Genetic toxicity data have traditionally been employed for qualitative, rather than quantitative evaluations of hazard. As a continuation of our earlier report that analyzed ethyl methanesulfonate (EMS) and methyl methanesulfonate (MMS) dose–response data (Gollapudi et al., 2013), here we present analyses of 1-ethyl-1-nitrosourea (ENU) and 1-methyl-1-nitrosourea (MNU) dose–response data and additional approaches for the determination of genetic toxicity point-of-departure (PoD) metrics. We previously described methods to determine the no-observed-genotoxic-effect-level (NOGEL), the breakpoint-dose (BPD; previously named Td), and the benchmark dose (BMDL) for genetic toxicity endpoints. In this study we employed those methods, along with a new approach, to determine the non-linear slope-transition-dose (STD), and alternative methods to determine the BPD and BMD, for the analyses of nine ENU and 22 MNU datasets across a range of in vitro and in vivo endpoints. The NOGEL, BMDL, and BMDL PoD metrics could be readily calculated for most gene mutation and chromosomal damage studies; however, BPDs and STDs could not always be derived due to
data limitations and constraints of the underlying statistical methods. The BMDL\textsubscript{10} values were often lower than the other PoDs, and the distribution of BMDL\textsubscript{10} values produced the lowest median PoD. Our observations indicate that, among the methods investigated in this study, the BMD approach is the preferred PoD for quantitatively describing genetic toxicology data. Once genetic toxicology PoDs are calculated via this approach, they can be used to derive reference doses and margin of exposure values that may be useful for evaluating human risk and regulatory decision making. Environ. Mol. Mutagen. 55:609–623, 2014. © 2014 The Authors. Environmental and Molecular Mutagenesis Published by Wiley Periodicals, Inc.

Key words: benchmark dose; ENU; MNU; alkylating agents; margin of exposure

INTRODUCTION

Until quite recently genotoxicity test results were employed almost exclusively for dichotomous qualitative evaluations (i.e., results classified as either a positive or negative), with studies routinely evaluating responses at very high doses (i.e., near the maximum tolerated dose or MTD). Interest in the manifestation of genotoxicity at low doses, as well as quantitative analyses of the dose–response data, has been limited (Pottenger and Gollapudi, 2009). In contrast, quantitative dose–response analyses and derivation of point-of-departure (PoD) metrics are routinely employed to assess other toxic effects that are not mediated by genotoxic mechanisms (Piersma et al., 2011) as well as carcinogenic risk (EFSA, 2009). Such PoD values are routinely used for risk assessment in conjunction with uncertainty factors to derive health-based guidance values and regulatory limits to assess and manage risks. In the case of genotoxic carcinogens, if the mode of action is not established, then a conservative, linear approach (i.e., linear low-dose extrapolation from the PoD to origin) is generally taken. Thus, there is theoretically no level of exposure for such a chemical that does not pose a small, but finite probability of generating a carcinogenic response. This assumption may not always hold true, because there is increasing recognition that non-linear dose responses are observed with genotoxic endpoints for at least some substances, and there is increasingly strong mechanistic evidence to support the calculation and use of biologically meaningful PoDs to inform regulatory decision making for genotoxic agents. For example, several recent publications have demonstrated that biologically meaningful, sub-linear dose–response functions exist for both non-DNA-reactive genotoxicants [mitotic spindle poisons (Johnson and Parry, 2008; Elhajoui et al., 2011)] and at least some DNA-reactive mutagens (Doak et al., 2007; Gocke and Wall, 2009; Johnson et al., 2009; Pottenger et al., 2009; Bryce et al., 2010; Gollapudi et al., 2013). This recognition has contributed to an increasing appreciation of the utility of the quantitative analysis of genetic toxicity dose–response relationships; and moreover, to employ quantitative methods and PoD determination for genotoxicity data to use in regulatory decision making.

The Quantitative Analysis Workgroup (QAW) of the Genetic Toxicology Technical Committee (GTTC) coordinated by the Health and Environmental Sciences Institute (HESI) of the International Life Sciences Institute (ILSI) is involved in the development and critical examination of methodologies for the quantitative analysis of in vitro and in vivo genotoxicity dose–response data, and the development of strategies for the use of PoD metrics to support regulatory evaluations and decision making (Gollapudi et al., 2011). In this report, the GTTC QAW extends the analyses presented earlier (Gollapudi et al., 2013) that addressed the applicability of several statistical methods for the analyses of genetic toxicology dose–response data. Collectively, this and our previous report contribute to a rapidly growing body of knowledge regarding the use of quantitative dose–response analyses to derive PoD metrics, and moreover, to employ PoD values to assess the risk of adverse health effects in humans and/or to determine exposure levels that would be associated with negligible risk.

Gollapudi et al. (2013) used data from studies of methyl methanesulfonate (MMS) and ethyl methanesulfonate (EMS) to investigate the utility of several metrics to define a PoD for use in determining regulatory limits associated with negligible risk of genotoxic effect. These metrics included (i) no-observed-genotoxic-effect-levels (NOGELs), (ii) statistically defined break points, now referred to as the breakpoint dose (BPD, previously termed “threshold” dose, Td), and (iii) benchmark dose (BMD) levels (Gollapudi et al., 2013). This report addresses the need to extend this initial investigation to encompass additional agents, to critically examine additional modeling techniques, to improve the biological understanding of the mode(s)-of-action (MOAs) that determine(s) the shape(s) of genotoxicity exposure-response curves, and to define strategies to employ PoD metrics for regulatory evaluations and decision making. Such work is essential, and will precede general acceptance regarding the determination and use of meaningful PoD metrics in genetic toxicology. The derivation of PoDs in genetic toxicology studies and their routine use for risk assessment and regulatory decision making will require the following:

1. Selection of appropriate mathematical models and statistical methods for reliable PoD determinations;
2. Determination of study design features that facilitate quantitative dose–response analyses, and,
3. Development of standardized methods for the incorporation of PoD metrics into human health risk assessment (e.g., application of uncertainty and/or safety factors and/or margin of exposure analysis to define exposures that are associated with negligible risk).

The current work extends our earlier efforts to address points 1, and 2 above, and to initiate discussions on point 3. This was carried out using the data on two potent alkylating mutagens with large data bases, 1-methyl-1-nitrosourea (MNU) and 1-ethyl-1-nitrosourea (ENU) and, building on our earlier analyses of the related, but less potent alkylating agents EMS and MMS, enabling further development of our “toolbox” for the derivation and use of PoD metrics in genetic toxicology.

**MATERIALS AND METHODS**

**Data Selection**

The database employed, referred to as the G4 database, was developed by the GTTC QAW, and details of the quality criteria used for data screening are in Gollapudi et al., (2013). In addition to the EMS and MMS datasets, the database contains a total of 45 datasets for ENU and MNU, including endpoints for gene mutation and chromosomal damage (measured as micronucleus formation) in vitro and in vivo. Since the goal of this work was to analyze datasets from which four PoD values could be derived (i.e., NOGEL, BMD, BPD, and STD), only datasets with data supporting PoD derivation were evaluated. Based on the recommendations of Lutz and Lutz (2009), we further restricted the analyses to datasets with ≥ 5 doses (including the negative control) to ensure that we could use the bilinear modeling approach (i.e., BPD modeling).

**Benchmark Dose Analysis**

The benchmark dose (BMD) is defined in the U.S. Environmental Protection Agency’s (EPA’s) risk assessment glossary as ‘A dose or concentration that produces a predetermined change in response rate of an adverse effect (called the benchmark response or BMR) compared to background’ (http://www.epa.gov/riskassessment/glossary.htm). For BMD determination, the latest versions of the PROAST BMD software and EPA’s Benchmark Dose Software (BMDS) were employed (EPA, 2013).

