>
<td>0.00039</td>
<td>0.00012</td>
</tr>
<tr>
<td>plant</td>
<td>0.00029</td>
<td>0.00011</td>
</tr>
<tr>
<td>years since hire</td>
<td>0.00029</td>
<td>0.00010</td>
</tr>
<tr>
<td>race</td>
<td>0.00026</td>
<td>0.00012</td>
</tr>
<tr>
<td>plant</td>
<td>0.00039</td>
<td>0.00012</td>
</tr>
<tr>
<td>calendar year</td>
<td>0.00028</td>
<td>0.00010</td>
</tr>
<tr>
<td>Sathiakumar et al. 2015 [11] – follow-up 2009: race, plant</td>
<td>0.00029</td>
<td>0.00010</td>
</tr>
<tr>
<td>Sathiakumar et al. 2021 [10] – follow-up 2009: race, plant, sex, age at hire, year of hire, ever hourly status</td>
<td>0.00026</td>
<td>0.00010</td>
</tr>
</tbody>
</table>

### Adjustment of non-exposure covariates

The SBR study included 6 non-exposure variables (i.e., age, years since hire, calendar year, sex, race, and plant) that can be tested for their potential confounding influence on the β-coefficients. Valdez-Flores et al. [12] examined whether the inclusion of these additional covariates improves the model fit after adjustment for age and BD HITs but did not observe a statistically significant improvement at a 1% significance level. The effects on the β-coefficients due to residual confounding of other covariates after adjustment for only age can be found in Table 6. Comparison to the β-coefficients of Table 1 (adjustment for age only) shows that additional adjustment for non-exposure variables does not alter the effect estimates and thus may not be needed. The SBR study had no information on the cigarette smoking habits of the workers, while it is likely that 50–80% of the workers may have smoked [32] and smoking itself is a source of non-occupational BD exposure [33]. Although leukemia has not been strongly linked to smoking [34,35], to the extent smoking remains as a possible independent risk factor for leukemia [36] the β-coefficients may be biased upwards if correction for smoking is not performed.
incidence data, as mortality is not solely determined by exposure but also, by the treatment choices accessible at any given point in time.

While bladder cancer is not a blood cancer, it has recently emerged as a heightened risk linked to BD exposure in the latest cohort update [12], with a comparable effect size to that of leukemia. This study estimated an increased mortality risk per ppm BD exposure per year to be 0.00028 for leukemia (Table 2) and 0.00031 for bladder cancer after both models were adjusted for age.

**DISCUSSION**

This impact analysis has provided an overview of the different statistical models that were employed in the SBR studies to estimate the exposure-response relationship between cumulative exposure to BD and leukemia mortality risk. The authors’ conclusion is that the variation in calculated exposure-response effects (β-coefficients) across different regression models, papers and cohort updates is relatively modest.

The authors’ recommendation to those who will use these β-coefficients as input for public health population metrics, such as the ELR or the occupational BD effective concentration (OEC) is as follows: use a β-coefficient that is based on the latest and largest update of the SBR study, is derived from Cox regression modelling for the dose-response relationship between non-lagged and non-transformed cumulative BD exposure (ppm-years) and leukemia mortality, with adjustment for age and peak exposure to BD (Table 8). This model is parsimonious, provides easy interpretation, has a good model fit and is one of the models with the largest statistical power. Valdez-Flores et al. [12] provided such a β-coefficient [38], i.e., 0.000132 (Table 2). Due to the relative stability of β-coefficients, the large variation found in estimated ELRs or OECs is likely attributed to the additional assumptions required for calculating these values [38–40]. These assumptions include life-table data like time and region-specific data on survival probabilities, mortality rates across age-ranges of the outcome of interest, the choice between assuming daily and lifelong exposure to BD (environmental exposure) or only daily exposure during ones’ working life (occupational exposure) and the decision of whether to use the maximum likelihood estimate of the β-coefficient itself as an input and/or the upper limit of 95% confidence interval (CI) of the β-coefficient, as some regulatory bodies have done.

Given the same β-coefficient, the life-table data will vary depending on the age-distribution and general health care and specific cancer treatment options available at a specific time and place. As the treatment options for
leukemia are continuously improving [41], the likelihood to die from leukemia has decreased. Even if all else would have remained stable, this means that the mortality and thus the ELRs must have decreased over time as well. This article highlights the crucial role of peak BD exposures in the modelling process. Here, the authors propose the hypothesis that peak exposure may be more important than long-term and cumulative exposure to BD. This is supported by the following observations: The studies by Sielken and Valdez-Flores [13,15,16] and Valdez-Flores et al. [12] found that the cumulative number of BD HITs is a more important predictor of leukemia mortality than cumulative BD ppm-years. Their statistical models focused on the relationship of leukemia mortality and BD ppm-years, while adjusting for BD HITs, because the general population is not expected to be exposed to BD peaks (Table 2). Sielken and Valdez-Flores [14] showed that restricting to person years with ≤300, ≤200 or ≤100 cumulative BD ppm-years, a low-exposure risk vs. cumulative BD ppm-years was observed (Table 4). Furthermore, Valdez-Flores et al. [12] found that the value of the β-coefficient for cumulative BD exposure more than halved after adjusting for BD HITs. Similarly, Cheng et al. [17] showed that all leukemia decedents in the SBR study exposed to BD had some peak exposures to BD, while no leukemia mortalities at all were found among exposed workers without peak exposures. Similarly, Cheng et al. [17] also found a positive exposure-response relationship between BD HITs alone and leukemia mortality risk. Given the apparent importance of peak exposures, future studies might consider modelling this variable independently as a predictor in the occupational settings, instead of relying solely on cumulative BD ppm-years.

CONCLUSIONS

This review offers evidence to endorse the utilization of the latest SBR study update. This study boasts a large cohort, longitudinal exposure quantification of 47 years, inclusion of both genders and a 65-year follow-up period covering over 50% of the cohort. Thus, it has the highest statistical power of available cohort studies to estimate potential effects of BD exposure. By acquiring and utilizing data from such high-quality human study available, there is no need for exposure-response extrapolations from animal studies for cancer endpoints.

The recommended approach involves employing the β-coefficient derived from Cox proportional hazard modelling for the relationship between non-lagged and non-transformed cumulative BD exposure and leukemia mortality, with adjustment for age and peak exposure to BD.

Conflict of interests

The sponsors had no role in the design, execution, interpretation, or writing of the study. The opinions expressed herein are those of the authors, not of their employers or funders.

Author contributions

Research concept: Evangelia E. Antoniou, Chris Kirman
Research methodology: Evangelia E. Antoniou, Chris Kirman
Interpretation of results: Evangelia E. Antoniou, Chris Kirman
References: Evangelia E. Antoniou

REFERENCES

13. Sielken RL, Valdez-Flores C. Quantitative Risk Assessment of Exposures to Butadiene in European Union Occupational Settings Based on the University of Alabama at Birmingham Epidemiological Study: All Leukemia, Acute Myelogenous Leukemia, Chronic Lymphocytic Leukemia, and Chronic Myelogenous Leukemia. Olefins Panel of the American Chemistry Council (ACC); 2008.
39. National Research Council (NRC). Health Risks of Radon and Other Internally Deposited Alpha-Emitters, Committee on the Biological Effects of Ionizing Radiation, Biological...
