# Toxicology Research

Accepted Manuscript






ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/toxicology

# New Approaches to Advance the use of Genetic Toxicology Analyses for Human Health Risk Assessment and Regulatory Decision-Making

**George E. Johnson**; Wout Slob, Shareen Doak, Mick Fellows, Bhaskar Gollapudi, Robert H. Heflich, Ben Rees, Lya G. Soeteman-Hernández, Jatin Verma, John Wills, Gareth Jenkins, Paul White,

**Abstract:**

Genetic toxicology testing has a crucial role in the safety assessment of new and existing substances of societal value by reducing/eliminating human exposure to potential somatic and germ cell mutagens. Genetic toxicology assays have historically been used in a qualitative manner to arrive at the binary decision of 'yes' or 'no' with regards to the mutagenic potential. However, the field is currently at a crossroads, with new methods being developed and new proposals being made to use genetic toxicity data in a more quantitative manner. Technological advances have made it possible to perform high-content, high-throughput and high-precision analysis to increase the number of "scored" events leading to increased statistical precision of the endpoint under evaluation. Automated flow cytometry and image analysis are providing significant advantages for the evaluation of gene mutations as well as cytogenetic damage both *in vitro* and *in vivo*. In addition, statistical methods such as the benchmark dose (BMD) approach can be used to identify point of departure (PoD) metrics for use in human health risk assessments, including estimation of reference dose (RfD) and margins of exposure (MOE) from *in vivo* data. Here we provide new data to compare different *in vitro* micronucleus approaches, observing that the flow based assay performs very well in defining a PoD for methyl methanesulfonate. We also present reanalysis of published *in vivo Pig-a* gene mutation data, to show how covariate analysis increases precision and reduces the effects of outliers when defining BMD values. Furthermore, we show how *in vivo* BMD metrics can be used to define RfD values, and then provide comparisons to other human exposure limit values such as permitted daily exposure (PDE). Finally, the principles of empirical correlation using BMD metrics are presented, with methods for derivation of BMD values for endpoint B, when using data from only endpoint A. These developments are opening the possibility of genetic toxicity data being used as an apical endpoint to define negligible risk in human health risk assessments. Expert groups consisting of stakeholders representing academia, industry and the government are now developing guidance on transforming genetic toxicology testing from a qualitative to a quantitative science, keeping in mind the 3R principles of animal welfare.

[†]The opinions and recommendations expressed in this publication are those of the authors, and do not necessarily reflect those of the institutions with which they may be affiliated. The information in these materials is not a formal dissemination of information by the U.S. Food and Drug Administration and does not represent agency position or policy.

## Introduction:

Recent studies have highlighted the limitations of assays currently used in genetic toxicology, as well as opportunities for improvement[1-5]. Of particular concern has been the sensitivity and specificity of the assays, which reflect their ability to correctly identify substances that experimentally induce cancer and/or germ cell mutations. In addition, genetic toxicity data have traditionally been used for hazard identification, and thus only interpreted in a qualitative dichotomous manner, i.e., a 'yes' or 'no' response with respect to a compound's potential to induce a genotoxic effect. However there is growing interest in the quantitative analysis and interpretation of genetic toxicity data, with a focus on the analysis of dose response functions and applications of the analyses in human health risk assessment [6-9]. There is also a growing appreciation that cancer is not the only apical endpoint of concern for genetic damage and adverse effects such as cardiovascular and neurological disorders, birth defects, and mitochondrial diseases are linked to mutagenicity *per se*[7]. Consequently, there is an increasing appreciation that genetic toxicity test results should be quantitatively scrutinized, and moreover, considered more widely for human health risk assessment (HHRA).

There is precedence for the quantitative use of toxicological dose-response data and of point of departure (PoD) values to define human dose equivalents that pose negligible risk of an adverse outcome. However, genetic toxicologists often argued that effects resulting from genetic damage are based upon a stochastic process, and as such, the dose response for induced effects is linear to zero dose, with no response threshold. For DNA reactive chemicals, some have argued that a single "hit" (e.g., adduct) in a single cell is adequate to increase the risk of the adverse outcome (i.e., cancer) in the exposed individual[10]. This paradigm assumes low dose linearity for induced effects with the absence of a response threshold below which responses could be termed negligible (i.e., below the limit of detection). On the other hand, there has been a growing acceptance of the existence of response thresholds for genotoxic chemicals that act via non-DNA targets, e.g., disruption of the mitotic/meiotic spindle by interfering with the polymerization of tubulin molecules[11, 12].

The debate on the existence of response thresholds for DNA reactive chemicals is a long-standing one. It is not possible to provide scientific evidence regarding the theoretical assertion that all DNA reactive molecules do not have response thresholds (i.e., are linear to zero dose). Similarly, no amount of experimentation will convince proponents of the linear-no threshold paradigm that response thresholds do indeed exist for at least some DNA reactive molecules. The challenge in both cases is the same, viz., collecting enough data to have adequate statistical precision at the low end of the dose-response curve. Expert groups such as the International Workgroup on Genotoxicity Testing (IWGT 2013) and the Genetic Toxicology Technical Committee of the Health and Environmental Science Institute of the International Life Science Institute (GTTC) are addressing the above issues and challenges with the aim of developing pragmatic approaches for establishing acceptable exposure limits for genotoxic agents (i.e., levels associated with negligible risk). The debate is gradually shifting away from the existence of thresholds towards pragmatic methods for the analysis of dose-response data, identification of PoD values, and extrapolation to environmentally relevant doses that are associated with negligible risk. Recent work coordinated by the GTTC and IWGT have demonstrated that PoDs can be determined using a variety of different statistical methods [6-8], and once defined, used for the calculation of exposure limits and margins of exposure (MOE) [9].

The aim of all toxicological screening for regulatory decision-making is to reduce the risk of harm to humans while minimizing the number of high value substances that are unnecessarily removed from the product development pipeline. In this work we argue that genetic toxicity should be regarded as a *bona fide* toxicological

endpoint, and as such, genetic toxicity screening and quantitative dose-response analyses should be designed to permit an analogously pragmatic, yet protective approach to substance screening. Moreover, novel experimental and data analysis methods should strive to minimise the use of experimental animals. It is our contention that pragmatic effective and protective genetic toxicity screening can be achieved through a combination of new techniques in genetic toxicology, along with quantitative dose-response analyses and the attendant extrapolations for HHRA.