**RIVM-PROAST Benchmark Dose Analysis**

BMD analysis was conducted using PROAST, the dose–response modeling software developed at the National Institute for Public Health and the Environment (RIVM) in The Netherlands [http://www.proast.nl; version 36.9 (Slob, 2002)]. The nested set of models used included the exponential and Hill models that are recommended by the European Food Safety Authority (EFSA, 2009). Note that the bilinear models used to define BPD metrics are not included, and thus are not assessed in these BMD packages (Slob and Setzer, 2013). The BMR examined was 10%, which corresponds to an increase equal to 10% of the background (negative control) level, as estimated by the fitted model used for continuous endpoints such as carcinogenicity data. The lower limit of the one-sided 90% confidence interval on the BMD is termed the BMDL, and the BMDL10 refers to the estimate of the lower 90% confidence limit of a dose that produces a 10% increase over the fitted background level for continuous endpoints. The BMDL10 is the upper limit of this 90% confidence interval. Model selection was performed based on a log-likelihood ratio test that assesses whether including additional parameters to the model results in a statistically significant improvement in model fit (Hernández et al., 2011a). The model with additional parameters is accepted only if the difference in log-likelihoods exceeds the critical value at \( P = 0.05 \). In addition, the log-likelihood is used to compare the “full” model (geometric means of the observations at each dose) to the selected model, to provide an indication of the goodness-of-fit. The distribution of the residual errors in PROAST is similar to that derived using the EPA’s BMDS (discussed below). For continuous data, the residual errors are assumed to be log-normally or normally distributed. There is an option in both software packages to choose one or the other distribution. In PROAST, the default assumption is that the standard deviation is proportional to the mean, and thus a log-normal distribution is applied to the continuous data, whereas in BMDS the default setting is the normal distribution.

**U.S. EPA Benchmark Dose Software**

BMD analysis also was conducted using the latest version of the EPA’s Benchmark Dose Software [i.e., BMDS v2.4 (EPA, 2013)]. The standard suite of continuous models (Hill, exponential, polynomial, linear, and power) was used along with constant and non-constant variance model assumptions. As was the case for RIVM-PROAST, the bilinear models used to define BPD metrics are not included, and thus are not assessed in this BMD package. The BMR chosen for the BMDS software was an increase equivalent to one standard deviation above the spontaneous (control) value; this was used in the calculation of a one-sided 95% lower confidence limit for this BMD value and designated BMDL10. This BMR is recognized to be equivalent to ~10% excess risk for individuals below and above the 2nd and 98th percentiles, respectively (Crump, 1995). Model selection was primarily based upon the \( P \)-value for goodness-of-fit to the data and the Akaike’s Information Criterion (AIC). For this study, the best-fitting model was selected among the suite of continuous models (see above). Log-transformation was used as a default for analysis of continuous data with PROAST (Slob and Setzer, 2013). For BMDS, continuous data were transformed in the same manner used for NOGEL, BPD, and smoothing regression spline modeling (discussed in next section).

**NOGEL, BPD, and Smoothing Regression Spline Analyses**

NOGEL, BPD, and Smoothing Regression Spline analyses were used to extend the previous effort (Gollapudi et al., 2013) to examine the utility of open source methodologies and additional approaches.

**Initial Statistical Evaluation**

All data sets were imported into R (R Development Core Team, 2011) and the following analyses conducted: Shapiro–Wilk Normality Test, Bartlett Test of Homogeneity of Variances, Jonckheere–Terpstra Test of Monotonic Trend (asymptotic version), and a Bonferroni test for outlier identification. Data were transformed in order to achieve normally distributed data and homogeneity of the within dose variances. In most cases where the original data were not normally distributed and/or the variances were heterogeneous, one of the data transformation processes (square root or logarithmic) performed on the response data
resulted in datasets with satisfactory distributions and variances, and the transformation essentially eliminated any significant outliers identified by the Bonferroni test. If transformation of data was necessary for a dataset, then all remaining analyses on that dataset were conducted on the transformed data. Based on these preliminary statistical tests, datasets that were non-normally distributed and with heterogeneous variances, and therefore was only used on data where the data transformation was successful. Although ‘segmented’ can be used in conjunction with weights to account for variance heterogeneity across dose groups, this approach was not used for analyses reported here. Therefore, ‘segmented’ was also only used on normally distributed data-sets with homogeneous variances.

**Smoothing Regression Spline**

We also used penalized smoothing splines to analyze the dose–response relationships. Penalized smoothing splines are a family of flexible techniques for estimating a continuous functional relationship without the need to assume linearity or any specific non-linear functional form. Wood (2006) has extensively developed the underlying theory. The ‘mgcv’ (Mixed GAM [generalized additive model] Computation Vehicle) package in R (Wood, 2006, 2011) was used to estimate the dose–response function using a default thin plate smoothing regression spline with degrees of freedom determined by generalized cross validation. Unlike the bilinear model used with the ‘segmented’ algorithm, ‘mgcv’ is commonly regarded as semi-parametric because it can be applied to non-normally distributed data with heterogeneous variances, and still provide an optimal solution from a cross-validation perspective. Thus, smoothing spline regression was used for PoD determination irrespective of whether the data required a parametric or non-parametric test (see Fig. 1).

The derived continuous non-linear dose–response function was subjected to finite differencing to calculate its first derivative, or slope, using the L&L and segmented models. At all lower doses, the slope of the dose–response function is the lowest dose at which the dose–response function has a slope that is significantly above zero with 95% confidence using a one-sided confidence limit. As above, this particular limit was chosen to match the output of PROAST, BMDS, and the L&L model. At all lower doses, the slope of the dose–response function is not statistically distinguishable from zero, therefore represents a flat line. When the lower bound CI on the STD (i.e., STDL) is \( \leq 0 \), the hypothesis that the slope is increasing significantly at dose = 0 (i.e., slope is above zero) cannot be rejected. This is conceptually different from the L&L and segmented bilinear models, where a BPDL \( \leq 0 \) means that linearity cannot be rejected (regardless of the slope of that linearity). The smoothing regression spline approach arrives at a PoD by directly assessing the slope of the dose–response relationship throughout the dose range to determine the dose where it initially becomes positive.

In addition to deriving potential PoDs, the smoothing regression spline model also was used to test the overall linearity of the dose response of a given dataset. To this end, we tested whether the smoothing regression spline model fit the overall dose response significantly better than the null model, i.e., the best-fitting two segment linear function where the first segment from zero dose to the breakpoint is horizontal (i.e., has zero slope) and the second segment has a positive slope. However, the segmented approach has several advantages: it is based on an open source, peer reviewed package available in R, and, unlike the L&L model, it does not require removal of top doses due to supra-linearity (saturation or high-dose toxicity). When the data were normally distributed but with heterogeneous variances, different pathways of statistical analyses and variances were compared using a post-hoc Dunnett’s Test (alpha = 0.05). Means for datasets that were normally distributed but with heterogeneous variances were compared using a post-hoc Dunnett’s T3 test (Field, 2009). Means from datasets that were non-normally distributed and with heterogeneous variances were compared using the non-parametric post-hoc Dunn’s Test (Laws et al., 2000).
| Chemical | Endpoint | Gene target | Type | Species | Cell type/tissue | Units | # of doses | # of replicates | Treatment regime | Expression timea | Study |
|----------|----------|-------------|------|---------|-----------------|-------|------------|----------------|-----------------|----------------|-------|
| ENU      | Gene mutation | Dlb1 | vv       | Mouse | SI | mg/kg | 5 | 3-4 | 1 d/i.p. | 49 d | van Delft et al. (1998) |
| ENU      | Gene mutation | LacZ | vv       | Mouse | SI | mg/kg | 5 | 4 | 1 d/i.p. | 49 d | van Delft et al. (1998) |
| ENU      | Gene mutation | LacZ | vv       | Mouse | Spleen | mg/kg | 5 | 4 | 1 d/i.p. | 49 d | van Delft et al. (1998) |
| ENU      | Gene mutation | Ptg-a | vv       | Mouse | RET | mg/kg | 7 | 5 | 1 d/i.p. | 14 d | Bhalli et al. (2011) |
| ENU      | Gene mutation | Ptg-a | vv       | Mouse | RBC | mg/kg | 7 | 5 | 1 d/i.p. | 28 d | Bhalli et al. (2011) |
| ENU      | Gene mutation | HPRT | vt      | Human | AHH-1 | μg/mL | 12 | 3 | 24 hr | 24 d | Bhalli et al. (2011) |
| MNU      | Gene mutation | Hprt | vv       | Mouse | Spleen | mg/kg | 5 | 9-10 | 1 d/i.p. | 21 d | Monroe et al. (1998) |
| MNU      | Gene mutation | LacI | vv       | Mouse | Spleen | mg/kg | 5 | 6-7 | 1 d/i.p. | 21 d | Monroe et al. (1998) |
| MNU      | Gene mutation | Ptg-a | vv       | Rat | RET | mg/kg | 7 | 3 | 28 d/gavage | 15 d | Lynch et al. (2011) |
| MNU      | Gene mutation | Ptg-a | vv       | Rat | RBC | mg/kg | 7 | 3 | 28 d/gavage | 29 d | Lynch et al. (2011) |
| MNU      | Gene mutation | Ptg-a | vv       | Rat | RBC | mg/kg | 5 | 6 | 28 d/gavage | 4, 15, 29 d | BMS (unpublished) |
| MNU      | Gene mutation | Hprt | vt      | Human | AHH-1 | μg/mL | 15 | 2-3 | 18 hr | 0 hr | Bhalli et al. (2011) |
| MNU      | Gene mutation | Hprt | vt      | Human | AHH-1 | μg/mL | 10 | 5 | 4 hr | 2 d | Pottenger et al. (2009) |
| MNU      | Gene mutation | Hprt | vt      | Human | AHH-1 | μg/mL | 10 | 5 | 24 hr | 13 d | Thomas et al. (2013) |
| MNU      | Gene mutation | Hprt | vt      | Human | AHH-1 | μg/mL | 8 | 6 | 4 d/gavage | 24 hr | LeBaron (2009) |
| MNU      | Gene mutation | Hprt | vt      | Human | AHH-1 | μg/mL | 8 | 6 | 4 d/gavage | 24 hr | LeBaron (2009) |
| MNU      | Gene mutation | Hprt | vt      | Human | AHH-1 | μg/mL | 8 | 6 | 28 d/gavage | 4, 29 d | Lynch et al. (2011) |
| MNU      | Gene mutation | Hprt | vt      | Human | AHH-1 | μg/mL | 8 | 6 | 28 d/gavage | 4, 29 d | Lynch et al. (2011) |
| MNU      | Gene mutation | Hprt | vt      | Human | AHH-1 | μg/mL | 8 | 6 | 28 d/gavage | 4, 29 d | BMS (unpublished) |
| MNU      | Gene mutation | Hprt | vt      | Human | AHH-1 | μg/mL | 9 | 4 | 18 hr | 0 hr | Doak et al. (2007) |

