Research

A Quantitative Meta-Analysis of the Relation between Occupational Benzene Exposure and Biomarkers of Cytogenetic Damage

Bernice Scholten,[1,2] Jelle Vlaanderen,[1] Rob Stierum,[2] Lützen Portengen,[1] Nat Rothman,[3] Qing Lan,[3] Anjoeka Pronk,[2] and Roel Vermeulen[1]

[1]Institute for Risk Assessment Sciences, Utrecht University, Utrecht, Netherlands
[2]Netherlands Organisation for Applied Scientific Research, Zeist, Netherlands
[3]Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Rockville, Maryland, USA

BACKGROUND: The genotoxicity of benzene has been investigated in dozens of biomonitoring studies, mainly by studying (classical) chromosomal aberrations (CAs) or micronuclei (MN) as markers of DNA damage. Both have been shown to be predictive of future cancer risk in cohort studies and could, therefore, potentially be used for risk assessment of genotoxicity-mediated cancers.

OBJECTIVES: We sought to estimate an exposure–response curve (ERC) and quantify between-study heterogeneity using all available quantitative evidence on the cytogenetic effects of benzene exposure on CAs and MN respectively.

METHODS: We carried out a systematic literature review and summarized all available data of sufficient quality using meta-analyses. We assessed the heterogeneity in slope estimates between studies and conducted additional sensitivity analyses to assess how various study characteristics impacted the estimated ERC.

RESULTS: Sixteen CA (1,356 individuals) and 13 MN studies (2,097 individuals) were found to be eligible for inclusion in a meta-analysis. Studies where benzene was the primary genotoxic exposure and that had adequate assessment of both exposure and outcomes were used for the primary analysis. Estimated slope estimates were an increase of 0.27% CA [(95% CI: 0.08%, 0.47%); based on the results from 4 studies] and 0.27% MN [(95% CI: −0.23%, 0.76%); based on the results from 7 studies] per parts-per-million benzene exposure. We observed considerable between-study heterogeneity for both end points ($I^2 > 90\%$).

DISCUSSION: Our study provides a systematic, transparent, and quantitative summary of the literature describing the strong association between benzene exposure and accepted markers of genotoxicity in humans. The derived consensus slope can be used as a best estimate of the quantitative relationship between real-life benzene exposure and genetic damage in future risk assessment. We also quantitate the large between-study heterogeneity that exists in this literature, a factor which is crucial for the interpretation of single-study or consensus slopes. https://doi.org/10.1289/EHP6404

Introduction

Benzene is a well-known environmental contaminant that was classified as a human carcinogen (IARC Group 1) in 1974 (IARC 1974), which was recently reconfirmed (IARC 2018). Benzene and its metabolites may cause leukemia via genotoxic effects on the pluripotent hematopoietic stem cells, resulting in chromosomal changes (IARC 2009). A substantial number of cross-sectional studies have demonstrated that benzene exposure is associated with the occurrence of (classical) chromosomal aberrations (CA) and micronuclei (MN) that are considered early markers of genotoxicity [see reviews by Zhang et al. (2002) and McHale et al. (2012)]. A CA is often defined as the appearance of missing, extra, or irregular portions of chromosomal DNA (NHGRI 2016), and an MN is considered to be the small nucleus that forms whenever a chromosome fragment or whole chromosome is not incorporated into the daughter nuclei after cell division (Fenech 2002). Information on intermediate end points is increasingly used in the risk assessment for many chemicals; for example, the International Agency for Research on Cancer (IARC) integrated mechanistic evidence with evidence from other data streams to support conclusions regarding carcinogenicity (Smith et al. 2016), and recently the Dutch Health Council Committee recommended a health-based occupational exposure limit for benzene that was (primarily) based on hematological effects in humans (Health Council of the Netherlands 2014).

To date, however, there have not been many attempts to assess in a systematic and quantitative way how exposure affects some of these intermediate end points, such as the induction of genetic damage. For example, Angelini et al. (2016) recently reviewed the published data on benzene exposure and MN, but the authors restricted their findings to a nonquantitative comparison between exposed and unexposed individuals.

We hypothesize that it would be useful to investigate the exposure–response curve (ERC) of benzene-induced CA or MN. An advantage of using biomarkers of effect for a (more) quantitative risk assessment is that they may offer more precise information on the exact shape of the exposure–response relation at low exposures, especially when the clinical outcomes of interest are relatively rare at low exposure levels. For benzene this is important because there is evidence that the exposure–response relation with cancers might be nonlinear (Kim et al. 2006; Rothman et al. 1998; Vlaanderen et al. 2010). CAs and MN are additionally relevant to study because both have been found to be predictive of an increased risk of cancer in large prospective cohort studies (Bonassi et al. 2008, 2007).

We aimed to summarize all currently available human observational data and use this information to quantitatively describe the relation between occupational exposure to benzene and CA and MN frequencies. This would be a first step in the use of preclinical cancer-predictive end points in risk assessment. We also assessed the between-study heterogeneity effect estimates and conducted sensitivity analyses to assess how various study characteristics affect the estimated ERC. This information is important when evidence from this literature (either from a single study or from a meta-analysis) is interpreted in a risk assessment context.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehponline@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

Methodology

Study Identification and Selection

A literature search was conducted from until August 2018 in PubMed and Scopus, using the following search terms for "benzene" AND chromosomal aberrations: ("chromosomal aberrations" OR "chromosome*"), and the following search terms for "benzene" AND micronuclei: ("micronuclei" OR "micronucleus"). References in all identified publications were checked for additional relevant studies. Publications were selected for inclusion in the meta-analysis if they had been peer reviewed and when they fulfilled the following criteria: *a*) the study population had been occupationally exposed to benzene; *b*) exposure assessment was based on quantitative benzene exposure estimates in air or relevant benzene biomarker measurements; and *c*) results from cytogenetic tests were available and shown per exposure category. Studies that were excluded were those that measured general air pollution (e.g., occupational exposure to traffic air), studies with considerable co-exposure to other carcinogenic substances (IARC classification 1 or 2a), studies that reported benzene exposure at levels that could not reasonably be distinguished from general background levels (~ 5 ppb; IARC 2018). Case–control studies were also excluded from further analysis. If more than one publication was published on the same cohort, the most recent update was included.

