

Event-Based Risk Modeling of Benzene Exposure: Implications for Acute Myeloid Leukemia Pathogenesis and Prevention

Dr. D. M. Alvarez¹, Dr. J. R. Becker², Dr. L. Petrov³, Dr. N. S. Rahman⁴, Dr. E. Kowalska⁵, Prof. T. H. Bennett^{1*}

¹ Department of Occupational and Environmental Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

² Institute of Epidemiology and Public Health, University of Munich (LMU), Munich, Germany

³ Department of Clinical Toxicology, Sechenov First Moscow State Medical University, Moscow, Russia

⁴ Division of Community Medicine and Global Health, University of Malaya, Kuala Lumpur, Malaysia

⁵ Department of Public Health and Preventive Medicine, Medical University of Warsaw, Warsaw, Poland

Abstract

Occupational exposure to benzene at levels of 10 ppm or more has been associated with increased risk of acute myeloid leukaemia (AML). The mode of action (MOA) for AML development leading to mortality is anticipated to include multiple earlier key events, which can be observed in hematotoxicity and genetic toxicity in peripheral blood of exposed workers. Prevention of these early events would lead to prevention of the apical, adverse outcomes, the morbidity and mortality caused by the myelodysplastic syndrome (MDS) and AML. Incorporation of key event information should modify the risk model, but few modification approaches have been suggested. To that end, two approaches to risk model modification are described that use sub-linear and segmented linear increases in risk below key events, while maintaining a linear increase in AML mortality risk beginning at 2 ppm, the lowest observed adverse effect concentration (LOAEC) identified for hemato- and geno- toxicity in high quality studies of human occupational exposure. Below 2 ppm two different modification approaches to quantitative risk models were applied: a continuously decreasing slope model and a segmented modification in slope. These two approaches provide greater flexibility to incorporate MOA information in risk model development and selection.

Background

Benzene is a naturally occurring hydrocarbon present in petroleum and some petroleum-derived products. Its potential to cause leukaemia and bone marrow toxicity in humans is known. Owing to benzene's hazard potential, different governmental and non-governmental entities globally have recommended limiting environmental and occupational benzene exposures. These recommendations vary depending on the population of interest, the method used for exposure limit development, and other factors used by the limit setting entity.

Over the course of 2018 and 2019 there were substantial scientific discussions in the European Union focused on recommendation of an occupational exposure limit for benzene. The Committee for Risk Assessment (RAC) of the European Chemicals Agency concluded there was no significant residual cancer risk and that other adverse health effects were avoided when occupational exposures were below 0.05 ppm 8 hour time-weighted average (8h TWA) (Committee for Risk Assessment, 2018). In later discussion the Advisory Committee for Safety and Health at Work considered both the RAC opinion and the benefit cost analysis, recommending a progressive reduction in the EU benzene OEL from 1 ppm to 0.5 ppm 8h TWA, then within two years following that decrease a second decrease to 0.2 ppm (Advisory

Committee on Safety and Health at Work, 2019), the same as the level recommended by the Dutch Expert Committee on Occupational Safety (2014). In parallel, the Lower Olefins and Aromatics (LOA) consortium, which manages the industry REACH benzene dossier for EU joint registrants, made an independent evaluation that supported a 0.25 ppm 8h TWA OEL. Both RAC and LOA proposals recommend a health-based limit that assumes negligible risk of health effects below the OEL. However, for benefit cost analysis it is important to define a risk function, not only a safe limit, so risk managers can make informed decisions about the balance of societal, environmental, health, and economic benefits and costs for implementing a limit.

Dose-response models commonly used in risk assessment to estimate effects are linear or curvilinear. The selected model frequently depends on the information available for identifying potential regions of the dose response where increasing exposure has little effect on response (i.e., exposures above the maximum effect size [plateau] and below the minimal effect size [threshold]). In occupational epidemiology it is rare to have information to identify plateaus, and thresholds are frequently debated due to limitations of statistical inference for vanishingly small effect sizes. As a consequence, it is common to see linear-no-threshold (LNT) models applied in estimating risk (i.e., zero exposure = zero risk), because in the context of risk assessment they have frequently been thought unlikely to underestimate risk. Implicit in that judgment is that they become likely to overestimate risk if the true response has a threshold. The appropriateness in general of such models has been reviewed elsewhere (Calabrese and Golden, 2019)..

Some entities have considered alternatives to LNT models for benzene (AGS, 2012; Dutch Expert Committee on Occupational Safety, 2014). Spline models are widely used alternatives to simple linear models in regression to better fit and predict responses. The statistician defines knots to introduce flexibility to the shape of the regression model. Stated simplistically, knots enable the statistical model to have multiple joined segments, with each knot being a point of change in the shape of the response and each segment having a different slope. However, the choices for number of knots and their placement are not always simple, especially with datasets limited to relatively few observations (e.g., <1000). Limited data set sizes make defining spline models for benzene responses challenging, as available data are often limited to tens of cases of leukaemia with few observations to inform the placement of knot. While not a spline approach, the concept of a kinked function may also have utility. The AGS opinion considered placing a “kink” at 0.42 ppm based on the 95th % lower limit benchmark dose hematotoxicity, reasoning that above that point an effect enhancement would likely lead to a higher slope for the risk function. In the end AGS decided against using the resulting sublinear model for a final risk model due to uncertainties related to mechanism and quantification in the selected study used for kink placement. Nonetheless, an approach that utilizes key events preceding the final adverse outcome to select points of change (kinks) in a risk model could allow assessors to incorporate information from multiple studies in a manner that increases confidence in the subsequent risk model, even if the precise risk cannot be observed empirically.

In the case of benzene, evaluation of the literature indicates there are biological effects that occur at earlier time, and lower exposures, than AML. Hematotoxicity and genotoxicity following occupational benzene exposure are reported in multiple independent studies. In the case of hematotoxicity, significant decreases in white blood cell counts can be observed within months of initial occupational exposure depending on exposure intensity (Cody et al., 1993), and are inversely correlated with average annual benzene air levels for exposed workers (Kipen et al., 1988)... Less literature characterizes the time course of genetic toxicity, but workers with a mean employment of 6.9 ± 4.9 years (0.7 to 16.5 years) at a 72 ppm average benzene exposure had significant increases in glycophorin A mutations, indicative of genotoxicity to bone marrow (Rothman et al., 1995).. Evaluation of the database of high quality studies

for both hematotoxicity and genetic toxicity suggest effects are observable at exposures of approximately 2 ppm (Schnatter et al., 2020), events which are both anticipated in benzene's MOA to precede leukaemia (North et al., 2020).

Here we applied risk modelling for benzene leukaemogenicity by considering that risks at exposures equal or greater than that causing hematotoxic or genotoxic effects (2 ppm) can be modelled using a LNT approach, but below that point a decreased leukaemia risk is more reflective of real world experience. We present two approaches to adjust the risk model below 2 ppm: continuously decreasing slope (CDS) or a segment modified slope (SMS) between 2 ppm and the 0.25 ppm OEL proposed by the LOA consortium.

Rationale

Multiple studies, in both humans and laboratory animals, support the carcinogenicity of benzene. Given that cancer risk modelling is generally performed for humans, and there is a body of scientific evidence to characterize human leukaemogenicity of benzene for OEL development, this review focuses on the human studies. Laboratory rodent studies collectively indicate that benzene exposure can result in dose-related increases in multiple tumour types in both rats and mice, while epidemiologic studies have reported excesses acute myeloid leukaemia (AML) and myelodysplastic syndrome (MDS), with other subtypes considerably less consistent.

In addition to the form of statistical models (LNT, thresholds, linear, non-linear), benzene carcinogenicity models can be based on different cancer types and on different latency or lag periods. These are critical aspects of carcinogenicity models which, if misspecified, can lead to spurious results. In addition, while recent reviews have suggested that basing occupational exposure limits on hematotoxicity and/or genotoxicity protects against benzene-induced carcinogenicity, review of exposure-response data for cancer is also warranted to either verify or refute that position. Below we briefly review each aspect (cancer type, lag period and exposure-response) as a basis for the preferred carcinogenicity models.