vv, in vivo; vt, in vitro; RET, reticulocytes; RBC, red blood cells; NCE, non-chromatic erythrocytes; PCE, polychromatic erythrocytes; ip, intraperitoneal injection; SI, small intestine; d, day; hr, hour.
aAssume post treatment unless multiple time points are listed, indicating time after initial dosing.
bStudy 2 from Lynch et al. (2011).
cBMS unpublished, Bristol–Myers Squibb unpublished data provided to HESI for incorporation into the G4 database.
| Study                                      | Endpoint Transformtion | Type     | Trend | Response | LOGEL | NOGEL | Sentinel | BMDL10 | BMDL1SD |
|-------------------------------------------|------------------------|----------|-------|----------|-------|-------|----------|--------|---------|
| VanDelft (1998) Dlb1_SI Mouse             | GM                     | Raw      | LogR  | NOGEL    | 7.0E-07 | 1.3E-06 | NA       | 1.55   | NA      |
| VanDelft (1998) L&L_SI Mouse              | GM                     | LogR     | NA    | NA       | 6.5E-07 | 1.3E-06 | 0        | 5.46   | 1.94    |
| VanDelft (1998) LacZ_SI Mouse             | GM                     | LogR     | NA    | NA       | 6.0E-07 | 1.3E-06 | 0        | 11.67  | 6.65    |
| VanDelft (1998) LacZ_Spleen               | GM                     | Raw      | 3.5E-02 | 0       | 3.8E-02 | 3.8E-02 | ID       | 12.0   | 5.20    |
| Doak (2007) AHH1_Human HPRT               | MN                     | LogR     | 4.5E-02 | 0       | 4.5E-02 | 4.5E-02 | 0        | 21     | 12      |
| Doak (2007) AHH1_Human HPRT               | MN                     | Raw      | 9.3E-02 | 0       | 9.3E-02 | 9.3E-02 | 0        | 20     | 12      |
| Bryce (2010) TK6_Human Expt 1            | MN                     | LogR     | 1.0E-02 | 0       | 1.0E-02 | 1.0E-02 | 0        | 21     | 12      |
| Bryce (2010) TK6_Human Expt 2            | MN                     | LogR     | 1.0E-02 | 0       | 1.0E-02 | 1.0E-02 | 0        | 21     | 12      |
| Bryce (2010) TK6_Human Expt 3            | MN                     | LogR     | 1.0E-02 | 0       | 1.0E-02 | 1.0E-02 | 0        | 21     | 12      |
| Bryce (2010) TK6_Human Expt 4            | MN                     | LogR     | 1.0E-02 | 0       | 1.0E-02 | 1.0E-02 | 0        | 21     | 12      |
| Bryce (2010) TK6_Human Expt 5            | MN                     | LogR     | 1.0E-02 | 0       | 1.0E-02 | 1.0E-02 | 0        | 21     | 12      |

**RESULTS**

Table I summarizes the individual study characteristics for the various genotoxicity datasets analyzed. A wide range of endpoints were analyzed, including in vivo and in vitro measures of micronuclei (MN), in various cell types and species. Similarly, gene mutations were assessed in multiple target genes including transgenes (e.g., LacZ) and endogenous genes (e.g., Hprt). The PoD metrics for nine ENU genotoxicity datasets are summarized in Table II. NOGEL values were obtained for all datasets with the exception of the Dlb1 mutation analyses in the small intestine, where the lowest study dose was significantly different from control. BMDL10 values were determined using PROAST for reasons stated in Gollapudi et al. (2013), whereas BMDL1SD values were determined using BMDS. The latter is the default metric used by the EPA for continuous data (U.S. EPA, 2012). In all cases, an exponential model provided the best fit in PROAST. In BMDS, the best fitting model was selected among the typical suite of continuous models. Overall, although not all PoD methods provided a good fit of the data to a statistical model, the two BMD modeling approaches provided estimates of the BMDL values that were lower than the corresponding NOGELs.

PoD values were derived for four of the ENU datasets using at least one of the three models; L&L, segmented, and smoothing regression spline. In all four cases, the slope below the LOGEL was not significantly different from zero, suggesting a bilinear dose response. Two of the datasets for which BPD values were derived were not predicted to have an STD by mgcv (van Delft et al., 1998; Doak et al., 2007). Examination of the mgcv plots of these datasets reveals an apparent lack of bi-linearity, consistent with the model results failing to identify an STD (Table II). In contrast, the two datasets from Bryce et al. (2010) were predicted to have STD values generated by smoothing regression spline analyses (i.e. exhibited non-linear dose responses with estimated STDL values).