We decided to restrict our analyses to classical CA assessments and excluded fluorescent *in situ* hybridization (FISH) CA studies. FISH technology is based on fluorescent DNA probes to allow for the detection aberrations in specific chromosomes (e.g., an alteration in the number of specific chromosomes or loss of particular chromosomal regions). FISH has been reported to be more sensitive in detecting the genotoxic effects of benzene, compared with classical CA analysis (Smith et al. 1998). Eleven publications (of which about half were from a single study) have employed FISH to report changes in specific chromosomes in relation to quantitatively assessed benzene exposure. However, the large heterogeneity in the various FISH protocols used (e.g., in terms of detecting specific chromosomes or measuring specific changes ranging from alterations in the number of specific chromosomes), acquisition of specific translocations or loss of particular regions of certain chromosomes, and assessments in different stages of the cell cycle (interphase vs. metaphase) prohibited the inclusion of these studies into this meta-analysis.

Study Evaluation

Each study to be included in the meta-analysis was evaluated for overall study quality and for the (expected) quality of the exposure and cytogenetic assessments. We selected a number of criteria for both end points.

Exposure assessment. A tiered approach was used to evaluate exposure assessment quality (see Table S1). First, we assessed whether benzene exposure was the main exposure or if other co-exposures may have existed. Studies designed primarily to evaluate the effects of benzene exposure, that is, in which benzene exposure was the main exposure were classified as "A." These were mostly studies in the shoemaking or petroleum industries. Studies that were classified as "B" were studies in jobs where significant co-exposure to other genotoxic (air) pollutants was not reported, but could not be ruled out, for example, in filling station attendants. Studies that were classified A were further divided based on the quality of their exposure assessment, with A^+ studies assessed exposure with the specific aim of exposure assessment for the respective study with at least some (personal or stationary) benzene exposure measurements or data on benzene metabolites that had been measured in blood or urine. All other A studies were classified as A^-.

Cytogenic assessment. The ranking of CA and MN assessments focused on the quality of the cytogenetic analyses. Studies that were classified A met the following requirements: For CA, the number of metaphases counted should be >100 and the culture time <50 h (Carrano and Natarajan 1988). For MN, the number of bi-nucleated cells scored should be >1,000, and cells had to be incubated initially for 44 h, followed by the addition of cytochalasin B, and, subsequently, another 24–48 h of incubation (Fenech 2007). Classification as an A-class study further required that both the frequency and variability [standard error (SE) or standard deviation (SD)] of CA or MN were reported. All studies not meeting these criteria were classified as "B" (see Table S1).

Primary study set. Studies that scored an A^+ for exposure assessment and an A for cytogenic assessment were included in the primary analyses. The other studies were included in sensitivity analyses (i.e., the full study set). The quality of eligible studies was evaluated jointly by four reviewers for exposure assessment (R.V., J.V., L.P., and B.S.) and three reviewers for cytogenetic analysis (J.V., R.S., and B.S.).

Extraction and Preparation of Data from Selected Papers

The following information was extracted from each relevant publication that was used: the country where the study was performed; the method used for exposure assessment and cytogenetic analyses, including culture time and the types of aberrations included; and the availability of biomonitoring data. For each exposure group within a study we also extracted the number of subjects, benzene exposure level, the recorded CA and/or MN frequencies with either the SE, SD, or range of frequencies. We also extracted information on benzene exposure levels and CA or MN frequencies separately by smoking status (yes/no) when this information was available.

Several studies lacked quantitative estimates of benzene exposure but did provide information on blood or urine levels of benzene or its metabolites. For those studies, we estimated benzene exposure levels through exposure reconstruction by using the physiologically based kinetic model by Knutsen et al. (2013). When the publication provided no benzene exposure levels for controls, we assumed a background level of 5 ppb.

The frequency of CA and MN were expressed as the number of cells with aberrations per 100 cells (a percentage). CA or MN that were reported in other units were calculated back to 100 cells. For studies that did not report the SE for CA or MN frequencies, we calculated the SE based on the reported SD and sample size. For the studies that did not report either the SE or SD (and, hence, were not included in the primary study set), we estimated the SE under the assumption that the SD equaled the median SD for the primary study set ($n = 5$ for CA).

Statistical Analysis

The main goal of this paper was to summarize the available information on the exposure–response relation between benzene and markers of cell damage. Because most of the individual studies reported aberration frequencies for only two groups (one of which is either unexposed or low-exposed), there was very little opportunity to investigate nonlinear exposure–response relations using this data, unless one was willing to make strong assumptions. We therefore decided to perform a meta-analysis on summary estimates of the available exposure–response information within each study (i.e., the study-specific slope), rather than attempt to build a (hierarchical) meta-regression model (for the group-level data).

The information available consisted, for most studies, of two average exposure estimates (denoted as $X0$ and $X1$) and two

estimates of the average aberration frequencies (denoted as $Y0$ and $Y1$) together with an estimate of the precision of these latter estimates (denoted as variance $V0$ and $V1$). Although the aberration frequency as a proportion could theoretically be modeled as a binomial outcome for an individual, this was not the case for the reported average aberration frequencies and would have required quasi-likelihood methods to account for overdispersion even if the individual-level data had been available. We therefore regarded the average aberration frequencies $Y0$ and $Y1$ as random variables with known variances. Standard algebra provides the rules to calculate the expected value and variance of the random variable that results from subtraction of $Y0$ from $Y1$ and dividing it by the difference between $X0$ and $X1$ (i.e., the slope factor), without reference to any specific (e.g., normal) distribution. The expected value of this random variable equals $(Y1 - Y0)/(X1 - X0)$, with variance $(V0 + V1)/(X1 - X0)2$.

