LETTER TO THE EDITOR

Letter to the editor in response to ‘Low-dose metabolism of benzene in humans: science and obfuscation’ Rappaport et al. (2013)

Paul S. Price*, Tim D. Rey, Donald D. Fontaine and Scott M. Arnold

The Dow Chemical Company, Midland, MI 48674, USA

*To whom correspondence should be addressed. Tel: +1 989 636 5030; Fax: +1 989 638 2425; Email: price@dow.com

Introduction

In the paper ‘Low-dose metabolism of benzene in humans: science and obfuscation’ (1), Dr Rappaport and his coauthors vigorously defend the findings of the 2006 publications (2, 3) against the points raised in our 2012 publication ‘A reanalysis of the evidence for increased efficiency in benzene metabolism at airborne exposure levels below 3 p.p.m.’ (4). The work leading to Price et al. was driven by the contradiction of the Kim et al. findings of 9- to 11-fold enhancement of metabolism of benzene at low doses and multiple published reports of lung clearance (5-7) that suggest only a 2-fold change in metabolism with dose. As shown below, this contradiction is confirmed by the information presented in Rappaport et al.

We support the right of Dr Rappaport and his coauthors to question the technical merits of our work or any work in the public domain and to defend Kim et al. (2, 3), but we strongly object to the authors’ wording in the title and in the paper’s concluding paragraph. Although the analyses in the Price et al. paper may have been complex, they were clearly presented and subjected to the peer review process. Furthermore, we published the raw data from the study so that experts could replicate the analyses and come to their own conclusions, which is something Kim et al. (2, 3) did not do. We included factors in our reanalysis that favor a finding of increased metabolism at low doses as well as others that do not. Thus, we clarified and made transparent legitimate technical issues with the analyses of Kim et al. (2, 3) and did not ‘obfuscate’ the original findings. It is unfortunate that Rappaport et al. attempted to disparage our analyses with allegations regarding both our motivations and our professional ethics in expressing our concerns with the Kim et al. data sets and analyses.

After reviewing Rappaport et al., we still contend that (i) methodological errors were made in Kim et al. (2, 3), (ii) the uncertainties in the predictions of the dose-specific metabolites are too large to determine changes in benzene metabolism at air levels below 0.2 p.p.m. and (iii) biases related to the definition of background are the likely cause of the finding of enhanced metabolism at low doses. As the following text demonstrates, Rappaport et al. have not refuted the first two findings and have confirmed the third finding.

Air concentration-related changes in metabolic fraction of benzene

We are largely in agreement with the quantitative findings on lung clearance presented in Rappaport et al. (1), but not the conclusions that the authors drew from the findings (see the Supplementary Information, available at Carcinogenesis Online, to this letter for additional information). The work in Rappaport et al. has greatly clarified the issues and we appreciate (i) the quantitative framework the paper provides and (ii) the collection and presentation of data from a number of published works (7–10), including the unpublished raw data from Egeghy et al. (8).

As shown in the supplemental materials to Rappaport et al., the rate of benzene entering the body by inhalation minus the rate of benzene lost by exhalation is a direct measure of metabolism $Q_{inh}$. Rappaport et al. establishes the relationship between lung clearance and the ‘metabolic fraction’ of benzene ($Q_{met}/Q_{inh}$). Figure 1 plots the predictions of metabolic fraction (linear scale) against the estimate of benzene air concentration as reported in Table S5 of Rappaport et al. (log scale) for four published studies of lung clearance (7–10). (Additional information on all of the figures in this letter is contained in the Supplementary Information, available at Carcinogenesis Online, to this letter.) We have added a simple linear regression model of metabolic fraction versus air concentration. There is a statistically significant trend of metabolic fraction with air concentration, with just under a 2-fold change in metabolic fraction over the range of 0.007–57 p.p.m. The trend, however, for measurements below 2 p.p.m. (Figure 2) is much smaller and was not found to be statistically significant.

At all levels of benzene exposure, the subjects had metabolic fractions of 0.4 or greater and the upper bound of the metabolic fraction is 1. These data suggest that any change in metabolic fraction with decreasing exposures below 60 p.p.m. would be no higher than 2.5 (10.4). This precludes the 9- or 11-fold change in metabolism at low doses proposed by Kim et al. (2, 3) and Rappaport et al.