The PoD metrics for 22 MNU genotoxicity datasets are summarized in Table III. As with ENU, the two BMD modeling approaches yielded lower PoD values than the
| Study                        | Type   | Endpoint | Response | Trend | Transformation | Slope < NOGEL | Linear < NOGEL | NOGEL Test | L&L | mgcv | segmented BPDL | BPDL | BMDS | BMDS SD | PROAST | BMDL10 | BMDL10 | Units |
|-----------------------------|--------|----------|----------|-------|----------------|--------------|---------------|-------------|-----|-----|----------------|------|------|---------|---------|--------|--------|-------|
| Monroe (1998) HPRT_Spleen   | vv     | GM       | LogR     | 2.80E+13 | 0               | yes          | Dunnett’s 5   | 0.78        | 1.40| 0.98| 3.90           | 2.06 | 2.32 | mg/kg   |
| Monroe (1998) LacI_Spleen   | vv     | GM       | LogR     | 0.0005  | 0               | yes          | Dunnett’s 15  | 8.22        | no  | STD | 10.62          | 11.53 | 5.26 | 3.09 mg/kg |
| Lynch (2011) Pig-a_RET Rat  | vv     | GM       | Raw      | 0.048   | NA              | NA           | Dunnett’s 2.5 | 0.77        | 0.10| 4.54 | mg/kg          |
| Lynch (2011) Pig-a LICENSE Rat  | vv     | GM       | SqrR     | 0.02    | 0               | yes          | Dunnett’s 1.25| 0.77        | 0.11| 2.61 | mg/kg          |
| BMS Pig-a_RBC Rat Day 4     | vv     | GM       | LogR     | 0.26    | NA              | NA           | Dunnett’s 5   | NA          | 3.75 | No DR | No DR mg/kg   |
| BMS Pig-a_RBC Rat Day 15    | vv     | GM       | LogR     | 3.00E-06| +               | yes          | Dunnett’s 2.5 | NA          | 1.57 | No DR | No DR mg/kg   |
| BMS Pig-a_RBC Rat Day 29    | vv     | GM       | LogR     | 8.50E-10| NA              | NA           | Dunnett’s None| no BPDL     | no  | STD | no BPDL        |
| BMS Pig-a_RET Rat Day 4     | vv     | GM       | LogR     | 0.53    | NA              | NA           | Dunnett’s 5   | NA          | 4.33 | No DR | No DR mg/kg   |
| BMS Pig-a_RET Rat Day 15    | vv     | GM       | LogR     | 1.70E-07| 0               | ID           | Dunnett’s 0.9 | no BPDL     | no  | STD | no BPDL        |
| BMS Pig-a_RET Rat Day 29    | vv     | GM       | LogR     | 3.50E-08| NA              | NA           | Dunnett’s None| no BPDL     | no  | STD | no BPDL        |
| LeBaron (2009) PCE Rat      | vv     | MN       | LogR\(^b\) | 9.20E-09| +               | yes          | Dunnett’s 1   | NA          | 0.60 | NA  | 0.08           | 0.02 | 2.42 | mg/kg   |
| LeBaron (2009) NCE Rat      | vv     | MN       | LogR\(^b\) | 0.75    | NA              | NA           | Dunnett’s 50  | NA          | no  | STD | 42.2 mg/kg     |
| Lynch (2011) PCE_Event 1 Rat| vv     | MN       | Raw      | 4.50E-06| +               | yes          | Dunnett’s 0.9 | 0.20        | 0.27| 0.16| 0.10 mg/kg    |
| Lynch (2011) PCE_Event 2 Rat| vv     | MN       | LogR     | 6.10E-07| +               | yes          | Dunnett’s 0.6 | 0.15        | 1.30| 0.73| 3.12 mg/kg    |
| Lynch (2011) NCE_Event 1 Rat| vv     | MN       | Raw      | 0.36    | NA              | NA           | Dunnett’s 2.5 | no BPDL     | 1.17| 0.30| 1.30 mg/kg    |
| Lynch (2011) NCE_Event 2 Rat| vv     | MN       |Raw      | 0.002   | +               | yes          | Dunnett’s 1.25| no BPDL     | no  | STD | no BPDL        |
| BMS Day 4 Rat               | vv     | MN       | LogR     | 8.80E-09| +               | ID           | Dunnett’s 1.25| NA          | 0.42| 0.30| 2.66 mg/kg    |
| BMS Day 29 Rat              | vv     | MN       | LogR     | 3.60E-09| NA              | NA           | Dunnett’s None| no BPDL     | no  | STD | no BPDL        |
| Doak (2007) AHG1_Human       | vt     | GM       | SqrR     | 1.00E-14| +               | yes          | Dunnett’s 0.005| no BPDL     | no  | STD | no BPDL        |
| Doak (2007) AHG1_Human       | vt     | GM       | LogR\(^b\) | 2.00E-08| 0               | yes          | Dunnett’s 0.69| 1.01        | 0.49| 1.03| 0.83 mg/mL    |
| Thomas (2013) AHG1_Human     | vt     | GM       | SqrR     | 0.0004  | 0               | no           | Dunnett’s 0.007| 0.002       | no  | STD | no BPDL        |
| Thomas (2013) AHG1_Human     | vt     | GM       | LogR\(^b\) | 0       | +               | no           | Dunnett’s 0.008| 0.006       | 1.33| 1.33 | \(\mu\)g/mL\ |
| Bryce (2010) TK6, Human Expt 1 | vt | MN | LogR\(^b\) | 0.45E-14 | +               | no           | Dunnett’s 0.23| NA          | 0.20| 0.46| 1.09 \(\mu\)g/mL\ |
| Bryce (2010) TK6, Human Expt 2 | vt | MN | LogR\(^b\) | 2.50E-09 | +               | yes          | Dunnett’s 0.025| no BPDL     | no  | STD | no BPDL        |

\(\text{vv}, \text{in vivo}\); \(\text{vt}, \text{in vitro}\); NA, not applicable; ID, insufficient doses; GM, gene mutation; MN, micronucleus; No DR, no dose response, BMS, Bristol-Myers Squibb unpublished data; SI, small intestine; +, positive gradient; NOGEL, no observed genotoxic effect level; BPDL, breakpoint dose; BPDSL, breakpoint dose lower confidence interval; STD, slope transition dose; STDSL, slope transition dose lower confidence interval; BMDLSD, benchmark dose 1 standard deviation lower confidence interval; BMDL10, benchmark dose 10 lower confidence interval, BMDL10, benchmark dose 10 upper confidence interval; L&L, Lutz and Lutz, 2009. Underlined PoD values were obtained after dropping high dose(s).

\(^a\) Poor fit for benchmark dose model, \(P < 0.05\).

\(^b\) Doses log transformed as well.

Response Transformation, same number added to ‘R’ to ensure all responses were above the value of 1 before transformation with Log or Sqrt.

\('\text{Slope < NOGEL}' tests whether slope up to and including the NOGEL differs significantly from zero.

\('\text{Linear < NOGEL}' tests whether slope up to and including the NOGEL is fit better by linear or nonlinear model (i.e., smoothing regression spline).
NOGEL. There were two MNU datasets for which both the bilinear and smoothing regression spline methods provided BPDLs and STDLs, respectively (Monroe et al., 1998; Pottenger et al., 2009). The PoDs associated with these data, calculated by the different aforementioned methods, were in remarkably close agreement for a given dataset, and the slopes up to and including the NOGEL were not significantly different from zero, demonstrating a good fit for the bilinear models (Fig. 2). There were eight datasets that required analysis with non-parametric methods, and the application of the smoothing regression spline methodology yielded STD values for five of these datasets, with the slopes from the negative control up to and including the NOGEL not being consistent with a zero slope. Therefore, these STD values should be interpreted with caution (see below).

Eighteen in vivo MNU datasets are summarized in Table III. Although the studies examined different endpoints in different species under different exposure scenarios, a quantitative comparison of the PoDs can still be conducted (Fig. 3). The medians (2.5 mg/kg/day) and distributions of NOGEL and LOGEL values were very similar (Fig. 3A). The BPD, STDL, and BMDL1SD were also in close agreement, ranging from 0.8 to 1.2 mg/kg/day. The median BPD, BMD10, and BMDL10 were 0.3, 0.2 and 0.1 mg/kg/day, respectively. To get a sense of how these PoD metrics compare to one another within each in vivo MNU dataset, each PoD was ‘normalized’ to the LOGEL of that dataset (if available). As expected, the ratio of the NOGEL to the LOGEL was less than unity (i.e., approximately 0.5; Fig. 3B). The ratios of the BMDL1SD and BPD to LOGELs were similar, as were the ratios of STDL and BPDL to LOGELs. The ratio of the BMDL1SD to the LOGEL was generally higher than the BPD and STD ratios, consistent with benchmark dose values being associated with a pre-defined increase in toxicological response. A similar pattern was observed for the BMDL10 relative to the BPDL and STDL. The median ratios of the BMDL10 values to the LOGELs were the lowest of all PoDs, and the BMDL10 values were often lower than the BPDL and STDL ratios (Figs. 3A–B).

DISCUSSION

The current work, which follows that of Gollapudi et al. (2013), is focused on advancing the development and application of statistical approaches to define PoDs for genotoxicity dose–response data. This is a necessary step in the path forward for the use of genotoxicity PoD metrics to inform regulatory decision making and/or risk assessment. The earlier work focused on the following PoD metrics: Dunnett’s approach for calculating NOGELs; a multi-step approach to define BPD values (Gocke and Wall, 2009); and the BMD approach (i.e., using PROAST and BMDS software). Here we present the next phase in these analyses, which involved a series of approaches (Table IV) to determine PoDs for the potent genotoxicants ENU and MNU. The data from the examined publications, in most cases, suggest that PoD values are generally lower and more difficult to define for ENU and MNU in comparison with EMS and MMS. This
is likely because ENU and MNU are more potent mutagens, probably related to the higher proportion of O6-alkylG and other pro-mutagenic adducts these SN1 alkylators form (Jenkins et al., 2005).