## New Developments in Experimental Approaches:

Reliable determinations of PoD values require more data than that routinely generated during the screening of new and existing test articles. This can be a challenge since many mammalian genetic toxicity studies currently tend to include only 3 treatment groups because the generation and collection for mutation and cytogenetic damage data can be very labour intensive. The utility of high-throughput, high-content and high-precision approaches has begun to overcome some of these limitations, and several emerging methods are ideally suited to generating the data required to adequately define dose response relationships and determine PoD values. Some of the most promising methods employ flow cytometry to enumerate *in vitro* and *in vivo* cytogenetic damage and gene mutation events. Assays that employ flow cytometry methods have benefitted both from miniaturisation and high throughput automation; these permit rapid, effective scoring of large numbers of cells and concomitant dramatic improvements in the precision of the response metric (e.g., frequency of cytogenetic damage in red blood cells). The dramatic increase in the number of cells that can be readily scored, relative to older manual scoring methods, improves statistical precision and the ability to define PoD values. With respect to *in vivo* studies, improved flexibility and the ability to integrate numerous toxicological endpoints into a single study has been recognised as a significant advantage that can contribute to the 3R's, i.e., replace, reduce and refine the use of experimental animals for toxicity assessment. For example, using different dyes, flow cytometry based methods now readily allow multiplexing of endpoints such as cell viability, gene mutation, and cytogenetic damage without the need for sacrifice of animals at each time point. Automated digital image analysis is also very promising, with systems such as Metafer™ (MetaSystems™, Zeiss™), In cell analyser™ (GE Healthcare™) and others providing platforms for automated or semi-automated slide-based cytogenetic analyses. These platforms enumerate the frequency of specific cellular characteristics that are identified based on software-specific cellular image classifiers.

Flow-based approaches have now become the preferred method for the *in vivo* micronucleus (MN) assay to assess cytogenetic damage in experimental animals. There are also *in vitro* versions of the MN assay, such as the Microflow™ (Litron, USA), which differs from that used previously [12], mainly through the addition of a cell lysis step. The Microflow™ system was recently used to elegantly characterise dose response functions for MN induction for a variety of agents[13]. In addition, there is a new semi-automated method (Metafer™) recently developed for the enumeration of MN frequency in cultured cells (**Figure 1**). In **Table 1** we summarise some of the advantages that the two aforementioned *in vitro* approaches have for dose response analysis, relative to traditional manual scoring, along with some preliminary data generated using the model chromosome breaking agent methyl methanesulfonate (MMS) (**Table 1**).

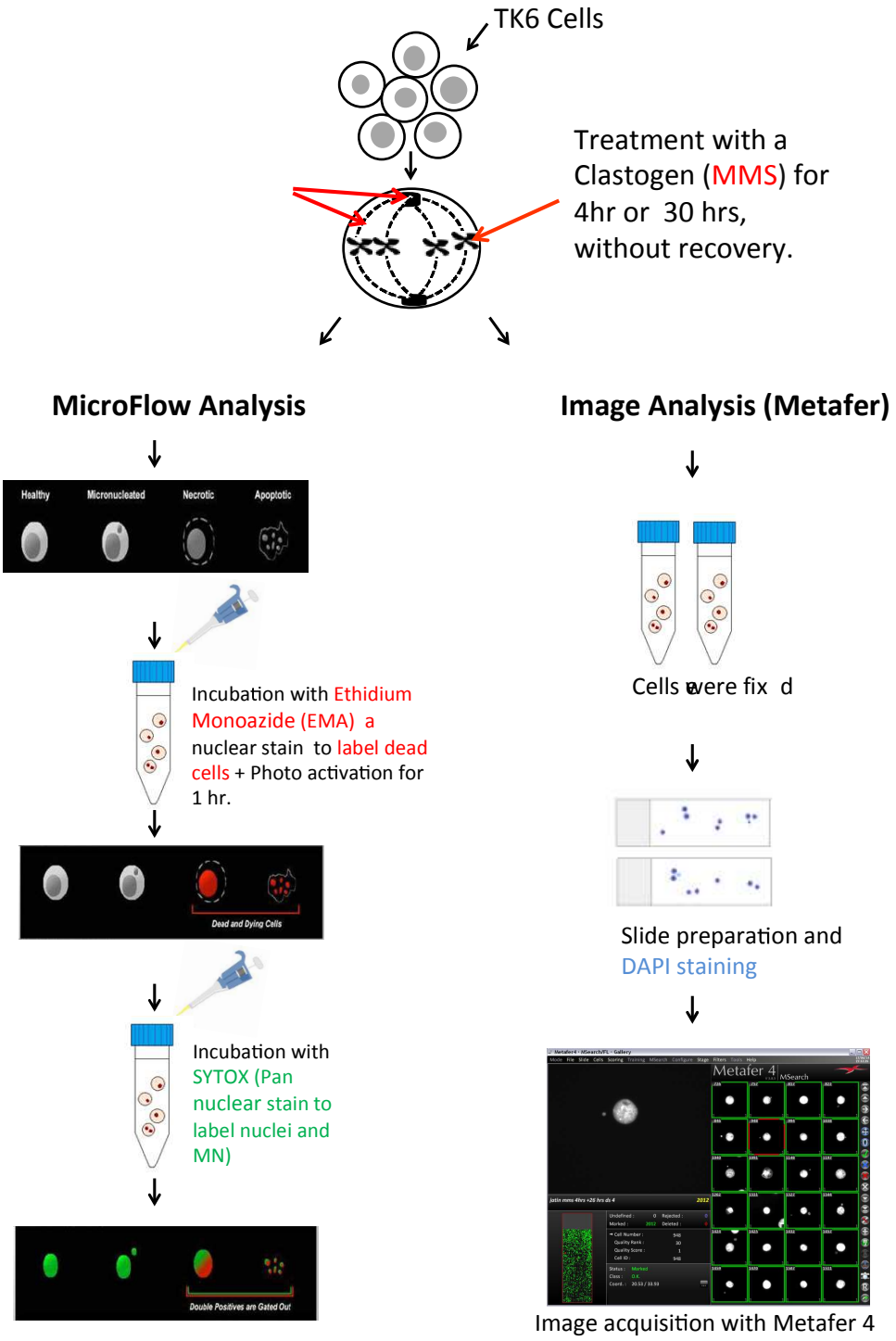


**Figure 1**: Overview of experimental protocols for two automated versions of the *in vitro* MN assay. The schematic on the right illustrates the use of the Metafer system for slide-based image analysis. The schematic on the left illustrates the use of Microflow™ kits available from Litron Laboratories (Rochester, NY, USA). Gating images from http://litronlabs.com/in_vitro_micronucleus.html).