Cancer type

The initial evidence of leukaemogenicity emerging in the 1970s and 1980s was inadequate to draw strong conclusions regarding likely causal relationships between benzene and leukaemia subtypes. More recently IARC (2018) evaluated evidence for multiple subtypes, concluding benzene causes AML in adults. They also noted associations have been observed for non-Hodgkin lymphoma (NHL), chronic lymphoid leukaemia (CLL), multiple myeloma (MM), chronic myeloid leukaemia (CML), AML in children, and cancer of the lung, but none of these subtypes were characterized by definitive evidence necessary to make a causal assertion. Causal effects between benzene and AML (in the absence of causal effects due to other cell types) are confirmed by other systematic reviews (Schnatter et al., 2005) focused on leukaemia cell type. However, until recently, few studies have provided information which is useful for quantitative risk assessment (Vlaanderen et al., 2008).

For MDS, the International Agency for Research on Cancer (2018) made no formal assessment, though recent studies (Copley et al., 2017; Linet et al., 2019; Schnatter et al., 2012) have found positive relationships between benzene and MDS, while Li and Schnatter (2018) considered that the evidence was sufficient to indicate causality. Yet few studies have used quantitative risk estimates that allow formal statistical modelling. Thus, MDS was not considered further in terms of quantitative risk projections.

In the benzene studies with leukaemia subtype information that report a statistically significant excess of total leukaemia, the excess is usually due to AML or ANLL (Crump, 1994; Glass et al., 2005; Linet et al.,

2019; Rhomberg et al., 2016; Wong et al., 2010). Less frequently, but consistent with dose-response and statistical significance, is MDS identified (Copley et al., 2017; Linet et al., 2019; Schnatter et al., 2012). This may further emphasize the myeloid character of benzene leukaemogenicity, as a significant fraction of MDS cases progress to AML. It is plausible that the aetiology of benzene-induced AML may include a subclinical stage manifesting as MDS before it progresses to AML.

While some argue that risk assessment should be based on total leukaemia since it is possible that other leukaemia cell types are related to benzene, our preference is to base models on proven causal relationships. If we call AML cases “x” and non-AML cases “y”, and aim to examine whether “y” has a relationship with benzene, it is fruitless to examine the relationship between total leukaemia (x+y) and benzene, which is frequently done. Models that are based on all leukaemia, rather than AML, are much more likely to include cases not related to benzene exposure. This would, in effect, assign incidental, likely lower exposure to benzene’s potency estimate, obscuring empirical thresholds. These non-AML lower-exposed cases, combined with the more highly exposed AML cases, give the misleading appearance of a smooth dose response for total leukaemia, although it is obvious that two different signals are being assessed when the AML dose response is compared to the non-AML dose response. In order to illustrate, Crump (1994) fit several statistical models to the Pliofilm cohort data. Despite similar results (in terms of potency estimates) for all leukaemia versus AML only, there was no relationship between benzene and non-AML leukaemias. Thus, using the Pliofilm data on total leukaemias could artificially result in similar potency estimates (rather than ‘dilute’ the AML dose response), and would unequivocally obscure empirical thresholds. In addition, it would attribute more leukaemia cases to benzene exposure, despite the lack of a dose-response relationship for the non-AML cases.

Lymphohematopoietic neoplasms other than forms of, or precursor conditions related to, AML (i.e., NHL, MM, and CLL) are inconsistently observed in association with benzene when considered in the light of dose-response or statistical significance (Steinmaus et al., 2008; Alexander et al. 2010). The difference in relative strength of evidence may be highlighted in the most recent IARC evaluation of benzene in 2018 (International Agency for Research on Cancer, 2018), in which they reported “the findings fully supported the previous conclusion that benzene causes ANLL – including AML – in adults, as well as the previous observations of limited evidence for chronic lymphocytic leukaemia, non-Hodgkin lymphoma, and multiple myeloma.” Other neoplasms (CML and lung cancer) were also considered by IARC, but also judged to be limited with regard to evidence. Given the strength of evidence favours association of benzene and AML, our focus is on dose-response assessment for AML.

Lag Period

Time windows of exposure may have important implications for exposure response modelling of cancer risk from benzene. If benzene cancer risk is attributable to primarily recent (<10 years) exposure, then models that treat exposure from all time periods with equal weight may misestimate the risk. However, it may remain relevant for benefit estimation for policy makers, and thus may be a topic worthy of future research.

One noteworthy study of the lag period is Richardson (2008), who examined cases in the Pliofilm cohort (Rinsky et al., 2002). The analysis reported that a model employing three exposure time windows (<10 years, 10 - <20 years, and 20+ years) fit the data much better than a lifetime cumulative risk model. The three periods showed that leukaemia risk was significantly high in the most recent 10 years prior to death or withdrawal (RR = 1.19 per 10 ppm-years) and was somewhat raised but not significantly in the 10-20 year exposure window (RR = 1.05 per 10 ppm-years). Exposure received in an exposure window of more than 20 years showed no association with leukaemia.

Hayes et al (1997) reported similar results when studying benzene-exposed workers from several industries in China. Notably, the risk of ANLL/MDS was significantly related to benzene exposure in a 1.5 to 10 year exposure window previous to diagnosis ($p=0.003$), but showed no relationship to “distant” exposure more than 10 years prior to diagnosis ($p=0.51$). More recently, Linet et al., (2019) confirmed this finding, reporting that exposures within the 10 years prior to diagnosis were more strongly associated with MDS/AML, (P_{trend} values of 0.08 and 0.06 for cumulative and average intensity exposure measures, respectively) versus exposures ≥ 10 years prior to diagnosis ($P_{\text{trend}} = 0.92$ and 0.42).

Schnatter et al (2012) did not find enhanced risk for recent benzene exposures for neither AML nor MDS, although this may have been due to much lower benzene exposure levels encountered in the populations studied.

The evidence indicates that benzene exposure risk is likely limited to 10-20 years before the diagnosis of AML. When time increases beyond 20 years of first exposure, there appears to be no excess risk for AML due to benzene exposure.

Exposure Response

Overall consideration of the carcinogenicity studies with regard to dose-response may be informative. Studies with quantitative exposure analysis reporting statistically significant excess AML mortality (Pliofilm and NCI-CAPM) did not identify statistically significant effects until average exposures were >10 ppm. Studies that did not find statistically significant excess AML mortality (pooled petroleum worker study, Monsanto, and Dow) were generally populations with average benzene exposures <10 ppm. The pooled petroleum worker study did identify a statistically significant increase in MDS for cumulative exposures >2.93 ppm-years, but not for the highest average exposure intensity (>0.259 ppm), though in the case of MDS the inclusion of both peak exposures >3 ppm with other exposure metrics was suggestive that peak exposures >3 ppm may be more influential for risk than cumulative or average exposures. The Shanghai Health Study used semi-quantitative maximum exposure groupings, identifying significant excess for AML for the highest exposed group, for which the lower limit in the category was ~ 3.1 ppm, but would have included much higher exposures as well. Considering the above, it appears there is a clear signal for AML for benzene exposures above 10 ppm and may be considered equivocal above 1 ppm. It is a given there are substantial uncertainties regarding the shape of the dose-response for benzene below 10 ppm.

The vast majority of epidemiologic occupational studies that measure exposure do so by estimating past or recent exposures and apply exposures to time periods to arrive at a concentration x time or cumulative exposure metric, often in ppm-years. Inherently, this calculation assumes that the manner in which a given quantity (in ppm-years) is attained is inconsequential – e.g. 10 ppm for 1 year is indistinguishable from 1 ppm for 10 years. However, for benzene exposure, there is conflicting data for this assumption (Collins et al., 2003; Schnatter et al., 2012; Schnatter et al., 1996; Seniori Costantini et al., 2003; Wong, 1987b) The data more often suggest that high concentrations are more impactful for subsequent risk from exposure. Thus, using cumulative exposure in risk models is likely to be a conservative strategy, in that the risk estimates from such calculations will result in a higher risk per unit of exposure.