Evaluation and Comparison of Different Approaches for Determination of Point of Departure Values

Table IV presents a comparative assessment of the advantages and disadvantages associated with each PoD metric. The NOGEL value is the highest experimental exposure level at which there is no statistically significant increase above the concurrent experimental control value (background level). NOGELs are, by definition, dependent on study design features such as dose selection and the statistical power to detect an increase at each dose. Furthermore, this approach does not permit calculation of PoD confidence intervals. When comparing NOGELs to the other PoD metrics, one can see that NOGELs are almost always higher than either BMDLs or BPDLS (Tables II and III; Fig. 3), and therefore may provide less conservative estimates than the other PoD values. They may also be less preferred because of their dependence on the specific doses tested.

As part of this effort to expand the number of data analysis tools for determination of PoDs, other approaches beyond the L&L and BMD approaches were examined and evaluated. The L&L and segmented models provided similar results (Tables II and III). However, the segmented PoD may be viewed as a more reliable metric since it does not require dataset censoring at the highest doses to address saturation or high-dose toxicity, and moreover, it is a well-documented R procedure. The segmented package also contains functions that directly account for variance heterogeneity and non-normality of residuals via weights, and conducts these analyses within the framework of generalized linear models (GLMs). GLMs can directly model data sets such as cell or colony counts, incidence frequencies, and other types of data that do not generally adhere to the normality and variance constancy assumptions. Therefore, we propose that the segmented package supersede the L&L package as the preferred bilinear modeling approach for assessing genetic toxicity data. The STD method, which uses a smoothing regression spline approach also within a GLM framework, directly assesses the slope of the dose-response relationship for a continuous non-linear dose-response function rather than for two linear segments (bilinear model); it provided PoD estimates that were consistent with those derived by other methods (Fig. 3). A potential advantage of the smoothing regression spline approach compared to bilinear modeling, is that it can be compared directly to different non-linear models such as the exponential, Hill and quadratic models. Thus we recommend the smoothing regression spline method applied here in preference to the segmented and L&L models. However, this method shares a major disadvantage similar to the BPD approaches, in that it is frequently not possible to derive a PoD with this method. For example, a PoD was determined for only 15/34 using smoothing regression spline and segmented methods (Table IV), whereas PoDs were determined for 30/34, 31/34, and 34/34 datasets when using NOGEL, BMD10, and BMD1SD, respectively (Tables II and III).

Similar to observations in the previous evaluations of EMS and MMS data (Gollapudi et al., 2013), the BMD...
TABLE IV. Descriptions, Advantages, Disadvantages, and Potential Limitations of the PoD Metrics Examined in This Study

| Full name | Definition | Defined using | Other output metrics | Potential limitations | Advantages |
|-----------|------------|---------------|----------------------|-----------------------|------------|
| NOGEL     | No observed genotoxic effect level | The highest tested dose for which there is no statistically significant increase in genotoxic effect compared to the control. | Dunnett’s, Dunnett’s Td, Dunn’s | LOGEL, $P$ value | Lower power tends to provide larger PoDs; Statistical assumptions must be met; Highly dependent on the study design, e.g., dose selection and dose spacing | Easy to apply; Does not require dose response modeling; Commonly defined. In Tables II and III, it was calculated for 30/34 datasets |
| BPD       | Breakpoint dose | The dose at which the slope changes from zero (horizontal) to positive, with its standard error forming the confidence bounds (90% CI) | L&L | BPDL, BPDU, $P$ value, $y$ intercept, gradient after BPD | Based on one model; Inflexible in terms of ability to account for other functional forms; Ability to define a BPD is highly dependent on the study design, e.g., dose selection and dose spacing. Not commonly defined. In Tables II and III it was calculated for only 8/34 datasets using L&L and 15/34 using segmented, furthermore when the prerequisites were considered, this was reduced to 2/34 and 3/34 respectively. | Lower power tends to provide smaller PoDs; May be appropriate when mechanistic data are available |
| STD       | Slope Transition Dose | The lowest dose for which the lower bound of the 95% confidence interval of the slope exceeds zero | mgcv | STDL, STDU, $P$ value | Plotting of the non-linear model is possibly too flexible compared to pre-defined models; Model is still being developed and validated; Ability to define an STD is highly dependent on the study design, e.g., dose selection and dose spacing. In Tables II and III it was calculated for only 15/34 datasets using L&L and 15/34 using segmented, furthermore when the prerequisites were considered, this was reduced to 2/34 and 3/34 respectively. | Lower power tends to provide smaller PoDs; May be appropriate when mechanistic data are available; Unlike the BPD, can readily be compared to other models e.g. quadratic; Less affected by distribution and variance |
| BMD<sub>10</sub> | Benchmark Dose 10 | A dose that produces a 10% increase over the fitted background | PROAST | BMDL<sub>10</sub>, BMDU<sub>10</sub>, BMDU/BMDL ratio | Requires consensus on appropriate biologically relevant benchmark response (BMR); Continuous and quantal data are modeled differently; Often produces very low BMDL metrics | Lower power tends to provide smaller PoDs; Fits function to entire dose–response, not just the tested doses; Currently used by many regulatory agencies; Suitable for use in comparing PoD metrics between different endpoints; Requires fewer data points than BPD and STD; Commonly defined. In Tables II and III it was calculated for 31/34 datasets |
| BMD<sub>1SD</sub> | Benchmark Dose 1 Standard Deviation | ~10% excess risk for individuals below and above the 2nd and 98th percentiles, respectively | BMDS | BMDL<sub>1SD</sub>, BMDU<sub>1SD</sub>, BMDU/BMDL ratio | Requires consensus on appropriate biologically relevant benchmark response (BMR); Continuous and quantal data are modeled differently; Comparisons between endpoints and historical datasets more influenced by background level and variance than the BMD<sub>10</sub> approach | Lower power tends to provide smaller PoDs; Fits function to entire dose–response, not just the tested doses; Currently used by many regulatory agencies; Requires fewer data points than BPD and STD; Commonly defined. In Tables II and III it was calculated for 34/34 datasets |

NOGEL, no observed genotoxic effect level; BPD, breakpoint dose; BPDL/U, breakpoint dose lower/upper confidence interval; STD, slope transition dose; STDL/U, slope transition dose lower/upper confidence interval; BMD<sub>1SD</sub>, benchmark dose 1 standard deviation; BMDL/U<sub>1SD</sub>, benchmark dose 1 standard deviation lower/upper confidence interval; BMD<sub>10</sub>, benchmark dose 10, BMDL/U<sub>10</sub>, benchmark dose 10 lower/upper confidence interval.

*When statistical power is reduced, the PoD is also reduced due to the increased variance in the dataset.*
was readily determined for almost all of the ENU and MNU datasets. Moreover, the BMD approach provides a number of advantages when compared to the NOGEL, BPD and STD methods. For example, the BMD methodology generally requires fewer doses in comparison to the BPD methods, BMDLs are readily defined [i.e., Tables II and III, and (Gollapudi et al., 2013)], and BMDL\(_{10}\) values, although generally lower and thus more conservative, are comparable to other PoDs for the datasets analyzed here. It is worth noting that the BMDL\(_{10}\) value, which represents a lower confidence limit of a 10% increase above the estimated background, is often below the BPD\(_{50}\), which represents a 0% change above background [Tables II and III and also in (Gollapudi et al., 2013)]. This is a consequence of these two PoD metrics being defined using different statistical models and approaches (Crump, 2011; Slob and Setzer, 2013).