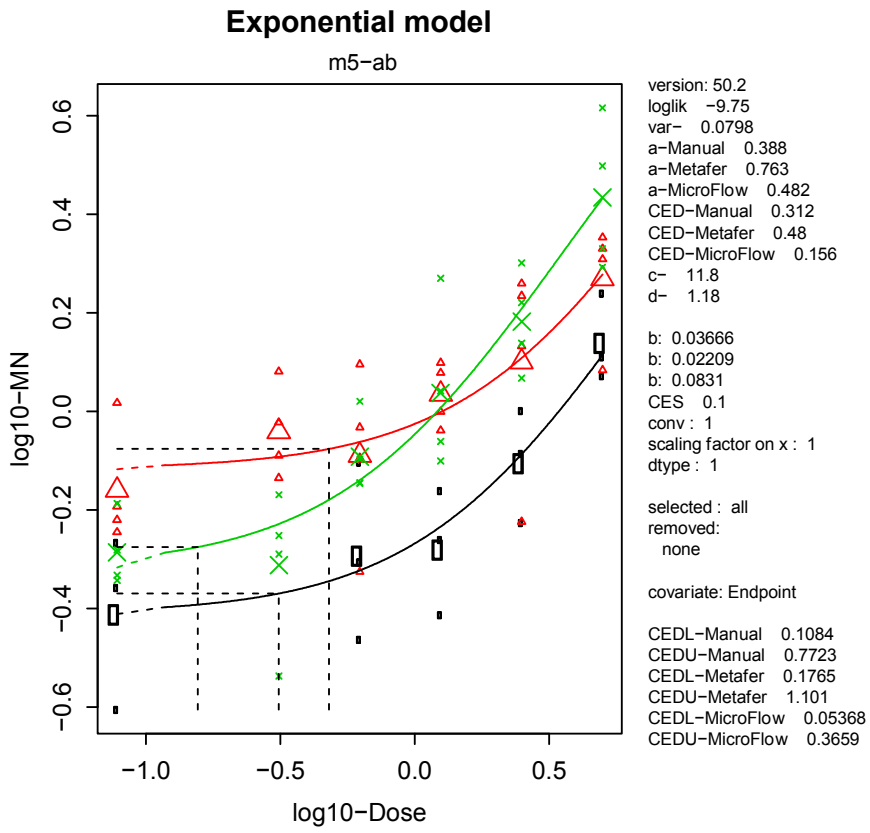


**Figure 2**: Induction of micronucleus formation in TK6 cells following treatment with MMS. 10,000 events were scored using the MicroFlow™ method following 30hours treatment (crosses + Green line), 4000 mono nucleated cells were scored with the Metafer system following 30hours treatment (triangles + Red line), and 6,000 mono nucleated cells were scored manually following 4hours treatment and 26 hours recovery (circles + Black line). PROAST version 50.3 was used for BMD (benchmark dose) analysis, with $BMD_{10}$ (CED), $BMDL_{10}$ (CEDL) and $BMDU_{10}$ metrics presented in the figure. The three dose responses were analysed together with MN scoring method as the covariate. These data are previously unpublished.

**Table 1:** Table summarising the advantages and disadvantages of different *in vitro* micronucleus protocols for dose-response analysis and point of departure (PoD) determination.

| MN scoring approach | Scoring platform | Advantages[1] | Disadvantages |
|---|---|---|---|
| Microscopy - Image analysis | Manual Scoring | • Simple, economical and adaptable<br>• Stained slides can be stored for re-analysis | • Laborious and time consuming<br>• 1,000 cells scored requires >15 minutes. |
| | Semi-automated Metafer Scoring | • High-content, high-throughput and high-precision<br>• Images of nuclei and MN can be stored for re-assessment by technician | • 4,000 cells can be scored in 10 minutes.<br>• Automated binucleate frequency scoring for proliferation index requires an additional 20 minutes per slide. |
| Flow cytometry | Fully-automated MicroFlow™ Scoring | • High-content, high-throughput and high-precision<br>• Permits endpoint multiplexing<br>• 10,000 events scored in 1-2 minutes. | • Cell lysis step removes ability for re-assessment of stored samples<br>• Over-scoring due to multi-MN cells or multi-nucleated cells with MN. |

[1]High-content - many events can be scored in a short period of time from a single sample. High-throughput - many samples can be assessed in a short period of time. High-precision - highly precise measurements can be taken. Multiplexing - different endpoints can be combined and analysed simultaneously or in sequence. MN - micronucleus.

Another high-throughput and high-precision genotoxicity assessment tool is the *in vivo Pig-a* assay, which in-directly measures mutation frequency at the *Pig-a* locus via flow cytometric determination of the loss of a GPI-anchored surface antigen. The X-linked phosphatidylinositol glycan-class A (i.e., *Pig-a*) gene has been used as a reporter for induction of somatic cell mutations using flow cytometric analysis to determine the frequency of cells that express glycosylphosphatidylinositol (GPI) anchored protein(s) on their surface [14]. Flow cytometric analysis uses fluorescent antibodies against GPI-anchored cell surface markers, such as CD59, to discriminate between wild-type (i.e., anchor-proficient) and mutant phenotype (i.e., anchor deficient) cells [15-22]. Although this assay is currently only carried out in blood to assess the mutagenic effects in haematopoietic tissue, it does provide an excellent means to defining PoDs for genotoxicants without the need to sacrifice the experimental animals. Therefore the endpoint can readily be multiplexed, and time course analyses can also be incorporated into genotoxicity assessments. The ability to apply the *Pig-a* mutant phenotype scoring principle to other tissues would constitute a major improvement to this assay; moreover, an *in vitro* version of the assay could offer high-content, high-throughput, high-precision multiplex capabilities for genetic toxicity screening in cultured mammalian cells. The latter could dramatically reduce time and cost relative to currently available assays (i.e., mutation induction at *TK* and *HPRT* loci) [23-26].