Fig. 1. Log-linear plot of metabolic fraction ($Q_{met}/Q_{inh}$) and air concentrations from data on lung clearance as reported in four studies of inhalation of benzene (7–10). The data points reflect data on single individuals with the exception of the value at 57 p.p.m. This value reflects the average metabolic fraction for six individuals exposed to levels between 52 and 62 p.p.m. The line presents a linear regression model of metabolic fraction against air concentration. The slope of the line is statistically different from zero.

Fig. 2. Log-linear plot of metabolic fraction ($Q_{met}/Q_{inh}$) and air concentrations from data on lung clearance as reported in studies with air levels below 2 p.p.m. (8, 9). The line presents a simple regression model of metabolic fraction against air concentration for the data in this figure. The slope of the line is not statistically different from zero.

© The Author 2013. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

Carcinogenesis vol.34 no.7 pp.1692–1696, 2013
doi:10.1093/carcin/bgt101
Advance Access publication June 8, 2013
In Rappaport et al. (1) the authors only focus on the changes in Dose Specific Metabolism (DSM) with the estimated air levels of benzene in the four studies. Measurements of DSM are related to metabolic fraction and, where breathing rates and urinary output are constant over a population, changes in DSM are a direct measure of changes in metabolic fraction. Rappaport et al. states correctly that DSM measurements in the four studies decrease by a factor of six with increasing air concentration. But as shown in the supplemental material of Rappaport et al., this change occurred not because of changes in metabolic fraction, but because of differences in the breathing rates for the workers. In Eggeghy et al. (8), an increase in breathing rates relative to the other three studies changed the relationship between air concentration and $Q_{inh}$ by a factor of 3.2. (See text in section 3 of the supplementary materials and footnote b of Table S5 of Rappaport et al.). We agree with these changes in breathing rates and the impact they would have on DSM, but believe that the resulting changes in DSM are irrelevant to the question of dose-related changes in the metabolic fraction of benzene.

Figure 3 provides a more relevant comparison between the inhalation results and the predictions of Kim et al. (3). The figure presents the same studies as Figure 6 of Rappaport et al. but presents the estimates of metabolic fraction for inhalation studies and the Kim et al. (3) results. The values of metabolic fraction for the four inhalation studies were taken from Table S5 of Rappaport et al. The DSM findings from Kim et al. (3) were converted using the same breathing rates used in Rappaport et al. for the Eggeghy et al. (8) data. Figure 3 clearly shows that the predictions of Kim et al. (3) are in conflict with the data from every one of the inhalation studies for air levels that are greater than 0.1 p.p.m. The conflict is greatest at doses above 30 p.p.m., where the Kim et al. (3) predictions of the metabolic fraction are 10% but the inhalation data indicate that they should be greater than 40%.

Based on these findings, we conclude that the new analyses from Rappaport et al. support the observation in Price et al. that the dose-related changes in the metabolism of benzene reported in the Kim et al. (2,3) publications are inconsistent with empirical measurements of the lung clearance of benzene.

Resolving the contradiction between lung clearance findings and the conclusions reached in Kim et al.

This leaves us with the question of how to resolve the findings of Kim et al. (2,3) with the data on lung clearance. In Price et al., we explored whether the disagreement might be due to (i) statistical errors in the natural spline modeling, the calibration modeling and the selection of the background levels of benzene metabolites, (ii) biases that resulted from the use of a calibration model derived from workers whose exposures were dominated by occupational exposures to workers whose exposures were dominated by environmental sources and use of tobacco products and (iii) spurious findings from random variations in the data.

Errors in the Kim et al. methodology

Two errors identified in Kim et al. (2,3) were the use of geometric means of the metabolites in the urine samples of the controls rather than the natural means when determining the background corrections and the absence of a bias correction factor in both the natural spline and calibration models. In Rappaport et al., the authors state:

Price et al. contend that natural-scale mean (hereafter, simply ‘mean’) values rather than GM values should have been used to investigate exposure-metabolite relationships. However, they offer neither a scientific rationale nor any supporting references for the conjecture that mean rather than median values ‘must’ be used to adjust metabolite concentrations for background values, particularly in light of non-linear relationships between benzene exposures and metabolite levels.