REVIEW OF STUDIES INCLUDED IN AML EXPOSURE RESPONSE ASSESSMENT

Pliofilm Cohort (Multiple studies)

The Pliofilm cohort is amongst the most studied populations of benzene-exposed workers (1,845 workers), and is possibly the most frequently used cohort for benzene quantitative risk assessment. Occupational exposures resulted from benzene used as a solvent in the production of film rubber from the late 1930s to 1976 in three facilities in Ohio, USA. The most recent update to cohort mortality (Rinsky et al., 2002), representing 53% (976 workers) of the cohort, indicated a SMR of 1.64 (95% CI 1.06-2.42) for lymphatic and hematopoietic cancers in the all workers group. Subgrouping based on tumour lineage was performed for leukaemia (SMR 2.47, 95% CI 1.38-4.07), MM (SMR 2.04, 95% CI 0.66-4.76), and NHL (SMR 0.96, 95% CI 0.31-2.25). Limiting analysis to white male workers only (90% of the cohort) results in marginally higher SMR and 95% CI values. Concentration response analysis was reported for leukaemias in white males, indicating a monotonically increasing SMR with benzene exposure. Twelve of 15 leukaemias in benzene-exposed workers were myeloid lineage. The two of three non-myeloid leukaemias were unspecified, and one was ALL. No less than four sets of exposure estimates have been developed for the cohort: (Rinsky et al, 1981); Crump and Allen 1984; Paustenbach et al., 1992; and Williams et al., 2003). While there is debate about the true exposures in this cohort (Utterback and Rinsky, 1995), there is some objective evidence that the Rinsky et al estimates may be too low (Kipen et al., 1986), while some of the Paustenbach estimates are likely to be too high (Williams and Paustenbach, 2003). In addition, the Crump and Allen estimates used a technique of increasing the Rinsky estimates in proportion to the TLV in effect at the time, a potentially overly simplistic technique. Schnatter et al. 1996 attempted to derive a more accurate estimate by using the median of the available estimates at that time (those predating the Williams and Paustenbach 2003 work), but these median estimates have not been used in risk assessments. Benefitting from several critiques over the years, more specific job history data, further monitoring results and applying more advanced probabilistic techniques, the Williams and Paustenbach (2003) estimates are likely the most accurate ones, although only a single study has employed these estimates (Rhomberg et al., 2016) in dose-response assessment. There are advantages to using the Pliofilm cohort since power was sufficient to show a dose-response relationship, there was sufficiently long follow-up, and there were few reported co-exposures to carcinogenic substances relative to other studies. All authors note a convincing dose-response for total leukaemia, however, there is marked heterogeneity in the relationship with benzene between different leukaemia subtypes.

NCI/CAPM China study (Hayes et al., Linet et al.)

A larger cohort study (74,828 exposed and 35,805 unexposed workers) was performed in collaboration between the United States National Cancer Institute and Chinese Academy of Preventive Medicine (Hayes et al., 1997; the NCI-CAPM cohort), with the most recent analysis published in 2019 (Linet et al., 2019). Quantitative benzene exposure estimates were reported in Hayes et al., (1997) and Linet et al., (2019), thus are the focus for this summary. In Hayes et al., the relative risk for leukaemia was significantly raised (RR 2.5, 95%CI 1.2 – 5.1). NHL was close to significance (RR 3.0, 95% CI 0.9 – 10.5). As with the Pliofilm study, there was no dose response trend for other leukaemia cell types. There was an excess of NHL only for some of the highest (> 25 ppm average exposure, > 100 ppm-years) exposure groups. MDS was also significant, but the RR was infinite, since there were no unexposed cases of MDS. With ANLL and MDS combined, the RR was significant at 4.1, with a strong dose-response relationship. Statistically significant increases in risk were reported for the middle category (10-25 ppm) of exposure. Similar to the Pliofilm study, the risk of ANLL/MDS was significantly related to benzene exposure in a 1.5 to 10 year exposure window previous to diagnosis (p=0.003), but showed no

relationship to “distant” exposure more than 10 years prior to diagnosis ($p=0.51$). NHL risk was seen for exposures more than 10 years from diagnosis ($P=0.005$), but not to more recent exposure ($P=0.15$).

The Linet et al., analysis focused primarily on combined MDS/AML, though some information on CML was also reported. For MDS/AML, a significantly increased hazard rate ratio (but referred to here as RR for simplicity) for cumulative exposure of 40-<100 ppm-years (RR 2.52, 95% CI 1.02 – 6.08) when considering all ages for first exposure without a time window was reported. RR was elevated, but not statistically significant, for cumulative exposures >100 ppm-years (RR 1.84, 95% CI 0.60 – 5.01). When analysed on the basis of average exposure no statistically significant increases in hazard rate ratio were identified, but RR was elevated for average exposure intensity of 5 ppm or greater (minimum was RR 1.68, 95% CI 0.64 – 4.21 and maximum was RR 2.59, 95% CI 0.97 – 6.57 among the exposure groups). For exposures of >0 to <5 ppm average intensity the RR was decreased, but not significantly (RR 0.80, 95% CI 0.31 – 1.97).

The potential influence of exposure time windows on RR was also assessed (Linet et al., 2019), finding exposures within the 10 years prior to diagnosis were more strongly associated with MDS/AML, and reporting P_{trend} values of 0.08 and 0.06 for cumulative and average intensity exposure measures, respectively. When considering exposures ≥ 10 years prior to diagnosis, the P_{trend} was non-significant (i.e., 0.92 and 0.42 for cumulative and average intensity exposure measures, respectively). Linet et al., further observed a potential difference in RR depending on the age at first exposure when further stratifying their analysis. The RR values are more than two times higher than those observed for the 40 to <100 and ≥ 100 ppm-year exposures groups for workers in the 2 to <10 year window and <30 years old at first exposure compared to workers with all ages and time windows. The RR values fluctuate from 0.78 to 1.22 without any clear exposure response character (P_{trend} is 0.55) when focused on cumulative exposures ≥ 10 year window for workers ≥ 30 years old at first exposure. The pattern is similar when considering average intensity of exposure.

Unlike MDS/AML, the pattern for CML in Linet et al., (2019) does not appear responsive to exposure. When limited to an exposed versus unexposed comparison, there is an increased (but not statistically significant) RR for CML (RR 2.64, 95% CI 0.76 – 11.41), but exposure response modelling fails to find a significant trend (P_{trend} values from 0.39 to 0.94 depending on the time window and exposure metric considered). The calculated risk coefficients for RR are negative, with confidence intervals that encompass zero, with a plausible interpretation being there is no benzene exposure response for CML in the NCI-CAPM cohort.

Thus, for the NCI-CAPM reports a clear dose-response for AML/MDS combined, as well as an indication that these tumours were related to relatively recent exposure window (from 1.5 to 10 years before incidence).

Stenehjem et al., 2015

By using the combination of the Cancer Registry of Norway and postal survey a cohort of 24,917 male Norwegian offshore workers were identified and compared for benzene exposure and incidence of various lymphohematopoietic cancers, finding 112 cases (Stenehjem et al., 2015). When evaluated on an ever/never exposed basis, neither statistically significant associations between benzene and lymphohematopoietic cancer, nor any subtype (including AML), were identified, though Hazard Ratio (HR) was elevated above 1 for multiple subtypes and groupings. Quantitative exposures were estimated based on a semi-quantitative assessment of activities likely to result in higher exposures. Dose-response analysis comparing unexposed workers with three increasing exposures tertiles was generally resulted in relatively flat responses, with HR values increased in the lowest tertile compared to unexposed, but

similar to the values observed for the highest tertile. Narrowing the subtype definition to myeloid cancers (which would combine AML, CML, MDS, and “other myeloid diseases”) generally increased the observed HR when comparing the same exposure tertiles, but was not consistently monotonically increasing, and trend tests did not identify a statistically significant effect. Even when narrowed to solely to AML the results were not statistically significant, whether evaluated on a pairwise or trend test basis. However, the AML analysis was closest to identifying a statistically significant trend ($p = 0.052$ using cumulative exposure, $p = 0.056$ when using average peak >3 ppm). The HR values, once adjusted for age, smoking, and benzene exposure from other work, in the highest tertile were 4.85 (95% CI 0.88-27; 0.124–0.948 ppm-years) for and 4.87 (95% CI 0.90- 26; quantitative metric not reported for average peak >3 ppm). A single statistically significant increase in HR for CLL was reported for the second tertile exposure group (adjusted HR 6.66, 95% CI 1.32-34), though this was not consistent with a monotonically increasing trend (adjusted HR values for the 1st and 3rd tertiles were 2.27 (95% CI 0.31-17) and 2.21 (95% CI 0.27-18), respectively). The only statistically significant trend was reported for MM when considered with cumulative exposure ($p = 0.024$), though the adjusted HR of 3.25 (95% CI 1.00-10) reported for the highest tertile included 1, and given the sheer number of statistical tests reported and no description of adjustment for multiple comparisons the result should be interpreted cautiously.