An essential feature of a PoD approach that can be broadly applied with ease and success is its ability to determine whether, and at what dose, there is a detectable increase in genotoxic effect above the spontaneous background level in a particular system. It has been stated that a BPD cannot be accurately defined unless the sample size is infinite (Crump, 2011; Slob and Setzer, 2013); therefore, when examining the low dose region, a PoD metric that is based on a specified increase above a selected background (e.g., the BMD) may be the more relevant approach (Slob and Setzer, 2013). Based on the analyses conducted here, the previous work of Gollapudi et al. (2013), and the work conducted at the RIVM (Hernández et al., 2010, 2011a, b, 2012, 2013; Slob and Setzer, 2013) we support a recommendation to use the BMD approach for assessing dose responses for continuous genetic toxicology data unless otherwise justified. This approach has the added advantage of unifying analyses of genetic toxicology data with analyses of other types of toxicology data.

As part of a related exercise, we employed Monte Carlo simulation to empirically assess the effect of dataset censoring (i.e., varying number of doses and dose spacing) on the probabilistic distribution of PoD values for a given dataset. The results obtained to date, which will be published separately, indicate that BPD determinations (e.g., the L\&L BPD) are far more sensitive to dataset censoring than are the BMD determinations. For example, analyses of \textit{in vitro} \textit{HPRT} gene mutation dose–response data for ENU (Doak et al., 2007) indicated that BPD determination is optimized when the dataset contains responses for three or more doses below the PoD and three or more doses above the PoD. No such requirement was noted for the same dataset with respect to the determination of a BMD\(_{10}\). Moreover, although not typically applied, censoring of data near the point of inflection can prevent successful determination of a BPD value. \textit{A priori} assessment of the number of doses most suitable for BMD analysis indicates that three or more doses and a control are a reasonable starting point, although additional doses will typically improve the precision of the estimated PoD. The precision of the BMD can be indicated by the BMDL to BMDU ratio.

The importance of being able to define a usable PoD metric for all datasets is underscored by the subsequent comparisons of BMDL values across the endpoints investigated. For example, the results presented in Table II suggest that significant increases in \textit{Dlb1} mutations in the small intestine occur at lower ENU doses than that required to elicit significant increases in \textit{LacZ} transgene mutations (Table II). For MNU, significant increases in \textit{Hprt} gene mutations in spleen tissue occur at lower doses in comparison with that required to elicit a significant increase in \textit{LacI} transgene mutations (Table III). These lower PoDs for \textit{Dlb1} and \textit{Hprt} mutation may reflect differences in assay sensitivity, and/or gene target differences (i.e., the ability to discriminate responses in treated from control), and/or differences in repair capacity in the different tissues. However, when focusing on genotoxicity potency ranking across the compounds investigated in this and the earlier Gollapudi et al study (i.e., lowest to highest BMD\(_{10}\)) (Table V), these differences have less impact than one might expect; the order of potency \textit{in vitro} and \textit{in vivo} for both gene mutation and MN endpoints are consistent despite variations in strains and endpoints. The ranking of substances from most potent to least potent is MNU > ENU > MMS > EMS for each endpoint based on the BMD\(_{10}\) value derived using PROAST. There are very limited carcinogenicity data for these compounds; these include the following: a 54-week MNU mouse study, a 104-week MNU rat study with one dose, a 113-week ENU rat study, and a MMS mouse study with one dose. There are no carcinogenicity studies for EMS (Carcinogenic Potency Database); therefore, it is difficult to compare carcinogenicity rankings. Nevertheless, the lowest genotoxicity BMD\(_{10}\) values for ENU and MNU, 0.001 and 0.85 mg/kg/day, respectively, are within two orders of magnitude as the reported TD\(_{50}\) values for MNU (i.e., 0.0927 and 1.23 mg/kg/day for rats and mice, respectively). Moreover, the TD\(_{50}\) values reported in the Carcinogenic Potency Database (Table V) yield a ranking from most potent to least potent of MNU > ENU > MMS. This corresponds to the aforementioned ranking based on genetic toxicity BMDs, and provides additional support for the use of genetic toxicology PoDs in human health risk assessment.

Incorporating Genetic Toxicity PoD Values into Human Health Risk Assessment

The main focus of this effort was evaluation of several methods for determining genetic toxicity PoDs. However, it is also important to highlight how genetic toxicology PoD metrics can be employed in human health risk...
assessment, e.g., to support determination of regulatory limits to reduce or to quantify the risk of adverse genotoxic effects in humans. We introduce the potential role of such PoDs in risk assessment briefly, while acknowledging that comprehensive recommendations on quantitative approaches for the use of genetic toxicity data in regulatory decision making will require additional analyses and discussion. In this regard, we note that the analyses and recommendations reported herein and by Gollapudi et al. (2013) have been considered by the recent Working Group on Quantitative Approaches to Genetic Toxicology Risk Assessment that met as part of the International Working Groups on Genotoxicity Testing (IWGT) in Brazil in November 2013 (IWGT, 2013). This group is preparing two publications that endorse many of our recommendations and provide additional recommendations for the use of PoD metrics in human health risk assessment.

**Mode of Action (MOA) Data to Support Extrapolation Below the PoD**

It is necessary to have chemical-specific MOA information to justify the assumption of different slopes in the dose–response curve below the PoD as compared to above the PoD. Conversely, the dose–response analyses can also be used to support the MOA information. For example, recent work by Johnson and colleagues has shown that DNA repair capabilities below the PoDs serve to counteract specific gene mutation and chromosomal damage induced by the alkylating agents MMS, EMS, MNU and ENU (Zair et al., 2011; Johnson et al., 2012; Thomas et al., 2013). The data reported by Zair et al., (2011) support a role for the DNA repair enzyme methylpurine DNA glycosylase (MPG) in repair of clastogenic lesions below the PoD; when MPG levels were reduced by RNA interference, the EMS PoD shifted to a sevenfold lower concentration (i.e., BMDL10 decreased from 1.19 µg/mL to 0.17 µg/mL). A similar decrease in PoD was shown with the methyl-guanine methyl-transferase (MGMT) DNA repair activity and the mutagenic lesion O6-alkylG, where prior MGMT inactivation using the nucleotide analogue O6-benzyl guanine reduced the MNU PoD for HPRT mutant frequency to an approximately 50-fold lower concentration (Thomas et al., 2013). Moreover, over-expression of MGMT was shown to significantly protect against, but not completely nullify, the effect of MNU in tumor initiation (Becker et al., 2013). Such data provide support for a biological mechanism underlying a non-linear dose response in the region around the PoD. Those studies focused on EMS and MNU, respectively, for clastogenicity and mutagenicity, but the similarity in the types of DNA adducts and mutation spectra for EMS, MMS, ENU, and MNU (Beranek, 1990; Jenkins et al., 2005; Jenkins et al., 2010; Sharm et al., 2014) suggests that this group of mono-functional alkylating agents have efficient DNA repair mechanisms operating below the PoD that dramatically diminish their genotoxic effects. These cellular processes provide a mechanism for differences in the dose–response slope below and above the PoD.

**Using BMDL10 to Support Regulatory Evaluation**

PoD metrics from toxicology endpoints are frequently used to support the determination of regulatory limits that can be employed to manage the risk of adverse health

### TABLE V: Table of Lowest BMDL Values Defined Using PROAST for MNU and ENU, Along with Previously Defined PoD Metrics from Gollapudi et al. (2013)

|                      | MNU | ENU | MMS | EMS |
|----------------------|-----|-----|-----|-----|
| Gene mutation        |     |     |     |     |
| *in vitro* (µM)      | 0.006c | 0.68a | 4.72b | 8.70a |
| *in vivo* (mg/kg)    | 0.0007c | 0.09d | 1.34c | 9.29f |
| Micronucleus         |     |     |     |     |
| *in vitro* (µM)      | 0.03a | 0.17a | 1.00g | 4.35g |
| *in vivo* (mg/kg)    | 0.02b | 1.36d | 1.74f | 56.68f |
| Cancer bioassay      |     |     |     |     |
| *in vivo* (mg/kg/day)| 0.093d | 0.95d | 31.0f | Not available |

Order of potency is MNU>ENU>MMS>EMS for genetic toxicity BMDL10, and is MNU>ENU>MMS for cancer bioassay TD50. Most potent to least potent PoDs are shown from left to right.