For example, for Bogadi (1997) the reported average proportion plus or minus the standard error (SE) of chromosomal aberrations was $1 \pm 0.12\%$ in the unexposed and $1.5 \pm 0.11\%$ in the exposed (with an average exposure of $5.9\,\mathrm{ppm}$). The estimated slope (β) is then $(1.5 - 1)/5.9 = 0.08$ (units %/ppm) with an estimated variance of $(0.122 + 0.112)/(5.92) = 0.0007$. This slope factor and associated variance was then used in the meta-analysis.

Meta-analyses were performed using the metafor package (Viechtbauer 2019) in R (version 3.6.1; R Development Core Team). Meta-analysis using this approach failed for the full MN study set because the range of sampling variances was extremely large, making the results of the algorithm used numerically unstable. We therefore chose to analyze the MN primary study set only. We calculated the I^2 statistic to assess the heterogeneity in slope estimates between studies. Funnel plots, the trim-and-fill method, and the Egger test were used to investigate potential publication bias.

Several sensitivity analyses were performed to explore the potential bias/variance tradeoff of including all studies (the full study set) in the analysis and to evaluate the impact of individual studies on the estimated meta-slope, or consensus ERC. In the first sensitivity analysis, we allowed the full set of studies to contribute to the estimation of the consensus ERC. The second (set of) sensitivity analyses was performed by leaving out one study at a time from the sample set (i.e., jackknifing) for both the primary and full study set.

To evaluate potential effect modification by smoking, we restricted the analyses to data from studies that reported exposure and genotoxic effects by smoking status, and we evaluated the interaction between smoking and benzene exposure on the frequency of CA and MN by including smoking as a moderator variable.

Results

Chromosomal Aberrations

Of the 745 identified studies on CA and benzene exposure, 16 (1,356 individuals) were considered eligible for inclusion in the meta-analysis. An overview of all eligible studies and a flowchart indicating how studies were selected is provided in Figure S1. Four studies (477 individuals) were included in the primary study set. Of the 16 selected studies, the vast majority ($n = 15$), across a range of occupations, showed higher CA frequencies in jobs involving benzene exposure when compared with unexposed jobs (see Table S2). Only the benzene-exposed fuel tanker drivers in Lovreglio et al. (2014) had lower CA frequencies in comparison with their control group. The individual ERCs for all 16 studies (the full study set) are presented in Figure 1, which clearly shows the large variation across studies in baseline CA frequencies and the derived benzene ERCs.

A meta-analysis of the slopes on the primary study set confirmed that benzene exposure was positively and significantly associated with an increase in CA [β = 0.27% (95% CI: 0.08%, 0.47%)] but with an I^2 of 94%, indicating very strong heterogeneity between studies (Figure 2). Expanding the analysis to include the full study set resulted in a similar positive but heterogenous association [β = 0.29% (95% CI: 0.16%, 0.42%), $I^2 = 91.1\%$].
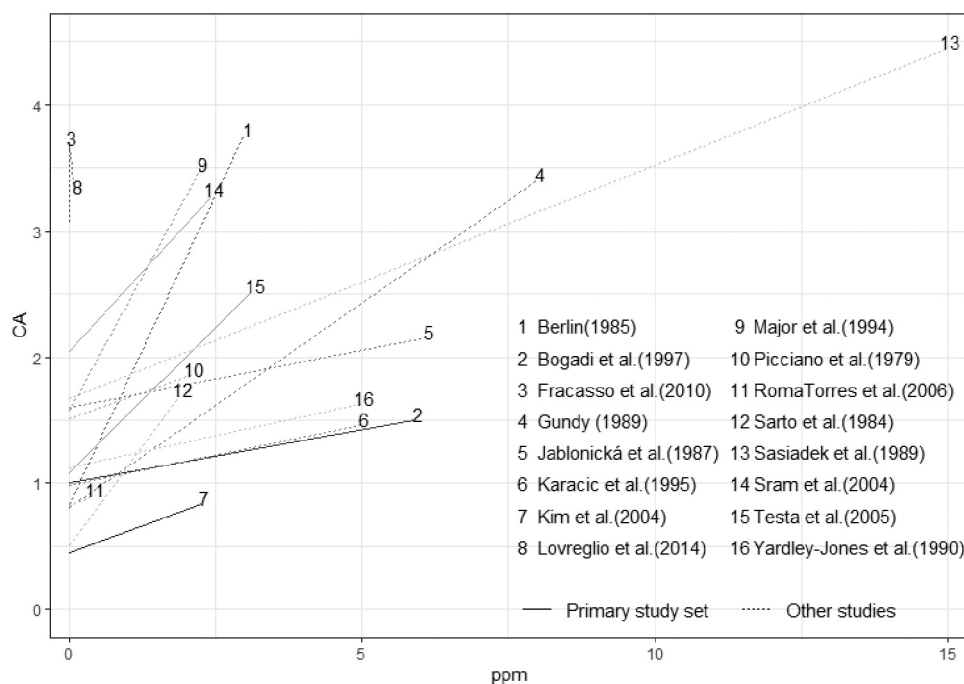
Figure 1. Chromosome aberrations (number of aberrations per 100 cells) and benzene exposure (ppm). The primary study set contained studies where benzene was the primary genotoxic exposure and that had adequate assessment of both exposure and outcomes. The term other studies refers to studies not meeting these criteria. The numbers shown with each line indicate the study name. Note: CA, chromosomal aberration.
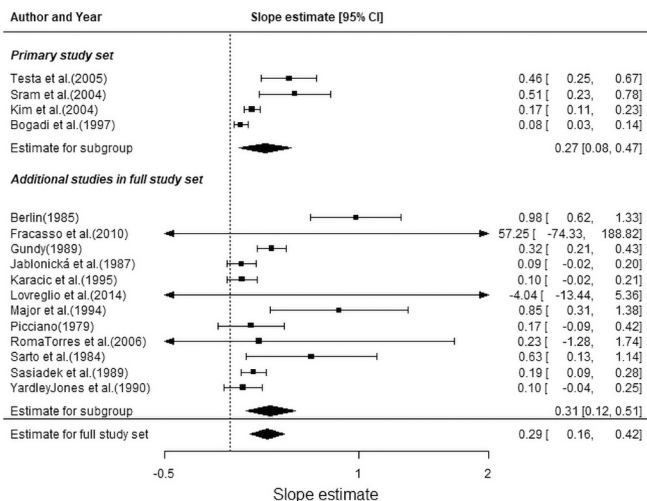
Figure 2. Meta-analysis of slope estimates (in units%/ppm) of benzene-induced chromosomal aberrations. The primary study set contained studies where benzene was the primary genotoxic exposure and that had adequate assessment of both exposure and outcomes. The term other studies refers to studies not meeting these criteria. The I^2 of the primary study set was 94%; the I^2 of the full study set was 91.1%. Note: CI, confidence interval.