Pooled Petroleum Worker Study (Rushton et al., 2014)

The pooled petroleum workers cohort updated and combined three nested case-control studies of benzene-exposed workers, identifying 370 lymphohematopoietic cases matched to 1587 controls from Australian, Canadian, and the United Kingdom cohorts (Schnatter et al., 2012). One important aspect of the study is the potential difference in exposure when compared to the Pliofilm and NCI-CAPM, with the pooled petroleum workers study experiencing lower benzene exposures. The dose-response for several different leukaemia subtype (AML, CML, and CLL) and myeloid disorders (MDS and MPD) was assessed. Various metrics for exposure were considered, including cumulative, average, maximum, duration, and jobs with repeated peak exposures >3 ppm. Little evidence for a dose-response relationship was observed for AML, CML, CLL, or MPD, but a significant increase in MDS when comparing the lowest and highest cumulative exposure tertiles for benzene exposure. Peak exposure (>3 ppm at least weekly for at least a year) also showed a strong relationship with MDS. When modelling both peak and cumulative exposure, MDS risk was more robust for peak exposure, as the metric remained strong even in the presence of cumulative exposure. Overall, exposures for these workers were lower (mean exposure intensity of 0.2 ppm, mean maximum exposure intensity of 0.7 ppm or less, and mean cumulative exposure of 5.15 ppm-years with 90% of participants with <20 ppm-years) than those observed for Pliofilm and NCI-CAPM. It should be noted that in the original studies, there were some suggestions of an AML excess, particularly in the Australian (Glass et al., 2003 and Glass et al., 2005) studies. The pooled study identified 60 case (241 matched controls), but a weaker, non-significant trend for AML exposure (Rushton et al., 2014). The authors suggest that upon more stringent pathology review, some AML cases were re-classified to MDS and/or MPD diagnoses, which could account for the difference in results compared to previous updates. It should also be noted for MDS and AML, the authors examined a more recent exposure window, and results did not support a higher risk for this time frame.

Dow Study (Collins et al., 2015)

A cohort of 2,266 benzene-exposed workers in Midland, Michigan, USA have been the subject of several studies, most recently in 2015 (Bloemen et al., 2004) (Collins et al., 2015). The most recent update identified five acute myeloid leukaemias and five acute non-lymphatic leukaemia, of which four of five in

both cases had a latency >30 years. Total leukaemia SMR is not significantly increased in these workers, nor for any specific leukaemia subtype. Dose-response analysis suggested a monotonic increase for total leukaemia and total myeloid leukaemia, but trend test results were not statistically significant. NHL risk decreased with cumulative exposure but was not statistically significant. It may be noteworthy that exposures for the Midland cohort appear lower than for Pliofilm and NCI-CAPM, as the cohort average cumulative exposure was 35.1 ppm-years, implying a working lifetime exposure average around 1 ppm.

Monsanto Study (Collins et al., 2003)

A study of 4,417 benzene-exposed workers from a chemical plant in Sauget, Illinois, USA has been updated several times (Collins et al., 2003). When evaluated on a cumulative benzene exposure metric, total leukaemia, MM, NHL, CLL, and ANLL SMR above 1 for multiple groups, but were not statistically significant for any groups. SMRs were increased in a near-monotonic manner for total leukaemia and ANLL (7 cases versus 4.9 expected), but not MM or CLL. SMRs were not significantly increased when considering peak benzene exposures as number of work days with exposures >100 ppm, and monotonic increases were not observed for any of the reported lymphohematopoietic neoplasms. However, the authors did conclude “number of peak exposures greater than 100 ppm to benzene [was] a better predictor of risk than cumulative exposure”, an interesting observation in the context of how the relationship between peak and cumulative exposure and health effects informs the likelihood of an effect threshold. Further, it may be noteworthy that the highest exposure group, >6 ppm-years, implies lower cohort average exposures than that observed in Pliofilm or NCI-CAPM.

In total, there is consistent evidence from epidemiology studies for benzene increasing the risk for AML. MDS is more plausible than any other non-AML potential subtype, but has not been as frequently evaluated. Both AML and MDS are myeloid in nature, and it is plausible AML and MDS are etiologically related if caused by benzene, with MDS being an earlier or marginally milder dysfunction in the event cascade that includes AML and aplastic anaemia at higher exposures or longer time periods. This myeloid character may initiate with myeloid lineage specific metabolism of hydroquinone or catechol to their respective reactive benzoquinone forms via myeloperoxidase catalysis. These reactive metabolites could act via either indirect genotoxicity from oxidative stress, inhibition of topoisomerase II, or cycles of selective immune-mediated depletion followed by homeostatic proliferation of that myeloid lineage. The likelihood of AML and/or MDS appears strongest amongst all cancer types in relation to benzene (North et al., 2020). In the context of potential for inclusion of leukaemia unrelated to benzene to bias potency estimates for benzene-caused leukaemia (discussed above), and the limited number of studies quantitatively considering MDS, it may appear prudent for analysts to consider focusing on benzene AML risk.

In the MOA for benzene leukaemogenicity there are earlier events than cancer: bone marrow dysfunction, manifested as haematological changes, and genetic damage observed in blood cells. In any MOA there are key events that are requirements for disease development, understood to occur in a series. Earlier events occur at lower exposures, generally with a less severe consequence (i.e., subclinical changes in white blood cell counts are a less severe effect than AML). In preventing the earlier, less severe event in the MOA as the key health effect, it is reasonably anticipated that the later, more severe effect of cancer can be prevented. Thus, as we propose an OEL based on consideration of both haematologic and genetic effects there is no further need to identify a point of departure based on carcinogenic effects. It is prudent to consider if the derived limit is objectively in the range without significant increase in risk. In the case of benzene, a limit of 0.25 ppm to 0.5 ppm is in the range where the literature consistently shows a significant excess AML risk is not observed (e.g., Pliofilm, NCI-CAPM, Norwegian offshore workers, Italian shoemakers).

Methods

Functions for the different models were written in R for Statistical Computing 3.5.1 (R Core Team, 2018). Plots were generated in ggplot2 (Wickham, 2016). Model code is available as supplementary information.

LNT Model

As a basis for establishing an upper anchor point for the LNT risk model a 1% risk (ED01) was sought. Estimates for ED01 were based on the method described in the German AGS report which estimates risk for total leukaemia after exposure to benzene based on epidemiological data (reported in Table 11 [p.38] of AGS, 2012). Benzene AML risk has been assessed in multiple populations globally, with several of the populations reassessed on more than one occasion. Thus, we sought to establish a comparable ED01 based on AML based on the latest updated studies rather than total leukaemia. Five of the six studies used by AGS have been updated since their report. The Pliofilm study by Rhomberg et al. (2016); the National Cancer Institute China study by Linet et al. (2019); the Glass et al. (2005) study was updated in a pooled analysis by Rushton et al. (2014); the Dow study by Collins et al. (2015); and the largest component population in the Wong (1987a) study by Collins et al. (2003). The Seniori Costantini et al. (2003) study was not used due to the lack of an analysis of AML. The Stenehjem et al. (2015) study was added due to the presence of an analysis on AML. Thus the updated analysis is also based on six studies. An average cumulative exposure (ppm-years) associated with a 1% increased risk of AML mortality was calculated in the same manner in which AGS calculated 10% increased risk from six studies listed in **Error! Reference source not found.** An assumption of 20 year latency (Richardson, 2008) was used to divide the average cumulative exposure to calculate an average yearly exposure (ppm), which provided an upper anchor point at the ED01 for the risk model.