## New Developments in Data Analysis:

The aforementioned *in vitro* approaches must have high sensitivity and specificity to accurately identify *in vivo* mutagens that would be highlighted for concern and control in an HHRA context[27-29]. This has already been realised in the cosmetics industry following EU directives that prohibit animal testing [30]. A reduction in animal numbers can also be achieved through improved statistical analysis, such as benchmark dose (BMD) modelling for *in vivo* PoD assessment[7], in this case leading to less than half the number of animals being needed to accurately define a PoD metric, as compared to other methods. A refinement of animal tests is also achievable since increased recognition of genetic toxicity as *bona fide* toxicological endpoints, and the use of BMD approaches to define genotoxicity PoD values, do not necessitate treatment at the maximum tolerated doses that are most likely to elicit animal suffering. Furthermore, the BMD can reduce the number of misleading results. Slob and Setzer (2014) discuss a novel approach, in which data sets are combined. This reduces the impact of outliers and the chance of misleading results, while resulting in more statistical precision, in particular smaller confidence intervals for BMDs that are used as measures of genotoxic potency. This methodology was successfully applied to data sets from the NTP-FDA database [31].

The aforementioned BMD is the dose that elicits a pre-specified change in response (i.e., the Benchmark Response or Critical Effect Size), and it is determined by fitting dose-response models to the data and subsequently interpolating for the desired response level[7, 8]. The pre-specified change may be specified in terms of percentage change in mean response relative to the mean background, or in terms of statistical distance beyond the background response (i.e., one standard deviation) [7]. The BMD approach generally includes calculation of a confidence interval around the BMD value, with the BMDL being the lower bound of the confidence interval (CI). This value, when properly adjusted, can serve as the PoD in chemical risk assessment. BMD values based on fixed percentage increases (i.e., BMDL 10% or $BMDL_{10}$) are the most appropriate potency metrics for investigating empirical relationships across endpoints (e.g., *in vivo* clastogenicity versus carcinogenicity). Such relationships, which have recently been published and promoted by Hernandez and colleagues[9, 31], permit more effective use of current and forthcoming genetic toxicity data for HHRA and regulatory decision-making, thus permitting the realisation of the assertions outlined in the introduction of this work.

Among the most comprehensive genetic toxicity studies to date, which employed the aforementioned high-throughput and high-precision techniques along with optimised study design and data analyses, is that of Cao et al (2014). This group enumerated *Pig-a* mutant frequency (MF) in haematopoietic tissue in the same EMS-treated animals analysed for transgene mutant frequency (i.e., *gpt*) and MN frequency. The work complements an earlier study on EMS *in vivo* genotoxicity [32-35], and confirms the similarity of manually-determined MN frequency values for haematopoietic tissue and transgene mutant frequency( i.e., *lacZ*) in selected tissues; and moreover, confirms the utility of the flow-derived *Pig-a* mutant frequency values presented by Cao et al (2014).

**Table 2:** $BMDL_{10}$ metrics calculated from published *in vivo* genetic toxicology dose-responses for three potent alkylating agents [7, 36] with focus on high-throughput and high-precision approaches. The PoD metrics were used to define reference doses (RfDs) for the human population using allometric dose scaling factors of 0.081 for mouse to human, and 0.16 for rat, and an additional uncertainty factor of 100 for inter- and intraspecies differences in sensitivity [37 38]. For example, for EMS *gpt*-delta MF in spleen, (0.35x0.081x60)/100 = mg/day. BMS, Bristol Myers Squibb data presented in Johnson et al (2014); citations for other data sets from which the $BMDL_{10}$s were derived [34, 36, 39, 40]; PCE - Polychromatic Erythrocytes. RET - Reticulotytes. RBC - Red Blood Cells. ; SI - Small intestine. NoDR - No significant dose response. MF - Mutant Frequency. MN - Micronucleus.

| Substance | Species | Scoring method | Tissue | Study | $BMDL_{10}$ mg/kg | RfDs (µg/day) |
|---|---|---|---|---|---|---|
| EMS | Mouse | Manual | Most sensitive tissue, Spleen | *Gpt* MF – Cao et al (2014) | **0.35** | **17.01** |
| EMS | Mouse | Manual | Bone Marrow | *Gpt* MF – Cao et al (2014) | 0.37 | 17.98 |
| EMS | Mouse | Flow | Peripheral blood, RET | *Pig-a* MF Day 29 – Cao et al (2014) | 1.18 | 57.35 |
| EMS | Mouse | Flow | Peripheral blood | MN Day 13 – Cao et al (2014) | 6.79 | 329.99 |
| EMS | Mouse | Flow | Peripheral blood | MN Day 29 – Cao et al (2014) | 8.26 | 401.44 |
| EMS | Mouse | Manual | Bone Marrow | *LacZ* MF – Gocke & Wall (2009) | 9.29 | 451.49 |
| | | | | | | |
| ENU | Mouse | Manual | Most sensitive tissue, SI | *Dlb* MF – VanDelft (1998) | **0.09** | **4.37** |
| ENU | Mouse | Flow | Peripheral blood, RBC | *Pig-a* MF Day 29 – Bhalli (2011) | 0.12 | 5.83 |
| ENU | Mouse | Flow | Peripheral blood, PCE | MN – Bhalli (2011) | 1.36 | 66.10 |
| | | | | | | |
| MNU | Mouse | Manual | Most sensitive tissue, Spleen | *Hprt* MF – Monroe (1998) | 2.06 | 100.12 |
| MNU | Rat | Flow cytometry | Peripheral blood, RBC | *Pig-a* MF Day 4 – BMS | NoDR | NoDR |
| MNU | Rat | Flow cytometry | Peripheral blood, RBC | *Pig-a* MF Day 15 – BMS | 0.2 | 19.20 |
| MNU | Rat | Flow cytometry | Peripheral blood, RBC | *Pig-a* MF Day 29 – BMS | 0.015 | 1.44 |
| MNU | Rat | Flow cytometry | Peripheral blood, RET | *Pig-a* MF Day 4 – BMS | NoDR | NoDR |
| MNU | Rat | Flow cytometry | Peripheral blood, RET | *Pig-a* MF Day 15 – BMS | 0.1 | 9.60 |
| MNU | Rat | Flow cytometry | Peripheral blood, RET | *Pig-a* MF Day 29 – BMS | **0.0007** | **0.07** |