To assess the sensitivity of the outcome of the meta-analysis, one study at the time was excluded from both the primary and the full study set using a jackknife approach. The results are presented in Table S3. These analyses showed that excluding a single study had a considerable impact on the point estimate of the meta-slope (range: 0.22–0.35) when limiting analysis to the primary study set. When considering all studies, Berlin et al. (1985) was found to be most influential on the meta-slope estimate [i.e., the average slope was 0.29, whereas excluding Berlin et al. (1985) would result in a slope estimate of 0.22, $I^2 = 78\%$].

Of the selected CA studies from the full study set, four (Jablonická et al. 1987; Major et al. 1994; Fracasso et al. 2010; Sram et al. 2004), contained specific information on CA aberrations for smokers and nonsmokers in the control and benzene-exposed group (see Figure S2). There was no statistical interaction effect of smoking ($p = 0.6$).

There was evidence of publication bias for the CA studies ($p = 0.02$). We attempted to adjust the results for this bias using the trim-and-fill method (see Figure S3).
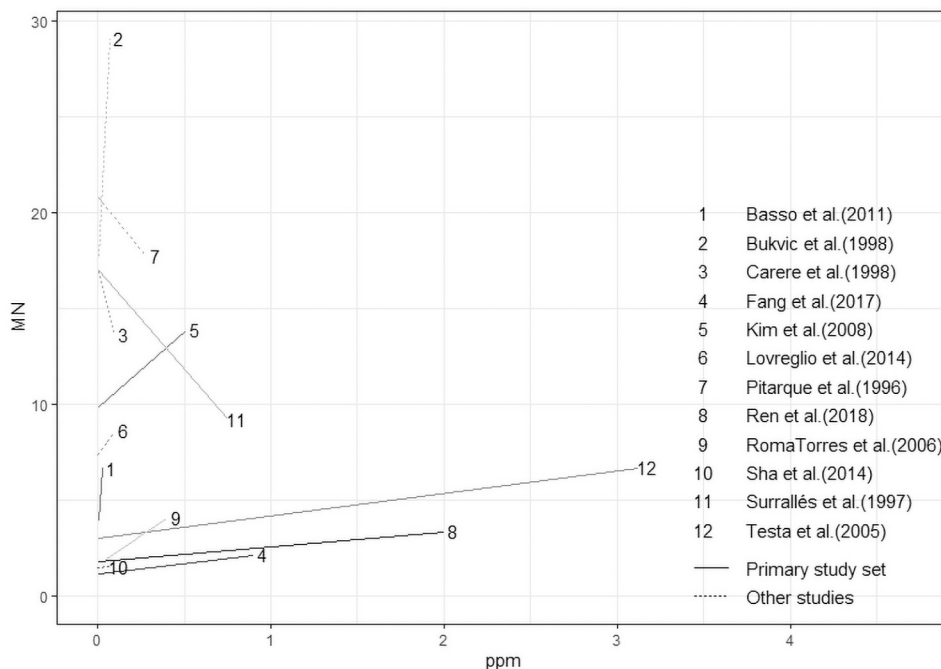
Micronuclei

Of the 315 MN studies that were found, 13 (2,097 individuals) were considered eligible for inclusion in the meta-analysis (see Table S4). Nine of these studies (1,672 individuals) were included in the primary study set (see Figure S4). Figure 3 indicates that most studies showed higher MN frequencies in the exposed group ($n = 9$), but the slopes varied considerably.

The meta-slope of the primary study set of MN studies is 0.27% (95% CI: −0.23%, 0.76%) with an I^2 of 99.5% (Figure 4). We were unable to derive a consensus slope for the full data set owing to the large range in sampling variances for some of the slope estimates.

When applying jackknifing on the primary study set, the range of slopes varied between 0.17 [when Basso et al. (2011) was excluded] to 1.03 [when Ren et al. (2018) was excluded] (see Table S5). Three MN studies from the full study set (Bukvic et al. 1998; Basso et al. 2011; Sha et al. 2014) contained information on smoking status and corresponding MN frequencies in both the benzene exposed and unexposed group. There was no statistical interaction effect of smoking ($p = 0.5$) (see Figure S5).

There was evidence of publication bias for the MN studies ($p = 0.001$). The trim-and-fill method pooled estimate was 0.24 with a corresponding SE of 0.98, whereas the meta-slope of the primary study set was 0.27 ± 0.19 (see Figure S6).