We find this as a strange statement because we did offer a simple algebraic proof for why one must use the mean as a correction for background. The proof appears as Appendix C of the Supplementary Information, available at Carcinogenesis Online, in our paper.

Our second point, on the introduction of bias when using predictions of models that were fitted to log-transformed data, and the
references (11–13) justifying the need for a correction are presented on pages 2095–2096 of Price et al. 2012. We concluded that the impacts of the two errors operated in opposite directions and thus partially canceled each other out; however, this partial cancelation does not mean that the errors did not occur or that they did not have a net impact on the findings.

Calibration models

Because air concentrations for the Tianjin control workers were too low to be measured, a calibration model is used to derive air concentrations from urinary concentrations of benzene. Biases could be introduced from the calibration model and could play a role in the reported increases in DSM at low doses. Rappaport et al. is correct in indicating that benzene in urine is correlated with workplace air and can be used to create a calibration model (14). The question is whether a calibration model developed for high occupational exposures can be applied to much lower exposures from non-occupational sources and smoking.

Rappaport et al. calls our objection to this application of the Kim et al. (3) calibration model to the control workers an ‘artificial’ barrier. There are, however, objective reasons for questioning the use of the calibration model on the control workers. Figure 4 presents seven calibration models reported by various investigators (14–19) as well as the calibration model from Kim et al. (2). The figure limits the models’ predictions to the range of benzene air levels in the data sets used to create the various models. The figure shows that air and urine data from different workplaces involving different sources yield different calibration models and different predictions of benzene air levels associated with the same levels in urine. Although some of these differences may be due to specific factors in the various studies (duration of the air monitoring and use of personal protection by workers), it is clear that the literature does not indicate that there is a single relationship that relates urine levels to air levels for all individuals, especially when stratified by smoking.

The impact of the variation in calibration models is significant for the Kim et al. (2,3) studies. A number of calibration models were developed for workers with urine levels that fall in the range of levels observed in the 79 control workers included by Kim et al. (2,3) and Rappaport et al. in the natural spline models. As shown in Figure 4, the predicted air concentration associated with a measurement of 10nM of benzene in urine can vary by a factor of approximately 10 depending on the calibration model used.

A second concern with the calibration models is the issue of modeling smokers and non-smokers. Ghittori et al., (15) study of workers found that smokers did not follow the same calibration model as non-smokers and the correlation between urine and air levels in smokers was much weaker than in non-smokers. Use of a single model by Kim et al. (2,3) to assess both smokers and non-smokers could produce bias in the estimates of DSM because the Tianjin control workers with the highest urinary benzene levels are smokers (2).

Based upon these findings, we do not believe that our concerns about the use of the calibration model in Kim et al. (2,3) and the uncertainty it introduces into the findings are ‘artificial.’ Instead we believe that there is ample evidence that the use of a calibration model developed from one group of workers in one workplace to another group of workers in a different workplace is not warranted unless the researcher can provide evidence to show that the model’s predictions will be valid. This has not been adequately done for the Tianjin control workers.

Accounting for major sources of uncertainty

In Price et al., we explored the degree of confidence that could be placed in the natural spline models’ estimates of DSM. We investigated three sources of uncertainty: (i) the uncertainty in the natural spline model’s predictions caused by having a limited number of data points, (ii) the uncertainty in the calibration model’s predictions caused by having a limited number of data points and (iii) the unexplained variance in the predictions of the calibration model. The uncertainty analysis in Kim et al. (2) considered the first two sources of uncertainty. Kim et al. (3) investigate the first source of uncertainty and Rappaport et al. extended the uncertainty analysis of Kim et al. (3) to consider both the first and second sources of uncertainty. Rappaport et al. reported that the uncertainty in the calibration model due to the second source of uncertainty was small and did not increase the uncertainty in the predictions of DSM. Based on this, Rappaport et al. questioned our finding of larger amounts of uncertainty in the estimates of DSM than those reported by Kim et al. (3).