As illustrated in the recent work by Johnson et al (2014), suitably adjusted genetic toxicity PoD values can be used to establish exposure limits such as an RfD (Reference Dose). For example, allometric scaling can be employed to convert rodent (i.e., rat or mouse) PoD values to human equivalents, and these values can then be suitably adjusted using the appropriate uncertainty/safety factors (e.g., 10 for inter-individual human variability and 10 for animal-to-human extrapolation). **Table 2** shows PoD values for 3 potent alkylating agents (i.e., EMS, ENU and MNU), defined using non-linear dose response models and the BMD method, and the RfD values calculated from the *in vivo* genetic toxicity $BMDL_{10}$ values[7]. These metrics are in the same order of magnitude as the 104μg/day permitted daily exposure (PDE) value for EMS derived by Roche following the Viracept contamination incident [41]. We acknowledge that some jurisdictions may employ different methods and adjustment/safety factors to determine human exposure limit values (e.g., RfD, PDE, TDI, ADI, etc), particularly for different types of agents (i.e., substances in food versus therapeutic products versus industrial chemicals, etc.). Moreover, it is important to note that tissue restrictions for some of the assessment endpoints (e.g., *Pig-a* mutant frequency in haematopoietic tissue) may be problematic, for example, for ingested substances that are not systemically distributed [7, 36]. Therefore, the PoD metrics and resulting RfDs shown may in fact be higher than those that would be obtained for target tissues collected from the same animals.

Other factors such as rodent strain, route of administration and sampling time can also affect the PoD value. For example, the EMS data presented in **Table 2** shows a $BMDL_{10}$ for *gpt* MF in bone marrow of 0.37 mg/kg/day (i.e., *gpt* delta mouse model) and a $BMDL_{10}$ for *LacZ* MF in bone marrow of 9.29 mg/kg/day (i.e., MutaMouse model). In addition, the data from [7],which investigated *Pig-a* MF in rat RBCs and RETs following MNU exposures for 4, 15 or 29 days, suggest that the PoD value shifts to lower doses with increased exposure duration, presumably due to mutation fixation and cell transit and turnover in peripheral blood [7]. However, when these data were reanalysed together with time as the covariate, the statistical precision was increased and this trend was no longer observed (**Figure 3**). This type of covariate analysis has the advantage of utilising dose-response information from related experiments to provide increased statistical precision by maximising the degrees of freedom and minimizing dose-response aberrations. Moreover, the main advantage of the analyses used for generating **Figure 3** was that in Day 15 for RETs in particular, there were outliers that caused a sudden stepped increase in dose response, which was smoothed out when using the covariate approach. Using a BMD approach, a $BMDL_{10}$ metric could be defined for Day 4 along with an infinite upper bound ($BMDU_{10}$, or $CEDU_{10}$), and this is much more useful information than that of 'no effect', as resulting from pairwise testing. After all, effects cannot be excluded had higher doses been applied, and this is reflected by the BMDL: below that dose the effect will be smaller than 10%, but it remains uncertain what would happen at higher doses. Covariate analysis was also used in generating **Figure 2**, where it provided extra information showing that the Microflow™ dose response had a greater gradient and maximum fold-change compared to the other 2 approaches, which was possibly due to over-scoring of cells. Moreover, the BMD metrics were derived using the same model for each, which therefore provides a more accurate comparison between systems with less variables (parameters) being changed[42].

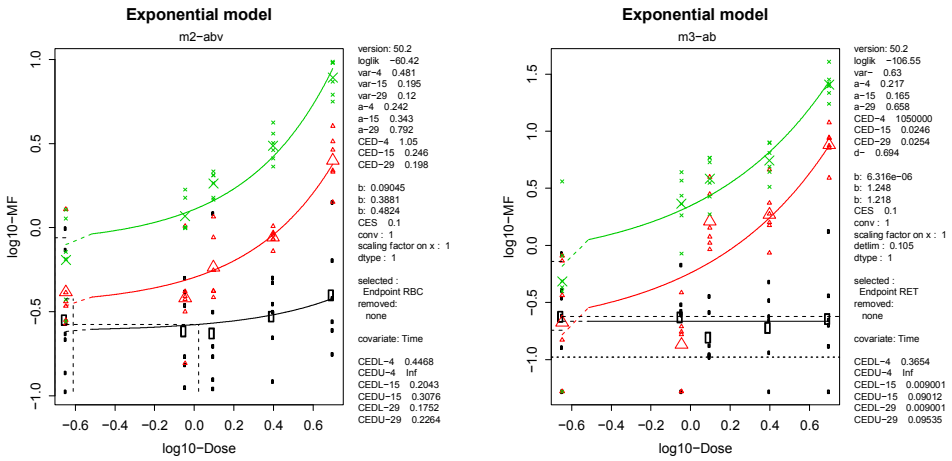


**Figure 3:** *Pig-a* mutant frequency dose responses from (a) RBC and (b) RET cells derived from rats following exposure to MNU, with sampling times at Day 4 (circles, Black line), 15 (triangles, Red line) and 29 (crosses, Green line). PROAST version 50.3 was used for the benchmark dose analysis, with $BMD_{10}$ (CED), $BMDL_{10}$ (CEDL) and $BMDU_{10}$ metrics presented in the figure. The three dose responses were analysed together with *Pig-a* scoring method as the covariate. These data were analysed previously using other PoD approaches [38].