Figure 3. Micronuclei and benzene exposure. The primary study set contained studies where benzene was the primary genotoxic exposure and that had adequate assessment of both exposure and outcomes. The term other studies refers to studies not meeting these criteria. The numbers shown with each line indicate the study name. The study by Liu et al. (1996) is not displayed in this figure because it was the only study that measured very high exposure levels (up to 60 ppm). Note: MN, micronuclei.
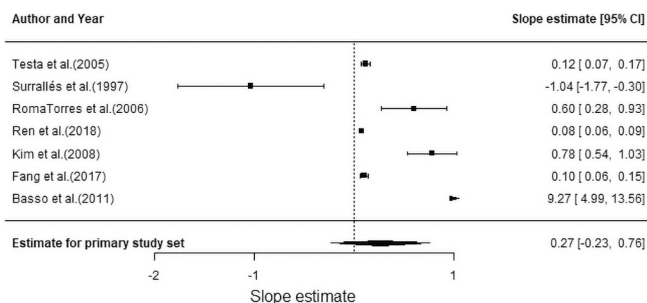
Figure 4. Meta-analysis of slope estimates (in units %/ppm) of benzene-induced micronuclei, based on the primary study set. The primary study set contained studies where benzene was the primary genotoxic exposure and that had adequate assessment of both exposure and outcomes. The I^2 of the study set was 99.5%. Note: CI, confidence interval; MN, micronuclei.

Discussion

A substantial number of occupational studies has been undertaken in the past decades to study the effect of benzene on genotoxicity, covering a range of occupations and entailing various co-exposures and different levels of benzene exposure. Although in the vast majority, CA and MN frequencies were associated with jobs involving benzene exposure (IARC 2018), a consensus quantitative relationship between exposure levels and CA or MN frequency is not yet defined.

To quantitatively summarize data on occupational benzene exposure and markers of genetic damage, we conducted a systematic review, followed by a meta-analysis and a set of sensitivity analyses. We found a positive slope for both CA [$\beta = 0.27\%$ (95% CI: 0.08%, 0.47%)] and MN [$\beta = 0.27\%$ (95% CI: −0.23%, 0.76%)] in relation to benzene exposure for the high quality studies (primary study set), but the between-study heterogeneity in slope estimates was large, with an estimated I^2 of over 90% for both end points.

Including also studies with less stringent quality criteria (the full study set) the slope was comparable for CA [$\beta = 0.29\%$ for all studies (95% CI: 0.16%, 0.42%), $I^2 = 91.1\%$] in comparison with the primary study set. We were unable to derive a meta-slope for the full data set for MN owing to the large range in variances for the study-specific slopes.

There are several possible explanations for the observed differences in individual slopes, including the small number of individuals per study (on average about 40), the limited dose groups (in most cases there was only one exposure estimate for the exposed group), differences in exposure assessment methods, and uncertainty in both exposure levels and CA or MN counts. This latter issue is reflected by the results of the HUman MicroNucleus project, which was established in early 2000 to gather MN data from 25 labs representative of many countries and populations (Bonassi et al. 2001). An overall median MN frequency in nonexposed individuals of 6.5 per thousand was found with an interquartile range of between 3 and 12 per thousand. The authors concluded that most of the observed total variance could be explained by laboratory methods. The issue of uncertainty in measuring DNA damage rates might be even more critical for classical CAs given that no harmonized protocol has been developed for this biomarker and that CAs are known to be more difficult to score in comparison with MN (Fenech 2002). Indeed, we found that analyzing and comparing all CA studies included in this meta-analysis was difficult because the interpretation of (total) chromosomal aberrations varied among studies. A range of structural aberrations were included by some studies, whereas others reported only a specific type of aberration, thereby making comparisons more difficult. Still, all authors aimed to collect the same information (i.e., total classical CAs), whereas this was not

the case for the FISH studies, where different end points in different chromosomes were selected in different division stages of the cells (e.g., interphase or metaphase). Hence, we decided to exclude studies based on FISH technology because they could not be sensibly combined. However, potentially, the observed analytical variance in CA measurements could be less for FISH-based studies because scoring is easier and can be automated. In addition, a chromosome-wide aneuploidy study described heterogeneity in monosomy and trisomy rates for specific chromosomes (Zhang et al. 2011). Focusing on the more sensitive chromosomes may then provide a more accurate estimate of benzene's genotoxic effect that maybe be partly masked in CA and MN assays where all chromosomes are considered.

Exposure assessments applied in the different studies were not equivalent. For example, some studies described the measurement of personal exposure in the breathing zone of workers, and others estimated exposure based on stationary samples. Although we did take uncertainty in measured CA or MN frequencies into account in this meta-analysis, this was not possible for the exposure estimates because information on the exposure assessment error was often missing (as commonly seen in meta-analyses). This could potentially add to the observed heterogeneity between studies. Note that it has been reported that I^2 has a substantial bias when the number of studies is small because an alternative confidence intervals could replace the I^2 estimates (von Hippel 2015). We provide both estimates in this paper.

The studies in the meta-analysis were weighed on the corresponding inverse variance. We did so because we assumed the SD described the variability of counted CAs or MN between individuals; this was confirmed by a number of authors of the studies used for meta-analysis in this paper (R. Sram and P. Lovreglio, personal communication). Due to the limited number of studies, the limited number of estimates per study, and the high heterogeneity between studies, we refrained from explorations on nonlinearity of the ERC. However, some evidence is available that the deleterious effects of benzene might, in fact, be nonlinear (Kim et al. 2006; Rothman et al. 1998).