The findings from Rappaport et al. do not call into question the conclusions from Price et al. because our analysis also considered the third source of uncertainty, the impact of the unexplained variance in the calibration model. The Kim et al. (2) calibration model had an $R^2$ of 0.428 indicating that more than half of the variance in the air concentration data was not captured by the model. A visual inspection of Figure 2 in Kim et al. (2) shows that the same urine level can occur in individuals with air concentrations that vary by more than a factor of 100. The impact of the residual error in the model was not captured in either Kim et al. (2,3) or Rappaport et al. Instead all three papers considered the air concentrations predictions of the calibration model as if they were as certain as actual

![Fig. 4. Predicted relationships between urinary levels of benzene (nM) and airborne levels of benzene (p.p.m.) from the Kim et al. (2) and seven other published calibration models for the range of air concentrations in the data sets used to create the models.](https://academic.oup.com/carcin/article-abstract/34/7/1692/2463385)
measurements. Thus, it is not surprising that Rappaport et al. failed to duplicate our findings of large uncertainties in the prediction of DSM that preclude the determination of changes in metabolism of benzene at low doses.

Alternative definitions of background and exposed workers
In Price et al., we investigated the impacts of alternative definitions of the control and modeled populations (Approaches B and C) and concluded that there was no evidence of enhanced metabolism under Approaches B and C. Despite differences in the methodology used, Rappaport et al. confirms this finding (see Figure 5 of Rappaport et al.). Thus, we are in agreement that only Approach A indicates enhanced metabolism at low doses.

We believe that the consideration of Approaches B and C for analysis of the Tianjin data are fully justified and that either approach should be preferred over the approach advocated in Kim et al. (2,3) and Rappaport et al. (Approach A). The reason for this is that the selection of the 60 individuals with the lowest urinary benzene levels as the basis for estimating the population background levels of each metabolite in the workers’ urine in Approach A is a plausible but arbitrary decision. As demonstrated in supplemental materials to this letter, the maximum contribution of the metabolites that occur from the air exposures for the control workers can be determined using Kim’s estimates of air levels (based on the workers’ urinary benzene levels and the Kim et al. (2) calibration model) and an assumption of a metabolic fraction of 1. The draft version of this letter, which was reviewed by Dr. Rappaport, contained an error in the calculations that affected the results presented in Figure 5 and in the final section of this document. The conversion of ppm to µg/l was incorrectly performed by dividing by the conversion factor (3.19 µg/l air-ppm) instead of multiplying; see equation 3 of the supplemental materials. The error affected the determination of the maximum levels of benzene metabolites that could have occurred in the 139 control workers as a result of benzene inhalation exposure. This version of our letter has corrected this error. When this was done, the maximum contribution from air exposure to the 139 control workers’ total benzene metabolites averaged only 2.9% of the observed levels of benzene metabolites (97.2% was due to background sources). Thus, there is no objective reason for not using the mean of all 139 control workers as an estimate background levels of the workers in factories that used benzene.

An examination of the maximum contribution from air provides a second reason why the Approach A should be rejected. Under Approach A, the 79 remaining control workers not used to estimate background levels are included in the natural spline models (3). There is an increase in total benzene metabolites in the 79 control workers versus the 60 workers with the lowest urinary levels of benzene. The mean of the lowest 60 is 86.3 µM and the mean of the remaining 79 control workers is 106.7 µM, or a difference of 20.4 µM, see Figure 5. The maximum contribution to these metabolites from air was determined for each of the workers in these two groups and found to be 0.31 and 4.7 µM, respectively. The reason for the 20.4 µM difference in the average of the total metabolites in the two groups is unclear, but only 4.4 µM (4.7 minus 0.31 µM) of the 20.4 µM can be explained by differences in exposures to airborne levels of benzene. Inclusion of the 79 control workers in the natural spline models in Approach A incorrectly counts the entire 20.4 µM increase as being due to the metabolism of 4.4 µM of inhaled benzene. This implies that on average, the 79 control workers produced 4.7 (20.6 µM / 4.4 µM) times the urinary metabolites that could be explained by the workers’ inhalation exposures to benzene. As a result Approach A should not be used in the analysis of the worker data.

References

Received February 14, 2013; accepted March 16, 2013