SMS Model

In the SMS model (Equation 1), risk was assumed to decrease at concentrations less than the LOAEC, but with fixed adjustments to the slope for the discrete intervals between the LOAEC, NOAEC, and OEL. The LOAEC (2 ppm, 6.38 mg/m³) and NOAEC (0.5 ppm, 1.6 mg/m³) for hematotoxicity and genotoxicity were selected from analysis of multiple high quality studies (Schnatter et al., 2020). Since there is no compelling data that shows either hematotoxic or genotoxic effects below the LOAEC, concentrations below 2 ppm were thought plausible as a region below which risk for leukaemia would be decreased. The modification factor applied corresponded to potential assessment factors commonly used in risk assessment that could be applied to the LOAEC to identify an OEL. Selection of the discrete intervals, and the associated assessment factor used for adjustment in that interval, should be identified. In the case of benzene the following change points and assessment factors were considered based on the rationale outline in Schnatter et al. (2020):

Discrete Interval	Assessment Factor
LOAEC to NOAEC	4 (composite of 2*2)
NOAEC to OEL	2
LOAEC to OEL	8 (composite of 4*2)

Equation 1. Segment Modified Slope Model

$$\begin{aligned}
 & \text{Risk}(c_{Benzene}) \\
 = & \left\{ \begin{array}{l}
 0.25 * \frac{0.01 - \text{Risk}\left(0.8 \frac{mg}{m^3}\right)}{60 \frac{mg}{m^3} - 0.32 \frac{mg}{m^3}} \left((c_{Benzene}) - 0.8 \frac{mg}{m^3} \right), \quad (c_{Benzene}) > 0.8 \frac{mg}{m^3} < 1.6 \frac{mg}{m^3} \\
 \text{Risk}\left(1.6 \frac{mg}{m^3}\right) + 0.5 * \frac{0.01 - \text{Risk}\left(1.6 \frac{mg}{m^3}\right)}{60 \frac{mg}{m^3} - 1.6 \frac{mg}{m^3}} \left((c_{Benzene}) - 1.6 \frac{mg}{m^3} \right), \quad (c_{Benzene}) \geq 1.6 \frac{mg}{m^3} < 6.38 \frac{mg}{m^3} \\
 \text{Risk}\left(6.38 \frac{mg}{m^3}\right) + \frac{0.01 - \text{Risk}\left(6.38 \frac{mg}{m^3}\right)}{60 \frac{mg}{m^3} - 6.38 \frac{mg}{m^3}} \left((c_{Benzene}) - 6.38 \frac{mg}{m^3} \right), \quad (c_{Benzene}) \geq 6.38 \frac{mg}{m^3} \\
 0, \quad \text{else}
 \end{array} \right.
 \end{aligned}$$

CDS Model

The CDS model (Equation 2) utilized the same response function as LNT above its lower anchor point, the LOAEC.

Risk estimates in the CDS model were assumed to decrease at concentrations less than the 2 ppm LOAEC for hematotoxicity and/or genotoxicity. In the absence of justification for a more complex shape for a slope reduction method, a simple linear function was applied. In the modification factor there is no reduction in slope at LOAEC (i.e., it is a value of one at the LOAEC), but as concentrations become closer to the OEL the modification factor progressively approaches zero until it reaches the OEL, where it becomes equal to zero.

$$\text{Modification Factor function} = mx + b$$

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{1 - 0}{LOAEC - OEL}$$

$$b = y - mx = 0 - (m \times OEL)$$

In this concept, the slope for the modification factor function is calculated using simple algebra: the lower point of the modification factor function occurs at the OEL where risk assumed to be zero (i.e., where y is the modification factor, $y_1 = 0$ and $x_1 = OEL$), with the upper point found where risk extrapolated from observed is unmodified (i.e., $y_2 = 1$ and $x_2 = LOAEC$). After the slope of the modification factor is calculated, the intercept is calculated using either of the OEL or LOAEC coordinates as input, though use of the OEL value is mathematically simpler because it makes $y = 0$. Once a modification factor function is defined for the region of interest (the region below the LOAEC for hematotoxicity and/or genetic toxicity), the linear risk function becomes modified.

$$\text{Risk} = (\text{Slope for Modification Factor} \times \text{Concentration} + \text{intercept}) \times \text{Linear Model Slope} \times \text{Concentration}$$

In this implementation, there are effectively three segments: the range below the OEL (0.25 ppm, 0.8 mg/m³) where risk is assumed to be zero, the range between the OEL and LOAEC in which the slope

goes from zero to an unmodified value, and the range above the LOAEC (2 ppm, 6.38 mg/m³) in which the slope is derived from the LNT model (Equation 2).

Equation 2. *Continuously Decreasing Slope Model*

$$\begin{aligned}
 & \text{Risk}(c_{\text{Benzene}}) \\
 = & \left\{ \begin{array}{l}
 (\text{Slope for Modification Factor} * c_{\text{Benzene}} + y \text{ intercept}) \times \frac{0.01}{\text{ED01}} \times \left(c_{\text{Benzene}} - 0.8 \frac{\text{mg}}{\text{m}^3} \right), \quad \left(c_{\text{Benzene}} \right) > 0.8 \frac{\text{mg}}{\text{m}^3} < 6.38 \frac{\text{mg}}{\text{m}^3} \\
 \frac{0.01 - \text{Risk} \left(6.38 \frac{\text{mg}}{\text{m}^3} \right)}{60 \frac{\text{mg}}{\text{m}^3} - 6.38 \frac{\text{mg}}{\text{m}^3}} \left(c_{\text{Benzene}} \right) - 6.38 \frac{\text{mg}}{\text{m}^3}, \quad \left(c_{\text{Benzene}} \right) \geq 6.38 \frac{\text{mg}}{\text{m}^3} \\
 0, \quad \text{else}
 \end{array} \right.
 \end{aligned}$$

Calculation of a Point of Departure (ED01 Value)

The starting point for determination of the point of departure was the AGS (2012) report, in which an ED10 was calculated for total leukaemia. The AGS analysis, in turn, was based on initial calculations from Roller et al. (2006) who considered benzene as one example for POD calculations derived from both human and animal studies for several substances. Roller et al. (2006) used the ED10 for human studies to be able to directly compare results with the animal data-derived ED10. However, Roller pointed out that use of the ED10 in human studies can require “upward extrapolation” when the background risk is 1% or less, which is true for both total leukaemia and AML. We chose to use the ED01 as the point of departure owing to the low background risk of AML and the fact that an excess risk of 1% is usually readily observable as a positive finding outside of chance variation in most cancer epidemiologic studies of suitable size (United States Environmental Protection Agency, 2005).

In Table 11 of their report, AGS used the following studies in their ED10 calculation: 1) the Pliofilm study (four different sets of authors), 2) an Italian shoe factory by Seniori Costantini et al. (2003), 3) chemical workers in China by Hayes et al. (1997), 4-5) chemical workers in the US by Bloemen et al. (2004) and Wong (1987a), 6) utility workers in France by Guénel et al. (2002), and 7) oil industry workers in Australia by Glass et al. (2003) and Glass et al. (2005). We selected these studies, while searching for updates of each study and also considering subsequent study quality guidelines (Vlaanderen et al., 2008) as referenced in AGS (2012). Since Guénel et al. (2002) was not based on underlying quantitative benzene data, it was excluded. The following studies were updated and had quantitative data on benzene exposure and AML risk: 1) Pliofilm by Rhomberg et al. (2016), 2) chemical workers in China by Linet et al. (2019), 3-4) chemical workers in the US by Collins et al. (2003) and Collins et al. (2015), and oil industry workers by Rushton et al. (2014). In addition, another study in Norwegian offshore workers (Stenehjem et al., 2015) was used, although only part of the cohort was based on quantitative benzene data. The Seniori Costantini et al. (2003) study could not be used due to an absence of a dose-response analysis for AML. Thus, six studies were used for a revised ED01 calculation.

We followed the approach outlined in the Appendix of the AGS report to calculate the ED01 for each study. Briefly, a simple linear regression was fit to categorical dose response data (cumulative exposure in ppm-years) versus AML risk for each study. The resulting regression line was used to solve for a risk that represented an absolute risk of 1% over either a background incidence of AML (3.08 for the incidence

studies, i.e., Linet et al., 2019; Rushton et al., 2014; Stenehjem et al., 2015) or a background mortality rate of AML (3.38 for the mortality studies, i.e., Collins et al., 2003; Collins et al., 2015; Rhomberg et al., 2016). Background leukaemia lifetime incidence and mortality risks were obtained from SEER data (Howlander et al., 2015), and fractions for AML were obtained from on line statistics of the American Cancer Society (2020). Table 1 shows the categorical cumulative exposure groupings used for each study; when more than one ED01 estimate was made for the same study, the two estimates were averaged. The quintile grouping favored by Rhomberg et al. (2016) and also the Rinsky et al. (2002) categorization previously used by AGS, but updated in the Rhomberg et al. (2016), study were used for the Pliofilm study. For the Linet et al. (2019) study, both lag periods (2-10 and >10) were summed to make the ED01 comparable to the other studies, in which all lag periods were used. For three of the six studies, two ED01 values were calculated: one using the zero exposure group and another using only positive exposure values, recognizing that the control group may not be ideal. An overall average of 378 ppm-years was calculated from the six study populations. Utilizing a 20 year cancer latency period (Richardson, 2008), this calculates to 18.9 ppm, or 60 mg/m³. On that basis, the risk models were derived using 60 mg/m³ as the ED01 upper anchor point.