- aDoak et al. (2007), Gene Mutation: HPRT gene, AHH-1 cells, 24 hr treatment. Micronucleus: AHH-1 cells, 18 hr treatment.
- bPottenger et al. (2009), Tk gene, L5178Y cells, 4 hr treatment.
- cBMS (unpublished data), Rat, Pig-a gene, RET and RET cells, 28 days gavage.
- dvan Delft et al. (1998), Mouse, Dbl gene, small intestine, 1 day i.p.
- eRoche (unpublished data); Rat, Pig-a gene, RBC cells, 28 days gavage.
- fGocke and Wall (2009), Gene Mutation: LacZ gene, MutaMouse, bone marrow cells, 28 days gavage. Micronucleus: bone marrow cells, 7 days gavage.
- gBryce et al. (2010), TK6 cells, 24–30 hr treatment.
- hLeBaron (2009), Rat, Blood, 4 days gavage.
- iLeBaron et al. (2010), Rat, Blood, 4 days gavage.
- jLowest TD50 from the Carcinogenic Potency Database (http://toxnet.nlm.nih.gov/cpdb/). Values adjusted for differences in treatment duration.
effects. We suggest the use of PoDs from in vivo genetic toxicology endpoints, in conjunction with, or in some circumstances in place of, PoDs estimated for other observed adverse effects. For example, the BMDL_{10} PoD values for ENU and MNU defined in this study (Tables II and III), revealed lowest in vivo BMDL_{10} values for ENU and MNU of 0.09 (male mice, Dlb-I^{+/0} gene mutation) (van Delft et al., 1998) and 0.0007 mg/kg/day (male rat Charles River Crl:CD (SD), RET Pig-a gene mutation) (BMS Pig-a RET Day 29), respectively. If the BMDL_{10} doses from the in vivo assays are converted to human equivalent doses by using the respective scaling factors of 0.081 for mouse to human, and 0.16 from rat to human (FDA, 2005), 7.3 and 0.11 μg/kg/day, or 438 and 6.6 μg/day for a 60 kg human, are obtained as the human equivalent doses associated with the aforementioned rodent BMDL_{10} values. These PoDs can be used in a similar manner to PoDs from other toxicity endpoints, e.g., to determine a regulatory limit such as a reference dose (RfD) after application of suitable uncertainty/safety factors. If, for example, a conservative safety factor of 100 (i.e., 10X for animal to human extrapolation and 10X for variability in human populations) is applied to the above PoDs, a calculable tolerable daily intake based on this endpoint would be 4.37 and 0.07 μg/person/day for ENU and MNU, respectively. If circumstances warranted, there could be a reason for the uncertainty factor related to human variability to be reduced (e.g., from 10 to 3) or even removed, based on the ability of the BMD approach to account for variability in the data as compared to, for example, the NOGEL. For example, a study with lower statistical power and greater variance would produce a lower BMD (and BPD/STD), but a higher NOGEL. If chemical-specific adjustment factors (CSAFs) are available for interspecies differences and human variability, their use also would be considered in the selection of uncertainty/safety factors (WHO/IPCS, 2005). Therefore, when considering methods for incorporation of PoD metrics into human health risk assessments, it is important to remember that some PoD metrics are more conservative than others. Moreover, the same uncertainty/safety factors would not necessarily be applicable when using the BMD compared to the NOGEL, and thus should not necessarily be applied in all instances.

A related approach that also uses the genetic toxicology PoDs estimated above involves the calculation of the increasingly used margin of exposure (MOE) metric. This approach is becoming preferable, as it incorporates estimated or actual human exposure information in the overall assessment. It is a straightforward method that involves comparison of the PoD and the current or predicted human exposure (i.e., a simple ratio of the PoD to human exposure). Regulatory decision making, and the requirement for risk management interventions, are based on the magnitude of the ratio; a larger MOE is less of a concern (e.g., MOE ≥ 10,000 may be considered to present minimal risk), while a smaller MOE may be less acceptable (e.g., MOE < 100). Other considerations that help determine the “acceptability” of the MOE approach include the severity of the effect, the MOA, the number of adverse effects observed, whether the observed effect(s) are from animal or human studies, the number of assumptions used in MOE estimations, the size of the affected population, and whether any susceptible subgroups have been identified.

Whether one uses the RfD approach or an MOE approach for genetic toxicity endpoints, the results can be evaluated with all the other available toxicity data to provide an improved and informed human health risk assessment. If determination of the PoD, and the subsequent comparative assessment indicates that genetic toxicity is the driving concern for human health considerations, then genetic toxicity data could become the basis for regulatory decision making.

CONCLUSIONS

1. MNU and ENU both elicit sub-linear dose responses that yield PoD metrics for gene mutation and chromosomal damage endpoints in vitro and in vivo.
2. Among the methods/approaches investigated here, the BMD approach yields the most conservative PoDs (i.e., BMDL_{10}).
3. The BMD_{10} is comparable to, and recommended alongside, the BMDL_{1SD} as the most suitable metrics for defining PoDs for continuous genetic toxicology data.
4. The BMD method is the preferred PoD determination method, followed by the NOGEL method, the smoothing regression spline to determine the STD, and then the segmented methods to determine the BPD (now superseded the L&L for BPD determination).
5. PoD metrics from genetic toxicology dose–response data, via the derivation of regulatory limits such as the RfD or risk management metrics such as the MOE, can be used for evaluations of human risk and regulatory decision making.

Routine determination of PoD metrics for genetic toxicity dose–response data, and routine use of genotoxicity PoD values for regulatory evaluations of new and existing substances will require application of the preferred methodology (i.e., the BMD approach) to a wider range of compounds with a diverse array of MOAs. We have already begun collecting and analyzing detailed dose–response data for other recognized genetic toxicants. Comparative analyses of the PoD values across a variety of endpoints, including carcinogenicity, as well as estimates of regulatory limits analogous to RfDs where appropriate, will enhance the foundation for the routine
interpretation of genetic toxicity dose–response data in a human health context.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Wout Slob (RIVM) and Dr. Kenny Crump for their expert input during the preparation of this manuscript. They would also like to thank the HESI Genetic Toxicology Technical Committee for its financial and intellectual support, and the International Workshop on Genetic Toxicology Working Group on Quantitative Approaches to Genetic Toxicology Risk Assessment (IWGT, 2013), who are developing two manuscripts that build upon this manuscript. Health Canada funding under the government of Canada’s Chemicals Management Plan; National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs, contract number NC/K500033/1) provided funding for Dr. Johnson. Bristol-Myers Squibb for providing us with previously unpublished data.

AUTHOR CONTRIBUTIONS

Drs. George Johnson, Bhaskar Gollapudi and Paul White provided leadership for the GTTC QAW during database construction, data collection, and data analysis, and led discussions of the group on results interpretation as well as preparation of the manuscript. Dr. Johnson is also the lead author of the paper. Drs. Greg Hixon, Chad Thompson and Liz Abraham developed and implemented algorithms for analyses of the dose–response data and determination of point of departure metrics. Drs. George Johnson and Lya Soeteman-Hernández also participated in the analyses of the dose–response data. Drs. Lya Soeteman-Hernández, Owen Bodger, Kerry Dearfield, Robert Heflich, Greg Hixon, David Lovell, Jim MacGregor, Lynn Pottinger, Chad Thompson, Liz Abraham, Veronica Thybaud, Jan van Benthem and Errol Zeiger provided insightful review of the data analyses, the data analysis methodologies and the interpretation of the results, and participated in preparation of the manuscript. Dr. Jennifer Young Tanir provided logistical, organizational and editorial support for the project. All authors contributed to addressing the comments from the internal reviewers.

REFERENCES

Becker K, Thomas AD, Kaina B. 2013. Does increase in DNA repair allow “Tolerance-to-Insult” in chemical carcinogenesis? Skin tumor experiments with MGMT-overexpressing mice. Environ Mol Mutagen 55:145–150.

Beranek DT. 1990. Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. Mutat Res 231:11–30.

Bryce SM, Avlasevich SL, Bemis JC, Phonephetpaw S, Dertinger SD. 2010. Miniaturized flow cytometric in vitro micronucleus assay represents an efficient tool for comprehensively characterizing genotoxicity dose–response relationships. Mutat Res 703:191–199.

Crump KS. 1995. Calculation of benchmark doses from continuous data. Risk Anal 15:79–89.