This covariate approach was also used by researchers at the RIVM (National Institute for Public Health and the Environment, Utrecht) to examine empirical relationships between genotoxic potency and carcinogenic potency [31, 43-46] (**Figure 4**). Two proof-of–principle analyses have been performed: in the first, a positive correlation was observed between the lowest BMD from *in vivo* genotoxicity tests (MN and transgenic gene mutation) and the tissue-matched carcinogenicity BMD for 18 compounds; in the second, a positive correlation was observed between the BMDs from the *in vivo* MN and BMDs from malignant tumors in 26 compounds. The later correlation was recently used by the RIVM in a risk assessment of an impurity in animal feed, for which an *in vivo* MN test had been carried out but no cancer study was available. This case illustrates the potential of using the potency correlation between *in vivo* MN BMDs and cancer BMDs. Based on the MN test for the impurity, the $BMD_{10}$ for cancer was predicted to lie in a range of 11-2000 mg/kg bw/day, and the lower bound of this range could be considered as a $BMDL_{10}$ for cancer, which in this case may be considered as quite conservative given the large uncertainty. Even then, it was found that the MOE for the calculated exposure was very high, and, together with the fact that exposure was only expected to occur for a short period of time, it could be concluded that cancer risk was negligible. Along with this novel way of deriving a predicted cancer BMDL, and its associated MOE, from *in vivo* MN data, MOE-type metrics can also be defined directly from the genetic toxicity PoDs themselves [7-9]. This latter approach constitutes a significant departure from the "cancer-centric" approach that currently forms the basis of most HHRAs, and concomitant regulatory decisions, of genotoxic agents. As already noted, once a "no reason for concern" MOE has been defined, chemical evaluators can use genetic toxicity BMD-derived exposure limits (e.g., RfD) for HHRA. For the approach based solely genetic toxicity endpoints, there is still a need to identify and justify the appropriate tissue, endpoint, study design and PoD metric[8], and moreover, the MOE that indicates "no reason for concern". Nevertheless, the aforementioned studies and the approaches described herein provide the foundation to start defining these factors.

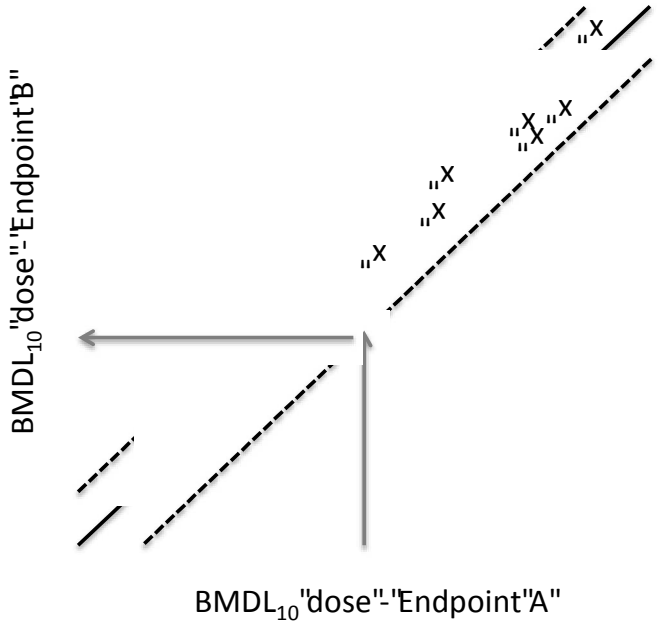


**Figure 4:** Representation of a hypothetical potency correlation between $BMDL_{10}$ metrics for Endpoint A and $BMDL_{10}$ metrics for Endpoint B across a range of different substances. The grey arrows show how a conservative $BMDL_{10}$ dose for Endpoint B can be predicted from an experimentally derived $BMDL_{10}$ from Endpoint A. The vertical distance between the dotted lines capturing the data points (x) reflect the hypothetical confidence interval as a function of the BMDL in endpoint A. Data points that would be far outside the dotted lines, i.e., in the upper left or lower right areas, may be considered as false positives or false negatives, respectively.

As part of recent collaborations involving the RIVM, Astra Zeneca and the US Food and Drug Administration (USFDA), we have also shown that BMD values from *in vitro* dose-response data correlate to *in vivo* MN BMD values in a similar manner to that illustrated in **Figure 4,** and these preliminary empirical correlations add to the growing body of *in vitro* to *in vivo* extrapolations (IVIVE) that are a cornerstone of *in vitro* only MOE values for HHRA [47 48]. In this proof-of-principle study, the applicability of using the *in vitro* micronucleus data from lymphoblastoid TK6 cells to derive cancer potency information was investigated with nineteen chemicals covering a very broad spectrum of modes of action and potencies. A clear correlation was observed between the BMDs from *in vitro* MN and BMDs for malignant tumours; as well as between BMDs from *in vitro* MN and *in vivo* MN. Although these results are very promising, some significant work remains to be completed before *in vitro* dose-response data can reliably be used for HHRA, and risk assessments based on carcinogenicity will remain an important part of regulatory decision-making. For instance, extending the number of compounds, and further investigating issues with regards to metabolic activation.

As noted in our recent work, when defining a PoD for HHRA, the BMD permits greater accuracy while using less data than other approaches[7]. In addition, deviations from the ideal data requirements have less of an impact on the BMD metric relative to PoD values determined via pairwise comparisons of control groups and treated group (i.e., the NOGEL or No Observed Genotoxic Effect Level). Therefore, we support the use of the BMD approach for defining a permitted daily exposure (PDE) for drug impurities and residual solvents (i.e., ICH M7). The preference for BMD metrics is likely to be scrutinised by governmental regulatory authorities, particularly following of the recent publication by the quantitative analyses workgroups operating under the auspices of the GTTC and IWGT

2013 [6-9]. It should also be noted that current efforts of the GTTC Quantitative Analysis Workgroup (QAW) are endeavouring to make the aforementioned statistical approaches more accessible through conference workshops, peer-reviewed articles, and user-friendly online versions of statistical tools used to define PoDs (i.e., http://www.MutAIT.org).

## Conclusions:

PoDs derived from *in vivo* genetic toxicity tests can now be used on a case-by-case basis for HHRA, with significant implications for reducing the reliance on cancer bioassays. Once quantitative analysis of *in vivo* endpoints definitely demonstrates the utility of genetic toxicity for HHRA and regulatory decision-making, the regulatory community will be poised to genuinely consider the utility of *in vitro* data. The future holds promise for high-content, high-throughput, high-precision automated genetic toxicity testing both *in vitro* and *in vivo*, and data generated using novel approaches could be used in conjunction with other molecular indicators of chemically-induced events to identify PoDs for HHRA, and moreover, to employ suitable PoDs for the determination of exposure limits (e.g., PDEs for pharmaceutical impurities)[47]. Adoption of such approaches will permit more timely and cost-effective generation of the data required for regulatory decision-making while concomitantly permitting reductions in animal usage. Although many challenges remain, the experimental and analytical approaches discussed in this work offer a multitude of advantages for regulatory authorities as well as stakeholders.