Table 1. ED01 Values (ppm-years) for Six Benzene Study Populations

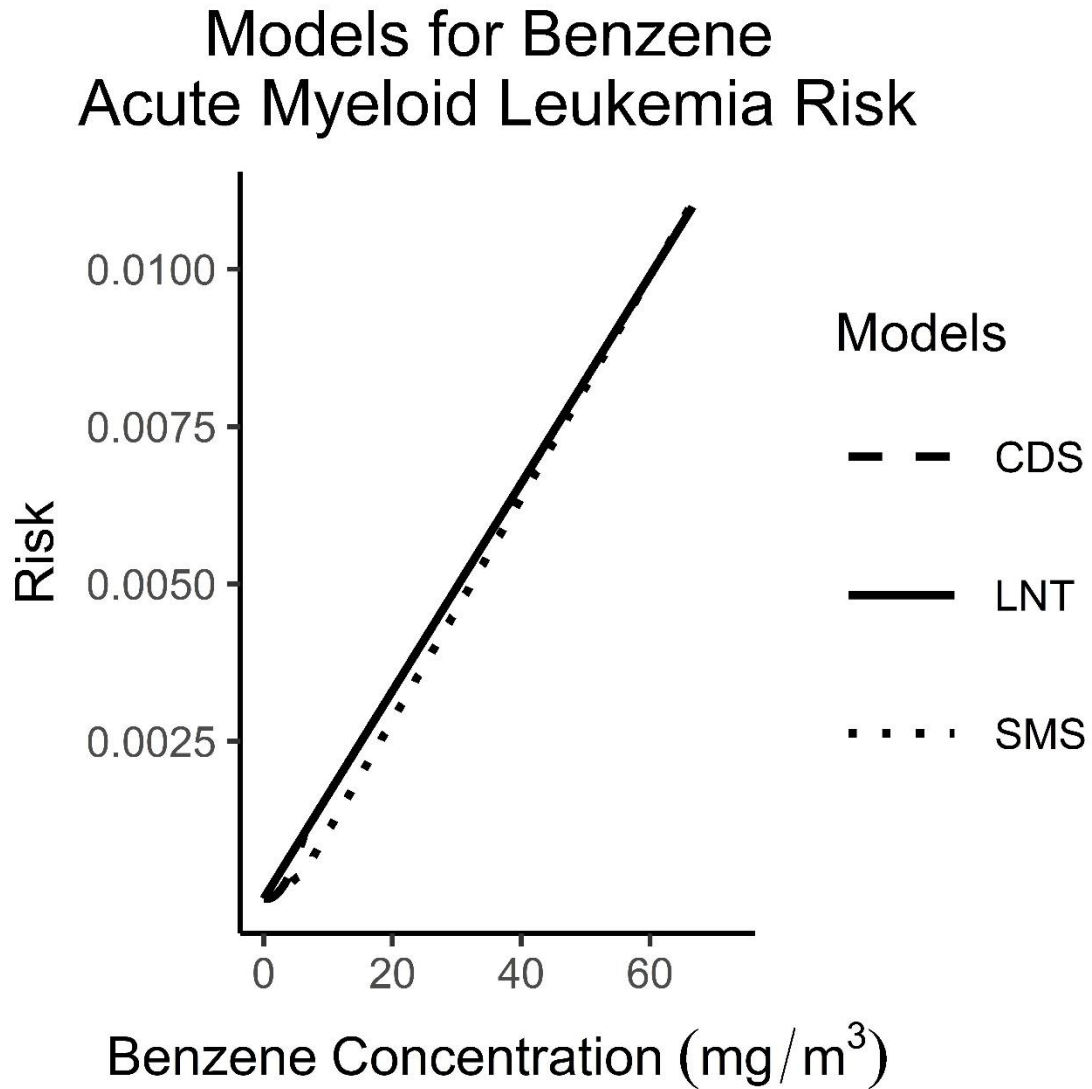
Study (ref)	Description First Value	Description Second Value	First Value	Second Value	Final Value
Pliofilm (Rhomberg et al., 2016)	Quintiles	Rinsky Categories	98.2	151	125
NCI/CAPM (Linet, 2019) ¹	Exclude zero exposure	Include zero exposure	1693	1625	1660
Pooled Petroleum (Rushton, 2014)	Tertiles				68.9
Dow (Collins 2015)	1-3.99, 4-24.99, 25+				347
Solutia (Collins 2003)	<1, 1-6, >6, exclude zero	<1, 1-6, >6, include zero			68.9
Norwegian Offshore (Stenehjem 2015)	Tertiles, exclude zero	Tertiles, include zero	0.334	0.332	0.333
TOTAL					378

¹ Based on aggregate 2-10 year and > 10 year lagged analysis for AML and MDS

Risk Model Outputs

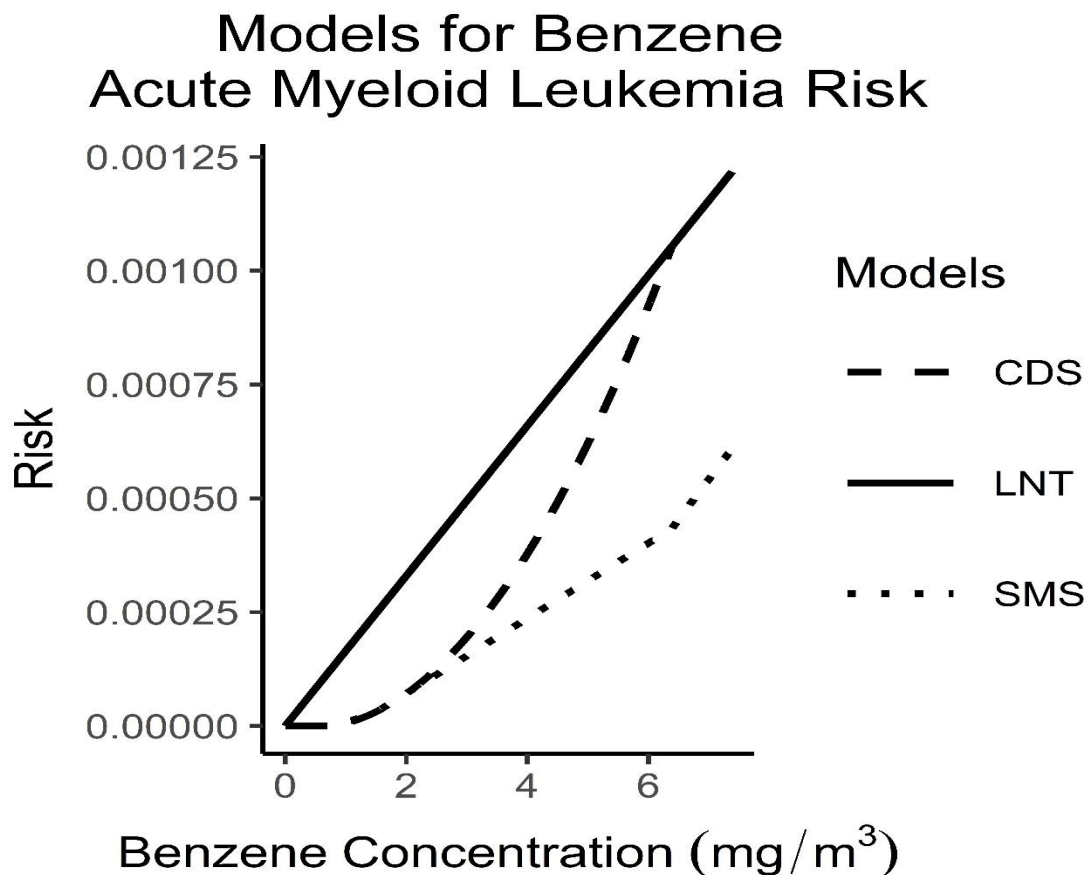
All three models are plotted over the range from zero to the ED01 value in Figure 1. Above the LOAEC, the LNT and CDS models are the same. At the LOAEC the SMS model begins at a lower risk point than LNT or CDS, but above the LOAEC has a steeper slope in order to converge with the LNT and CDS models at the ED01 value.