We also investigated the potential interaction effect between smoking and benzene exposure. The results were not significant for either CA or MN; however, the number of studies available with information on CA or MN frequencies specifically for smokers and nonsmokers was limited. There was some evidence of publication bias, most notably for the MN studies where the trim-and-fill method indicated a (relatively small) change in pooled estimate but a very large increase in the corresponding SE (i.e., the slope estimated decreased from 0.27 to 0.24), and the corresponding SE became much larger (from 0.19 to 0.98). This effect was mainly driven by the study of Basso et al. (2011), where the observed outcome was much larger for the reported SE in comparison with other studies (see Figure S6). Publication bias thus influenced the MN slope estimate, adding to the uncertainty of the meta-slope. Evidence for publication bias for MN was also found by Angelini et al. (2016). Finally, in the past few years, several studies explored possibilities beyond measuring classical CA or MN frequencies, that is, studying other biomarkers [e.g., telomere length assays (Bassig et al. 2014)], or other cell types [e.g., cells closer to acute myeloid leukemia/myelodysplastic syndrome-related stem cells (Zhang et al. 2012)]. Potentially, these studies could also be used to derive ERCs, but we chose to focus on CA and MN assays because the evidence base for these studies is larger and their association to benzene exposure is widely accepted.

We quantitively estimated the relationship between benzene exposure and induced genetic damage to inform future risk assessments; for example, because both CAs and MN have been found to be predictive of an increase in future cancer risk (although not for

specific hematopoietic malignancies, given the lack of power to investigate these associations), our derived ERC might be used to inform the shape of the benzene–leukemia curve in exposure ranges where epidemiological data is limited. One strategy to include our ERC in a benzene–leukemia risk assessment would involve a Bayesian approach that uses information on the meta-slope of CA and MN as prior in estimating the benzene–leukemia ERC. We believe the inclusion of biomarkers could potentially enhance risk assessment because these studies can be informative for early biological effects (which might not be included in cohort studies on, e.g., cancer occurrence) or relevant biological pathways, and in addition they can inform the ERC at lower exposure levels because biomonitoring studies have generally more power than outcome studies at lower exposure levels.

In conclusion, we summarized all available data on CA and MN frequencies in benzene-exposed workers and provide a quantitative estimate of the exposure–response relationship. As far as the authors are aware, no other study applied the quantitative approach used in this paper for other chemicals. We believe that further insight into the quantitative relationship between real-life benzene exposure and genetic damage could provide a starting point for further inclusion of this data in benzene risk assessments; however, the considerable observed heterogeneity between studies should be carefully considered.

Acknowledgments

References

Angelini S, Bermejo JL, Ravegnini G, Sammarini G, Hrelia P. 2016. Application of the lymphocyte cytokinesis-block micronucleus assay to populations exposed to petroleum and its derivatives: results from a systematic review and meta-analysis. Mutat Res 770(pt A):58–72, PMID: 27894691, https://doi.org/10.1016/j.mrrev.2016.03.001.

Bassig BA, Zhang L, Cawthon RM, Smith MT, Yin S, Li G, et al. 2014. Alterations in leukocyte telomere length in workers occupationally exposed to benzene. Environ Mol Mutagen 55(8):673–678, PMID: 24945723, https://doi.org/10.1002/em.21880.

Basso E, Cevoli C, Papacchini M, Tranfo G, Mansi A, Testa A. 2011. Cytogenetic biomonitoring on a group of petroleum refinery workers. Environ Mol Mutagen 52(6):440–447, PMID: 21370282, https://doi.org/10.1002/em.20641.

Berlin M. 1985. Low level benzene exposure in Sweden: effect on blood elements and body burden of benzene. Am J Ind Med 7(5–6):365–373, PMID: 4003399, https://doi.org/10.1002/ajim.4700070503.

Bogadi-Šare A, Vlatka B, Rajka T, Višnja K, Marija Z. 1997. Genotoxic effects in workers exposed to benzene: with special reference to exposure biomarkers and confounding factors. Ind Health 35(3):367–373, PMID: 9248220, https://doi.org/10.2486/indhealth.35.367.

Bonassi S, Fenech M, Lando C, Lin YP, Ceppi M, Chang WP, et al. 2001. HUman MicroNucleus project: international database comparison for results with the cytokinesis-block micronucleus assay in human lymphocytes: I. Effect of laboratory protocol, scoring criteria, and host factors on the frequency of micronuclei. Environ Mol Mutagen 37(1):31–45, PMID: 11170240, https://doi.org/10.1002/1098-2280(2001)37:1<31::AID-EM1004>3.0.CO;2-P.

Bonassi S, Norppa H, Ceppi M, Strömberg U, Vermeulen R, Znaor A, et al. 2008. Chromosomal aberration frequency in lymphocytes predicts the risk of cancer: results from a pooled cohort study of 22 358 subjects in 11 countries. Carcinogenesis 29(6):1178–1183, PMID: 18356148, https://doi.org/10.1093/carcin/bgn075.

Bonassi S, Znaor A, Ceppi M, Lando C, Chang WP, Holland N, et al. 2007. An increased micronucleus frequency in peripheral blood lymphocytes predicts the risk of cancer in humans. Carcinogenesis 28(3):625–631, PMID: 16973674, https://doi.org/10.1093/carcin/bgl177.

Bukvic N, Bavaro P, Elia G, Cassano F, Fanelli M, Guanti G. 1998. Sister chromatid exchange (SCE) and micronucleus (MN) frequencies in lymphocytes of gasoline station attendants. Mutat Res 415(1–2):25–33, PMID: 9711259, https://doi.org/10.1016/S1383-5718(98)00045-X.

Carere A, Antoccia A, Crebelli R, Di Chiara D, Fuselli S, Iavarone I, et al. 1995. Exposure to benzene and genotoxic effects among filling station attendants. Epidemiologia e Prevenzione 19(62):105–119, PMID: 7601233, https://europepmc.org/abstract/med/7601233.

Carrano AV, Natarajan AT. 1988. International Commission for Protection against Environmental Mutagens and Carcinogens. ICPEMC publication no. 14. Considerations for population monitoring using cytogenetic techniques. Mutat Res 204(3):379–406, PMID: 3347212, https://doi.org/10.1016/0165-1218(88)90058-5.