## Acknowledgements:

The authors would like to thank the HESI Genetic Toxicology Technical Committee for is intellectual support. Dr Andreas Zeller and Dr Melanie Guerard for their expert input. National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs, contract number NC/K500033/1) provided funding for Dr. Johnson. Engineering and Physical Science Research Council and GlaxoSmithKline (EPSRC-GSK case studentship EP/J502248/1) provided funding for Ben Rees. Jatin Verma would like to thank Dr. Lila A. Lodha for providing his funding. Additional funding was provided by the government of Canada under the Chemicals Management Plan. We would like to acknowledge Litron laboratories for their gifts of the Microflow kits as well as expert input from Dr. Steve Bryce and colleagues during this work.

## References:

1. K. L. Dearfield, V. Thybaud, M. C. Cimino, L. Custer, A. Czich, J. S. Harvey, S. Hester, J. H. Kim, D. Kirkland, D. D. Levy, E. Lorge, M. M. Moore, G. Ouedraogo-Arras, M. Schuler, W. Suter, K. Sweder, K. Tarlo, J. van Benthem, F. van Goethem and K. L. Witt, *Environ Mol Mutagen*, 2011, **52**, 177-204.
2. P. Fowler, R. Smith, K. Smith, L. Jeffrey, J. Young, P. Carmichael, S. Pfuler, M. Aardema, W. Diembeck, R. Fautz, J. Harvey, N. Hewitt, A. Latil, G. Ouedraogo, K. Reisinger, M. K. Fairley and D. Kirkland, EEMS, 2010.
3. D. Kirkland, M. Aardema, L. Henderson and L. Muller, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 2005, **584**, 1-256.
4. D. Kirkland, P. Kasper, L. Muller, R. Corvi and G. Speit, *Mutation research*, 2008, **653**, 99-108.

5. D. Kirkland, S. Pfuhler, D. Tweats, M. Aardema, R. Corvi, F. Darroudi, A. Elhajouji, H. Glatt, P. Hastwell, M. Hayashi, P. Kasper, S. Kirchner, A. Lynch, D. Marzin, D. Maurici, J. R. Meunier, L. Muller, G. Nohynek, J. Parry, E. Parry, V. Thybaud, R. Tice, J. van Benthem, P. Vanparys and P. White, *Mutat Res*, 2007, **628**, 31-55.
6. B. B. Gollapudi, G. E. Johnson, L. G. Hernandez, L. H. Pottenger, K. L. Dearfield, A. M. Jeffrey, E. Julien, J. H. Kim, D. P. Lovell, J. T. Macgregor, M. M. Moore, J. van Benthem, P. A. White, E. Zeiger and V. Thybaud, *Environ Mol Mutagen*, 2013, **54**, 8-18.
7. G. E. Johnson, L. G. Soeteman-Hernandez, B. B. Gollapudi, O. G. Bodger, K. L. Dearfield, R. H. Heflich, J. G. Hixon, D. P. Lovell, J. T. Macgregor, L. H. Pottenger, C. M. Thompson, L. Abraham, V. Thybaud, J. Y. Tanir, E. Zeiger, J. van Benthem and P. A. White, *Environ Mol Mutagen*, 2014.
8. J. T. MacGregor, R. Frötschl, P. A. White, K. S. Crump, D. A. Eastmond, S. Fukushima, M. Guérard, M. Hayashi, L. G. Soeteman-Hernandez, T. Kasamatsu, D. Levy, T. Morita, L. Müller, R. Schoeny, M. J. Schuler, V. Thybaud and G. E. Johnson, *Mutation Research - Genetic Toxicology*, 2014, **10.1016/j.mrgentox.2014.09.011**.
9. J. T. MacGregor, R. Frötschl, P. A. White, K. S. Crump, D. A. Eastmond, S. Fukushima, M. Guérard, M. Hayashi, L. G. Soeteman-Hernandez, G. E. Johnson, T. Kasamatsu, D. Levy, T. Morita, L. Müller, R. Schoeny, M. J. Schuler and V. Thybaud, *Mutation Research - Genetic Toxicology*, 2014, **In press**.
10. M. Kirsch-Volders, M. Aardema and A. Elhajouji, *Mutat Res*, 2000, **464**, 3-11.
11. A. Elhajouji, F. Tibaldi and M. Kirsch-Volders, *Mutagenesis*, 1997, **12**, 133-140.
12. A. Elhajouji, P. Van Hummelen and M. Kirsch-Volders, *Environ Mol Mutagen*, 1995, **26**, 292-304.
13. S. M. Bryce, S. L. Avlasevich, J. C. Bemis, S. Phonethepswath and S. D. Dertinger, *Mutat Res*, 2010, **703**, 191-199.
14. B. B. Gollapudi, A. M. Lynch, R. H. Heflich, S. D. Dertinger, V. N. Dobrovolsky, R. Froetschl, K. Horibata, M. O. Kenyon, T. Kimoto, D. P. Lovell, L. F. Stankowski Jr, P. A. White, K. L. Witt and J. Y. Tanir, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 2014, **DOI: 10.1016/j.mrgentox.2014.09.007**.
15. S. D. Dertinger, S. M. Bryce, S. Phonethepswath and S. L. Avlasevich, *Mutat Res*, 2011, **721**, 163-170.
16. S. D. Dertinger and R. H. Heflich, *Environ Mol Mutagen*, 2011, **52**, 681-684.
17. S. D. Dertinger, S. Phonethepswath, D. Franklin, P. Weller, D. K. Torous, S. M. Bryce, S. Avlasevich, J. C. Bemis, O. Hyrien, J. Palis and J. T. MacGregor, *Toxicological sciences : an official journal of the Society of Toxicology*, 2010, **115**, 401-411.
18. S. D. Dertinger, S. Phonethepswath, P. Weller, S. Avlasevich, D. K. Torous, J. A. Mereness, S. M. Bryce, J. C. Bemis, S. Bell, S. Portugal, M. Aylott and J. T. MacGregor, *Environmental and Molecular Mutagenesis*, 2011, **52**, 748-755.
19. S. D. Dertinger, S. Phonethepswath, P. Weller, J. Nicolette, J. Murray, P. Sonders, H. W. Vohr, J. Shi, L. Krsmanovic, C. Gleason, L. Custer, A. Henwood, K. Sweder, L. F. Stankowski, Jr., D. J. Roberts, A. Giddings, J.