Crump KS. 2011. Use of threshold and mode of action in risk assessment. Crit Rev Toxicol 41:637–650.

Doak SH, Jenkins GI, Johnson GE, Quick E, Parry EM, Parry JM. 2007. Mechanistic influences on mutation induction curves after exposure to DNA-reactive carcinogens. Cancer Res 67:3904–3911.

EFSA. 2009. Use of benchmark dose approach in risk assessment: Guidance of the Scientific Committee. 2009 1150(843Q-2005-232):1–72.

El Hajjouji A, Lukamowicz M, Cammerer Z, Kirsch-Volders M. 2011. Potential thresholds for genotoxic effects by micronucleus scoring. Mutagenesis 26:199–204.

Field A. 2009. Discovering statistics using SPSS. 3rd edition. London: Sage publications.

Gocke E, Wall M. 2009. In vivo genotoxicity of EMS: Statistical assessment of the dose response curves. Toxicol Lett 190:298–302.

Gollapudi BB, Johnson GE, Hernandez LG, Pottinger LH, Dearfield KL, Jeffrey AM, Julien E, Kim JH, Lovell DP, Macgregor JT, Moore MM, van Benthem J, White PA, Zeiger E, Thybaud V. 2013. Quantitative approaches for assessing dose–response relationships in genetic toxicity studies. Environ Mol Mutagen 54:8–18.

Gollapudi BB, Thybaud V, Kim JH, Holsapple M. 2011. Strategies for the follow-up of positive results in the in vitro genotoxicity assays—an international collaborative initiative. Environ Mol Mutagen 52:174–176.

Hernández LG, Slob W, van Steeg H, van Benthem J. 2010. Comparison of carcinogenic potency estimates to in vivo genotoxic potencies from the micronucleus, transgenic rodent mutation and comet assay using the benchmark dose approach. Environ Mol Mutag 51:707.

Hernández LG, Slob W, van Steeg H, van Benthem J. 2011a. Can carcinogenic potency be predicted from in vivo genotoxicity data? A meta-analysis of historical data. Environ Mol Mutag 52:518–528.

Hernández LG, Van Benthem J, Johnson GE. 2013. A mode-of-action approach for the identification of genotoxic carcinogens. PLoS One. doi:10.1371/journal.pone.0064532:e64532.

Hernández LG, Van Benthem J, Slob W. 2012. Estimating the carcinogenic potency of chemicals from the in vivo micronucleus test: RIVM Report 340700007/2012. In: Environment NIPfHat, editor. Bilthoven: RIVM.

Hernández LG, van Steeg H, Luijten M, van Benthem J. 2011b. Mechanisms of non-genotoxic carcinogens and importance of a weight of evidence approach. Mutat Res 682:94–109.

Hixon G, Bichetler A. 2013. dsmoother: dose–response modeling with smoothing splines. http://CRAN.R-project.org/package=dsmoother.

IWGT. 2013. International Workshop of Genetic Toxicology. Working Group on Quantitative Approaches to Genetic Toxicology Risk Assessment. Foz do Iguassu, Brazil, Oct 2013, http://www.iaemgs.org/IWGT2013_Quantitative_WorkGroup_Summary.pdf. Accessed March 2014.

Jenkins GJ, Doak SH, Johnson GE, Quick E, Waters EM, Parry JM. 2005. Do dose response thresholds exist for genotoxic alkylating agents? Mutagenesis 20:389–398.

Jenkins GJ, Zair Z, Johnson GE, Doak SH. 2010. Genotoxic thresholds, DNA repair, and susceptibility in human populations. Toxicology 278:305–310.

Johnson GE, Doak SH, Griffiths SM, Quick EL, Skibinski DOF, Zair ZM, Jenkins GI. 2009. Non-linear dose–response of DNA-reactive genotoxins: Recommendations for data analysis. Mutat Res 678:95–100.

Johnson GE, Parry EM. 2008. Mechanistic investigations of low dose exposures to the genotoxic compounds bisphenol-A and rotenone. Mutat Res 651:56–63.
Johnson GE, Zair Z, Bodger OG, Lewis PD, Rees BJ, Verma JR, Thomas AD, Doak SH, Jenkins GJS. 2012. Investigating mechanisms for non-linear genotoxic responses, and analysing their effects in binary combination. Genes Environ 34:179–185.

Laws SC, Carey SA, Ferrell JM, Bodman GJ, Cooper RL. 2000. Estrogenic activity of octylphenol, nonylphenol, bisphenol a and methoxychlor in rats. Toxicol Sci 54:154–167.

Lebaron MJ, Kan HL, Geter DR, Pottenger LH, Zhang F, Schisler MR, Bartels MJ, Gollapudi BB. 2008. Non-linear low dose response in methyl methanesulfonate (MMS)-treated rats as measured by multiple endpoints. 47th Annual Meeting of the Society of Toxicology Abstract no. 1429. Toxicologist 62:293.

LeBaron MJ. 2009. Evidence of Low Dose Non-Linearity and Biological Thresholds for Multiple Endpoints in Rats Administered Alkylating Agents; Summer Toxicology Forum Aspen.

Monroe JJ, Kort KL, Miller JE, Marino DR, Skopek TR. 1998. A comparative study of in vivo mutation assays: Analysis of hprt, lacI, and cII/cI as mutational targets for N-nitroso-N-methylurea and benzo[a]pyrene in Big Blue(TM) mice. Mutat Res 421:121–136.

Muggeo V. 2008. Segmented: An R Package to Fit Regression Models with Broken-Line Relationships. R-News: http://cran.r-project.org/doc/Rnews/. p 20–25.

Piersma AH, Hernandez LG, van Benthem J, Muller JJ, van Leeuwen FX, Vermeire TG, van Raaij MT. 2011. Reproductive toxicants have a threshold of adversity. Crit Rev Toxicol 41:545–554.

Pottenger LH, Gollapudi BB. 2009. A case for a new paradigm in genetic toxicity testing. Mutat Res 678:148–151.

Pottenger LH, Schisler MR, Zhang F, Bartels MJ, Fontaine DD, McFadden LG, Bhaskar B. 2009. dose–response and operational thresholds/NOAELs for in vitro mutagenic effects from DNA-reactive mutagens, MMS and MNU. Mutat Res 678:138–147.

Sharma V, Collins LB, Clement JM, Zhang Z, Nakamura J, Swenberg JA (2014) Molecular dosimetry of endogenous and exogenous O6-methyl-dG and N7-methyl-G addacts following low dose [D3]-Methylxometasourea exposures in cultured human cells. Chem Res Toxicol. doi:10.1021/tr5000602

Slob W. 2002. PROAST: Software for dose–response modeling and benchmark dose analysis. http://www.rivm.nl/en/Library/Scientific/Models/PROAST; RIVM.

Slob W, Setzer RW. 2013. Shape and steepness of toxicological dose–response relationships of continuous endpoints. Crit Rev Toxicol. doi:10.3109/10408444.2013.853726.

Thomas AD, Jenkins GI, Kaina B, Bodger OG, Tomaszowski KH, Lewis PD, Doak SH, Johnson GE. 2011. Influence of DNA repair on non-linear dose–responses for mutation. Toxicol Sci 132:87–95.

USEPA. 2012. Benchmark Dose Technical Guidance EPA/100/R-12/001.

USEPA. 2013. BMDS v2.2.

USFDA. 2005. Guidance for industry estimating the maximum safe starting dose in initial clinical trials for therapeutics in adult healthy volunteers. U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER), Rockville.

van Delft JH, Bergmans A, van Dam FJ, Tate AD, Howard L, Winton DJ, Baan RA. 1998. Gene-mutation assays in lambda lacZ transgenic mice: comparison of lacZ with endogenous genes in spleenocytes and small intestinal epithelium. Mutat Res 415:85–96.

Wood SN. 2006. Generalized additive models: An introduction with R. 1st edition. Boca Raton, Florida: Chapman and Hall/CRC.

Wood SN. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J Roy Stat Soc 73:3–36.

Zair ZM, Jenkins GI, Doak SH, Singh R, Brown K, Johnson GE. 2011. N-methylpurine DNA glycosylase plays a pivotal role in the threshold response of ethyl methanesulfonate-induced chromosome damage. Toxicol Sci 119:346–358.

Accepted by—
M. Honma