Fang Y, Wu H-T, Ye Y-J, Zhou LF, Hu W, Zhang G-H, et al. 2017. Association between polymorphisms of metabolic enzyme genes and chromosomal damage in benzene-exposed workers in China. J Occup Environ Med 59(11):e215–e220, PMID: 29116991, https://doi.org/10.1097/JOM.0000000000001148.

Fenech M. 2002. Biomarkers of genetic damage for cancer epidemiology. Toxicology 181–182:411–416, PMID: 12505344, https://doi.org/10.1016/S0300-483X(02)00480-8.

Fenech M. 2007. Cytokinesis-block micronucleus cytome assay. Nat Protoc 2(5):1084–1104, PMID: 17546000, https://doi.org/10.1038/nprot.2007.77.

Fracasso ME, Doria D, Bartolucci GB, Carrieri M, Lovreglio P, Ballini A, et al. 2010. Low air levels of benzene: correlation between biomarkers of exposure and genotoxic effects. Toxicol Lett 192(1):22–28, PMID: 19427373, https://doi.org/10.1016/j.toxlet.2009.04.028.

Gundy S. 1989. Cytogenetical studies on a large control population and on persons occupationally exposed to radiation and/or to chemicals. Ann Lst Super Sanita 25(4):549–556, PMID: 2631621.

Health Council of the Netherlands. 2014. *Benzene: Health-Based Recommended Occupational Exposure Limit* [in Dutch]. https://www.healthcouncil.nl/binaries/healthcouncil/documents/advisory-reports/2014/02/21/benzene-health-based-recommended-occupational-exposure-limit/advisory-report-benzene-health-based-recommended-occupational-exposure-limit.pdf [accessed 21 July 2020].

IARC (International Agency for Research on Cancer). 1974. *IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Man. Some Anti-Thyroid and Related Substances, Nitrofurans and Industrial Chemicals*, volume 7. https://publications.iarc.fr/25 [accessed 21 July 2020].

IARC. 2009. *Chemical Agents and Related Occupations. Benzene. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, vol. 100F. Lyon, France: IARC. https://publications.iarc.fr/_publications/media/download/3076/73443059d4ec0adde733204bab30939c7470dd2b.pdf [accessed 21 July 2020].

IARC. 2018. *Benzene. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, vol. 120. Lyon, France: IARC. https://publications.iarc.fr/_publications/media/download/5630/48902d7035dcd9f6f64ac780ac64cabaf033871b.pdf [accessed 21 July 2020].

Jablonická A, Vargová M, Karelová J. 1987. Cytogenetic analysis of peripheral blood lymphocytes in workers exposed to benzene. J Hyg Epidemiol Microbiol Immunol 31(2):127–132, PMID: 3611757.

Karačić V, Ljiljana S, Branka B-C, Ana B-S. 1995. Possible genotoxicity in low level benzene exposure. Am J Ind Med 27(3):379–388, PMID: 7747744, https://doi.org/10.1002/ajim.4700270307.

Kim S, Vermeulen R, Waidyanatha S, Johnson BA, Lan Q, Rothman N, et al. 2006. Using urinary biomarkers to elucidate dose-related patterns of human benzene metabolism. Carcinogenesis 27(4):772–781, PMID: 16339183, https://doi.org/10.1093/carcin/bgi297.

Kim YJ, Yoon HC, Domyung P, Hai WC. 2004. Determination of chromosome aberrations in workers in a petroleum refining factory. J Toxicol Environ Health A. 67(23–24):1915–1922, PMID: 15513892, https://doi.org/10.1080/15287390490513319.

Kim YJ, Jun YC, Domyung P, Hai WC. 2008. Association of the NQO1, MPO, and XRCC1 polymorphisms and chromosome damage among workers at a petroleum refinery. J Toxicol Environ Health A 71(5):333–341, PMID: 18214807, https://doi.org/10.1080/15287390701738558.

Knutsen JS, Kerger BD, Finley B, Paustenbach DJ. 2013. A calibrated human PBPK model for benzene inhalation with urinary bladder and bone marrow compartments. Risk Anal 33(7):1237–1251, PMID: 23278103, https://doi.org/10.1111/j.1539-6924.2012.01927.x.

Liu L, Zhang Q, Feng J, Deng L, Zeng N, Yang A, et al. 1996. The study of DNA oxidative damage in benzene-exposed workers. Muta Res Genet Toxicol 370(3–4):145–150, PMID: 8917660, https://doi.org/10.1016/S0165-1218(96)00048-1.

Lovreglio P, Maffei F, Carrieri M, D'Errico MN, Drago I, Hrelia P, et al. 2014. Evaluation of chromosome aberration and micronucleus frequencies in blood lymphocytes of workers exposed to low concentrations of benzene. Mutat Res Genet Toxicol Environ Mutagen 770:55–60, PMID: 25344164, https://doi.org/10.1016/j.mrgentox.2014.04.022.

Major J, Jakab M, Kiss G, Tompa A. 1994. Chromosome aberration, sister-chromatid exchange, proliferative rate index, and serum thiocyanate concentration in smokers exposed to low-dose benzene. Environ Mol Mutagen 23(2):137–142, PMID: 8143702, https://doi.org/10.1002/em.2850230211.

McHale CM, Zhang L, Smith MT. 2012. Current understanding of the mechanism of benzene-induced leukemia in humans: implications for risk assessment.

Carcinogenesis 33(2):240–252, PMID: 22166497, https://doi.org/10.1093/carcin/bgr297.

NHGRI (National Human Genome Research Institute). 2016. Chromosome Abnormalities Fact Sheet. https://www.genome.gov/about-genomics/fact-sheets/Chromosome-Abnormalities-Fact-Sheet [accessed 21 July 2020].