Kenny, A. M. Lynch, C. Defrain, F. Nesslany, B. J. van der Leede, T. Van Doninck, A. Schuermans, K. Tanaka, Y. Hiwata, O. Tajima, E. Wilde, A. Elhajouji, W. C. Gunther, C. J. Thiffeault, T. J. Shutsky, R. D. Fiedler, T. Kimoto, J. A. Bhalli, R. H. Heflich and J. T. MacGregor, *Environ Mol Mutagen*, 2011, **52**, 690-698.

20. V. N. Dobrovolsky, S. Y. Boctor, N. C. Twaddle, D. R. Doerge, M. E. Bishop, M. G. Manjanatha, T. Kimoto, D. Miura, R. H. Heflich and S. A. Ferguson, *Environ Mol Mutagen*, 2010, **51**, 138-145.
21. V. N. Dobrovolsky, R. K. Elespuru, C. A. Bigger, T. W. Robison and R. H. Heflich, *Environ Mol Mutagen*, 2011, **52**, 784-794.
22. V. N. Dobrovolsky, R. H. Heflich and S. A. Ferguson, *Environ Mol Mutagen*, 2012, **53**, 440-450.
23. B. J. Rees, G. J. S. Jenkins, A. M. Lynch and G. E. Johnson, *Mutagenesis*, 2012, **27**, 789-816.
24. B. J. Rees, M. Tate, A. Lynch, C. Thornton, G. J. S. Jenkins, R. M. Walmsley and G. E. Johnson, 2014, p. In Prep.
25. B. J. Rees, G. Jenkins, A. Lynch, J. Kenny and G. E. Johnson, *Mutagenesis*, 2012, **27**, 103-137.
26. B. J. Rees, A. M. Giddings, D. Dertinger and A. M. Lynch, *Mutagenesis*, 2010, **25**, 631-658.
27. A. M. Lynch, J. C. Sasaki, R. Elespuru, D. Jacobson-Kram, V. Thybaud, M. De Boeck, M. J. Aardema, J. Aubrecht, R. D. Benz, S. D. Dertinger, G. R. Douglas, P. A. White, P. A. Escobar, A. Fornace, Jr., M. Honma, R. T. Naven, J. F. Rusling, R. H. Schiestl, R. M. Walmsley, E. Yamamura, J. van Benthem and J. H. Kim, *Environ Mol Mutagen*, 2011, **52**, 205-223.
28. J. T. MacGregor, ICEM/ EEMS Firenze-Italy, 2009.
29. V. Thybaud, M. Aardema, J. Clements, K. Dearfield, S. Galloway, M. Hayashi, D. Jacobson-Kram, D. Kirkland, J. T. MacGregor, D. Marzin, W. Ohyama, M. Schuler, H. Suzuki and E. Zeiger, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 2007, **627**, 41-58.
30. EU, in *Press release*, ed. europa.eu, http://europa.eu/rapid/press-release_IP-13-210_en.htm, 2013, vol. 2014, p. Press release outlining the EU ban.
31. L. G. Hernández, J. Van Benthem and W. Slob, *RIVM Report 340700007/2012*, 2012.
32. E. Gocke, M. Ballantyne, J. Whitwell and L. Muller, *Toxicology letters*, 2009, **190**, 286-297.
33. E. Gocke and L. Muller, *Mutat Res*, 2009, **678**, 101-107.
34. E. Gocke and M. Wall, *Toxicology letters*, 2009, **190**, 298-302.
35. L. Müller, E. Gocke, T. Lavé and T. Pfister, *Toxicol. Lett.*, 2009, **doi:10.1016/j.toxlet.2009.04.003**
36. X. Cao, R. A. Mittelstaedt, M. G. Pearce, B. C. Allen, L. G. Soeteman-Hernández, G. E. Johnson, C. A. H. Bigger and R. H. Heflich, *Environmental and Molecular Mutagenesis*, 2014, **55**, 385-399.
37. FDA, *U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER)*, 2005.
38. G. E. Johnson, L. G. Soeteman-Hernandez, B. B. Gollapudi, O. G. Bodger, K. L. Dearfield, R. H. Heflich, J. G. Hixon, D. P. Lovell, J. T. MacGregor, L. H.

Pottenger, C. M. Thompson, L. Abraham, V. Thybaud, J. Y. Tanir, E. Zeiger, J. van Benthem and P. A. White, *Environ Mol Mutagen*, 2014, **55**, 609-623.
39. J. J. Monroe, K. L. Kort, J. E. Miller, D. R. Marino and T. R. Skopek, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 1998, **421**, 121-136.
40. J. A. Bhalli, J. G. Shaddock, M. G. Pearce, V. N. Dobrovolsky, X. Cao, R. H. Heflich and H.-W. Vohr, *Environmental and Molecular Mutagenesis*, 2011, **52**, 731-737.
41. L. Muller and E. Gocke, *Toxicology letters*, 2009, **190**, 330-332.
42. EFSA, *2009*, 2009, **1150**, 1-72.
43. L. G. Hernández, G. E. Johnson, L. H. Pottenger and J. van Benthem, *Environmental and Molecular Mutagenesis*, 2011, **52**, S26-S26.
44. L. G. Hernández, W. Slob, H. van Steeg and J. van Benthem, *Environmental and Molecular Mutagenesis*, 2011, **52**, 518-528.
45. L. G. Hernández, W. Slob, H. van Steeg and J. van Benthem, *Environmental and Molecular Mutagenesis*, 2010, **51**, 707-707.
46. L. G. Hernández, J. Van Benthem and G. E. Johnson, *PLoS One*, 2013, **DOI: 10.1371/journal.pone.0064532** , e64532.
47. R. S. Thomas, M. A. Philbert, S. S. Auerbach, B. A. Wetmore, M. J. Devito, I. Cote, J. C. Rowlands, M. P. Whelan, S. M. Hays, M. E. Andersen, M. E. Meek, L. W. Reiter, J. C. Lambert, H. J. Clewell, 3rd, M. L. Stephens, Q. J. Zhao, S. C. Wesselkamper, L. Flowers, E. W. Carney, T. P. Pastoor, D. D. Petersen, C. L. Yauk and A. Nong, *Toxicological sciences : an official journal of the Society of Toxicology*, 2013, **136**, 4-18.
48. L. Soeteman-Hernandez, M. Fellows, W. Slob and G. E. Johnson, *Toxicological Sciences*, 2014, **In press**.