Figure 1. Risk Models from OEL to ED01



In Figure 2 all three models are shown in the region between zero and the LOAEC, making the difference in the models below the LOAEC more visible. As noted above, in the CDS model the slope of the line gradually decreases from being equal to the LNT model until it reaches zero at the OEL.

Figure 2. Risk Models in region near LOAEC and OEL



For comparison, the risk values from the three models are shown at specific points of interest in the model (Table 2) and in reference to specific risk values used in some risk assessments (Table 3).

Table 2. Risk estimates at specific change points in models (rounded to single digit)

AML Risk	LNT	CDS	SMS
LOAEC	0.001	0.001	0.0004
NOAEC	0.0002	0.00003	0.00003
OEL	0.0001	0	0

Table 3. Concentrations (mg/m³) at specific risk metrics by models

AML Risk	LNT	CDS	SMS
4 in 10,000	2.42	4.099	5.982
1 in 10,000	0.61	2.281	2.345
4 in 100,000	0.24	1.6286	1.6174
1 in 100,000	0.061	1.1039	1.04
1 in 1,000,000	0.006	0.8378	0.821

Discussion

Risk assessments for benzene require model selection, and historically relied on LNT models. While understandable, data accumulating since the 2000s have increasingly built the case for a threshold risk of AML from occupational benzene exposure (Committee for Risk Assessment, 2018; Dutch Expert Committee on Occupational Safety, 2014; North et al., 2020). Risk modelers may wish to incorporate a threshold in models used for benefit cost assessment, but the observed data for mortality may leave uncertainty in the lower exposure range. One option is to shift the lower anchor point from the origin to the OEL (or other exposure limit), but this approach discards potentially useful information about key events in the causal chain to an adverse outcome. In scenarios where such key events can be identified, it would be reasonable to modify the estimated response based on that information that introduces flexibility into the risk model.

The introduction of flexible response curves is not novel, but knot selection in the response curve has not typically been informed by key events. Such an approach was illustrated by the German Committee for Hazardous Substances (Ausschuss für Gefahrstoffe) in their risk modelling of benzene exposure and leukaemia mortality (AGS, 2012), in which a kink function was described using a benchmark dose lower limit value derived by the United States Agency for Toxic Substances and Disease Registry based on hematotoxicity (2007). The CDS and SMS models are informed by their approach, but expand that application to consider LOAELs, NOAELs, and OELs as points of change.

One distinction from the AGS approach and that taken here is in the utilization of an ED01 instead of ED10 as an anchor point. Estimation of an ED10 value requires extrapolation to levels of risk well above that observed all the studies except the Pliofilm cohort. Utilization of an ED01 value reduces the amount of extrapolation needed for risk estimates, placing the benchmark within the observed range for increased risk instead above it, thus potentially reducing the chance of misestimating the point of departure.

A simple approach (i.e., linearly decreasing function) for modifying the response slope was applied in the illustration of the CDS model above. Assessors could certainly apply other functions to increase or decrease the rate of slope reduction. Such selection could likely be informed by an understanding of the dose-response connecting the LOAEL, NOAEL, and OEL. A steep dose-response for key events might justify a function for the modification factor that more slowly reduces slope, whereas a shallow dose-response for the key event might justify a function that more rapidly reduces slope.

As expected, there is a difference in risk estimates for exposures <2 ppm between the LNT model compared with the CDS and SMS models at risk metrics used in some cancer risk assessments. The reduced risk estimates in the CDS and SMS models reflect an expectation by preventing early key events in the MOA one reduces the risk of later adverse outcomes. Some may find the rationale of AGS, exposures equal or greater than that causing hematotoxicity lead to an enhancement of the disease process, compelling. That enhancement eventually leads to observed increases in AML risk for benzene-exposed workers. By that same rationale, exposures below the point of enhancement are expected to be lower risk. Above the LOAEC for earlier key events, in regions where epidemiology observes cancer risk, the risk model relies on cancer risk. Below the LOAEC for earlier key events, where cancer risk is not or cannot be observed, the risk model takes the assumption that a “damping” of risk occurs until it is at background levels. The CDS and SMS models offer a flexible approach to defining how quickly risk estimates change below the point of enhancement.

Application of an SMS model requires selection of the slope modification factor over each segment. Without a clear rationale the most likely selection method would seem to be selecting the smallest modification factor near the LOAEL, permitting the modification factor to become larger as the response

nears the NOAEL and OEL. As with the CDS model, information about the dose response for the key event in the range between the NOAEL and LOAEL may provide a useful in justification of a slope modification factor. Information in this context could come from a range of sources, including animal studies, in vitro experiments, or even in silico modelling. Undoubtedly the further the information source gets from the observed human data the more challenging it may be to incorporate.

In the future, use of defined benchmark responses as change points instead of NOAECs and LOAECs may be an improvement. Benchmark dose modelling has significant technical advantages over a NOAEC/LOAEC approach, but can be challenging to apply to some datasets. The AGS analysis used such an approach which could have been expanded to include more than a single benchmark response. For example, the modeller could use non-adverse, small adverse, and moderate adverse response as additional change points in a risk model. In this regard, the general theory of effect size may be a useful guide to defining a small, medium, or large effect size for benchmark dose modelling (Slob, 2017).

Looking to the future, as modellers increasingly look to integrate in vitro and in silico results to predict whole organism responses, the need for informed approaches to dose-response modelling may move beyond simply empirical curve fitting. While the models described above are informed by key events observed in vivo, a future model could consider combining dose-response information from in vitro genetic damage or hematotoxicity responses in ex vivo bone marrow cultures with physiologically-based pharmacokinetic models to synthesize a prediction for the LOAEL, NOAEL, or BMD of a response. In turn, these derived key event points could inform change point selection or slope modification factors applied to high dose in vivo response models.

In closing, we illustrate the use of key events to modify dose-response models used in risk assessment with the aim of refining risk estimates in the lower range of the dose-response model. Such models may prove useful in future risk assessment.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The LOA REACH Consortia provided financial support for Penman Consulting staff and consultant participation, and the employers of the other authors provided salary and travel support in the normal course of their work.

References

- Advisory Committee on Safety and Health at Work, 2019. Opinion on an EU Binding Occupational Exposure Limit Value (BOEL) for Benzene under the Carcinogens and Mutagens Directive 2004/37/EC.
- Agency for Toxic Substances and Disease Registry, 2007. Toxicological Profile for Benzene.
- AGS, 2012. Begründung zu Benzol in BekGS.
- American Cancer Society, 2020. Key Statistics for Acute Myeloid Leukaemia (AML). vol. 2020.
- Bloemen, L.J., Youk, A., Bradley, T.D., Bodner, K.M., Marsh, G., 2004. Lymphohaematopoietic cancer risk among chemical workers exposed to benzene. *Occupational and Environmental Medicine* 61, 270-274.
- Calabrese, E., Golden, R., 2019. Assessing the Scientific Basis of the Linear No Threshold (LNT) Model with Threshold Models for Cancer Risk Assessment of Radiation and Chemicals.