Picciano D. 1979. Cytogenetic study of workers exposed to benzene. Environ Res 19(1):33–38, PMID: 510268, https://doi.org/10.1016/0013-9351(79)90031-8.

Pitarque M, Carbonell E, Lapeña N, Marsá M, Torres M, Creus A, et al 1996. No increase in micronuclei frequency in cultured blood lymphocytes from a group of filling station attendants. Mutat Res Genet Toxicol 367(3):161–167, https://doi.org/10.1016/0165-1218(95)00091-7.

Ren J-C, Wu Y-X, Wu Z, Zhang G-H, Wang H, Liu H, et al. 2018. MTHFR gene polymorphism is associated with DNA hypomethylation and genetic damage among benzene-exposed workers in southeast China. J Occup Environ Med 60(4):e188–e192, PMID: 29370017, https://doi.org/10.1097/JOM.0000000000001288.

Roma-Torres J, João PT, Susana S, Blanca L, Luís MC, Josefina M, et al. 2006. Evaluation of genotoxicity in a group of workers from a petroleum refinery aromatics plant. Mutat Res 604(1–2):19–27, PMID: 16431152, https://doi.org/10.1016/j.mrgentox.2005.12.005.

Rothman N, Bechtold WE, Yin SN, Dosemeci M, Li GL, Wang YZ, et al. 1998. Urinary excretion of phenol, catechol, hydroquinone, and muconic acid by workers occupationally exposed to benzene. Occup Environ Med 55(10):705–711, PMID: 9930093, https://doi.org/10.1136/oem.55.10.705.

Sarto F, Cominato I, Pinton AM, Brovedani PG, Merier E, Peruzzi M, et al. 1984. A cytogenetic study on workers exposed to low concentrations of benzene. Carcinogenesis 5(6): 827–832, PMID: 6722989, https://doi.org/10.1093/carcin/5.6.827.

Sasiadek M, Jagielski J, Smolik R. 1989. Localization of breakpoints in the karyotype of workers professionally exposed to benzene. Mutat Res 224(2):235–240, PMID: 2797038, https://doi.org/10.1016/0165-1218(89)90161-4.

Sha Y, Zhou W, Yang Z, Zhu X, Xiang Y, Li T, et al. 2014. Changes in poly (ADP-ribosyl)ation patterns in workers exposed to BTX. PLoS One 9(9):e106146, PMID: 25215535, https://doi.org/10.1371/journal.pone.0106146.

Smith MT, Guyton KZ, Gibbons CF, Fritz JM, Portier CJ, Rusyn I, et al. 2016. Key characteristics of carcinogens as a basis for organizing data on mechanisms of carcinogenesis. Environ Health Perspect 124(6):713–721, PMID: 26600562, https://doi.org/10.1289/ehp.1509912.

Smith MT, Zhang L, Wang Y, Hayes RB, Li G, Wiemels J, et al. 1998. Increased translocations and aneusomy in chromosomes 8 and 21 among workers exposed to benzene. Cancer Res 58(10):2176–2181, PMID: 9605763.

Sram RJ, Beskid O, Binkova B, Rossner P, Smerhovsky Z. 2004. Cytogenetic analysis using fluorescence in situ hybridization (FISH) to evaluate occupational exposure to carcinogens. Toxicol Lett 149(1–3):335–344, PMID: 15093279, https://doi.org/10.1016/j.toxlet.2003.12.043.

Surrallés J, Autio K, Nylund L, Järventaus H, Norppa H, Veidebaum T, et al. 1997. Molecular cytogenetic analysis of buccal cells and lymphocytes from benzene-exposed workers. Carcinogenesis 18(4):817–823, PMID: 9111220, https://doi.org/10.1093/carcin/18.4.817.

Testa A, Festa F, Ranaldi R, Giachelia M, Tirindelli D, De Marco A, et al. 2005. A multi-biomarker analysis of DNA damage in automobile painters. Environ Mol Mutagen 46(3):182–188, PMID: 16206220, https://doi.org/10.1002/em.20147.

Viechtbauer W. 2019. Meta-Analysis Package for R. https://cran.r-project.org/web/packages/metafor/metafor.pdf [accessed 21 July 2020].

Vlaanderen J, Rothman L, Portengen N, Lan Q, Kromhout H, Vermeulen R. 2010. Flexible meta-regression to assess the shape of the benzene–leukemia exposure–response curve. Environ Health Perspect 118(4):526–532, PMID: 20064779, https://doi.org/10.1289/ehp.0901127.

von Hippel PT. 2015. The heterogeneity statistic I^2 can be biased in small meta-analyses. BMC Med Res Methodol 15:35, PMID: 25880989, https://doi.org/10.1186/s12874-015-0024-z.

Yardley-Jones A, Anderson D, Lovell DP, Jenkinson PC. 1990. Analysis of chromosomal aberrations in workers exposed to low level benzene. Br J Ind Med 47(1):48–51, PMID: 2310707, https://doi.org/10.1136/oem.47.1.48.

Zhang L, Eastmond DA, Smith MT. 2002. The nature of chromosomal aberrations detected in humans exposed to benzene. Crit Rev Toxicol 32(1):1–42, PMID: 11846214, https://doi.org/10.1080/20024091064165.

Zhang L, Lan Q, Guo W, Hubbard AE, Li G, Rappaport SM, et al. 2011. Chromosome-wide aneuploidy study (CWAS) in workers exposed to an established leukemogen, benzene. Carcinogenesis 32(4):605–612, PMID: 21216845, https://doi.org/10.1093/carcin/bgq286.

Zhang L, Lan Q, Ji Z, Li G, Shen M, Vermeulen R, et al. 2012. Leukemia-related chromosomal loss detected in hematopoietic progenitor cells of benzene-exposed workers. Leukemia 26(12):2494–2498, PMID: 22643707, https://doi.org/10.1038/leu.2012.143.