- Cody, R.P., Strawderman, W.W., Kipen, H.M., 1993. Hematologic Effects of Benzene: Job-Specific Trends during the First Year of Employment among a Cohort of Benzene-Exposed Rubber Workers. *Journal of Occupational and Environmental Medicine* 35, 776-782.
- Collins, J., Ireland, B., Buckley, C., Shepperly, D., 2003. Lymphohaematopoietic cancer mortality among workers with benzene exposure. *Occupational and environmental medicine* 60, 676-679.
- Collins, J.J., Anteau, S.E., Swaen, G.M., Bodner, K.M., Bodnar, C.M., 2015. Lymphatic and hematopoietic cancers among benzene-exposed workers. *Journal of occupational and environmental medicine* 57, 159-163.
- Committee for Risk Assessment, 2018. Opinion on scientific evaluation of occupational exposure limits for Benzene. European Chemicals Agency.
- Copley, G.B., Schnatter, A.R., Armstrong, T.W., Irons, R.D., Chen, M., Wang, X.Q., Kerzic, P., 2017. Hospital-Based Case-Control Study of MDS Subtypes and Benzene Exposure in Shanghai. *J Occup Environ Med*.
- Crump, K.S., 1994. Risk of benzene-induced leukaemia: A sensitivity analysis of the pliofilm cohort with additional follow-up and new exposure estimates. *Journal of Toxicology and Environmental Health, Part A Current Issues* 42, 219-242.
- Dutch Expert Committee on Occupational Safety, 2014. Benzene - Health-based recommended occupational exposure limit. Health Council of the Netherlands,.
- Glass, D.C., Gray, C.N., Jolley, D.J., Gibbons, C., Sim, M.R., 2005. Health Watch exposure estimates: Do they underestimate benzene exposure? *Chemico-biological interactions* 153-154, 23-32.
- Glass, D.C., Gray, C.N., Jolley, D.J., Gibbons, C., Sim, M.R., Fritschi, L., Adams, G.G., Bisby, J.A., Manuell, R., 2003. Leukaemia Risk Associated With Low-Level Benzene Exposure. *Epidemiology* 14, 569-577.
- Guénel, P., Imbernon, E., Chevalier, A., Crinquand-Calastreng, A., Goldberg, M., 2002. Leukaemia in relation to occupational exposures to benzene and other agents: A case-control study nested in a cohort of gas and electric utility workers. *American journal of industrial medicine* 42, 87-97.
- Hayes, R.B., Dosemeci, M., Wacholder, S., Travis, L.B., Rothman, N., Hoover, R.N., Linet, M.S., Yin, S.-N., Li, G.-L., Li, C.-Y., 1997. Benzene and the dose-related incidence of hematologic neoplasms in China. *Journal of the National Cancer Institute* 89, 1065-1071.
- Howlader, N., Noone, A., Krapcho, M., Garshell, J., Miller, D., Altekruse, S., Kosary, C., Yu, M., Ruhl, J., Tatalovich, Z., Mariotto, A., Lewis, D., Chen, H., Feuer, E., Cronin, K., 2015. SEER Cancer Statistics Review, 1975-2012, National Cancer Institute. Bethesda, MD.
- International Agency for Research on Cancer, 2018. Benzene. International Agency for Research on Cancer, pp. 301.
- Kipen, H.M., Cody, R.P., Crump, K.S., Allen, B.C., Goldstein, B.D., 1988. Hematologic effects of benzene: a thirty-five year longitudinal study of rubber workers. *Toxicol Ind Health* 4, 411-430.
- Li, W., Schnatter, A.R., 2018. Benzene risk assessment: does new evidence on myelodysplastic syndrome justify a new approach? *Critical reviews in toxicology* 48, 417-432.
- Linet, M.S., Gilbert, E.S., Vermeulen, R., Dores, G.M., Yin, S.-N., Portengen, L., Hayes, R.B., Ji, B.-T., Lan, Q., Li, G.-L., Rothman, N., Control, C.C.f.D., Group, P.U.N.C.I.B.S., 2019. Benzene Exposure Response and Risk of Myeloid Neoplasms in Chinese Workers: A Multicenter Case-Cohort Study. *JNCI: Journal of the National Cancer Institute* 111, 465-474.
- North, C.M., Rooseboom, M., Kocabas, N.A., Schnatter, A.R., Faulhammer, F., Williams, S.D., 2020. Modes of Action Considerations in Threshold Expectations for Health Effects of Benzene. Submitted to *Toxicology Letters*.
- R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rhomberg, L., Goodman, J., Tao, G., Zu, K., Chandalia, J., Williams, P.R.D., Allen, B., 2016. Evaluation of Acute Nonlymphocytic Leukaemia and Its Subtypes With Updated Benzene Exposure and Mortality Estimates: A Lifetable Analysis of the PlioFilm Cohort. *Journal of Occupational and Environmental Medicine* 58, 414-420.

- Richardson, D.B., 2008. Temporal Variation in the Association between Benzene and Leukaemia Mortality. *Environmental Health Perspectives* 116, 370-374.
- Rinsky, R.A., Hornung, R.W., Silver, S.R., Tseng, C.Y., 2002. Benzene exposure and hematopoietic mortality: A long-term epidemiologic risk assessment. *American Journal of Industrial Medicine* 42, 474-480.
- Roller, M., Akkan, Z., Hassauer, M., Kalberlah, F., 2006. Risikoextrapolation vom Versuchstier auf den Menschen bei Kanzerogenen. *Wirtschaftsverl. NW., Verlag für Neue Wiss.*
- Rothman, N., Haas, R., Hayes, R.B., Li, G.L., Wiemels, J., Campleman, S., Quintana, P.J., Xi, L.J., Dosemeci, M., Titenko-Holland, N., 1995. Benzene induces gene-duplicating but not gene-inactivating mutations at the glycophorin A locus in exposed humans. *Proceedings of the National Academy of Sciences* 92, 4069-4073.
- Rushton, L., Schnatter, A.R., Tang, G., Glass, D.C., 2014. Acute myeloid and chronic lymphoid leukaemias and exposure to low-level benzene among petroleum workers. *British Journal of Cancer* 110, 783-787.
- Schnatter, A.R., Glass, D.C., Tang, G., Irons, R.D., Rushton, L., 2012. Myelodysplastic syndrome and benzene exposure among petroleum workers: an international pooled analysis. *Journal of the National Cancer Institute* 104, 1724-1737.
- Schnatter, A.R., Nicolich, M.J., Bird, M.G., 1996. Determination of leukemogenic benzene exposure concentrations: Refined analyses of the Pliofilm cohort. *Risk analysis* 16, 833-840.
- Schnatter, A.R., Rooseboom, M., Kocabas, N.A., North, C.M., Dalzell, A., Twisk, J.J., Faulhammer, F., Rushton, E., Boogard, P.J., Ostapenkaite, V., Williams, S.D., 2020. Derivation of an Occupational Exposure Limit for Benzene Using Epidemiological Study Quality Assessment Tools *Toxicology Letters*.
- Schnatter, A.R., Rosamilia, K., Wojcik, N.C., 2005. Review of the literature on benzene exposure and leukaemia subtypes. *Chemico-biological interactions* 153-154, 9-21.
- Seniori Costantini, A., Quinn, M., Consonni, D., Zappa, M., 2003. Exposure to benzene and risk of leukaemia among shoe factory workers. *Scandinavian journal of work, environment & health*, 51-59.
- Slob, W., 2017. A general theory of effect size, and its consequences for defining the benchmark response (BMR) for continuous endpoints. *Critical reviews in toxicology* 47, 342-351.
- Stenhjem, J.S., Kjarheim, K., Bratveit, M., Samuelsen, S.O., Barone-Adesi, F., Rothman, N., Lan, Q., Grimsrud, T.K., 2015. Benzene exposure and risk of lymphohaematopoietic cancers in 25,000 offshore oil industry workers. *Br J Cancer* 113, 1641-1641.
- United States Environmental Protection Agency, 2005. *Guidelines for Carcinogen Risk Assessment*.
- Vlaanderen, J., Vermeulen, R., Heederik, D., Kromhout, H., European Union Network of Excellence, E.I.R.A.G., 2008. *Guidelines to Evaluate Human Observational Studies for Quantitative Risk Assessment*. *Environmental Health Perspectives* 116, 1700-1705.
- Wickham, H., 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.
- Wong, O., 1987a. An industry wide mortality study of chemical workers occupationally exposed to benzene. I. General results. *British Journal of Industrial Medicine* 44, 365-381.
- Wong, O., 1987b. An industry wide mortality study of chemical workers occupationally exposed to benzene. II. Dose response analyses. *British Journal of Industrial Medicine* 44, 382-395.
- Wong, O., Harris, F., Armstrong, T.W., Hua, F., 2010. A hospital-based case-control study of acute myeloid leukaemia in Shanghai: analysis of environmental and occupational risk factors by subtypes of the WHO classification. *Chemico-biological interactions* 184, 112-128.