THIS IS A DRAFT PAPER FOR DISCUSSION. IT SHOULD NOT BE QUOTED, CITED OR REPRODUCED

MUT/2016/07

COMMITTEE ON MUTAGENICITY OF CHEMICALS IN FOOD, CONSUMER PRODUCTS AND THE ENVIRONMENT (COM)

Quantitative approaches to the assessment of genotoxicity data

The COM first considered quantitative approaches to assessing genotoxicity data and how they may be used in chemical risk assessment at its Horizon Scanning exercise in June 2013. In 2014, this topic was discussed in a workshop organised by the ILSI HESI Genetic Toxicology Technical Committee (GTTC), in collaboration with the European EMS/UKEMS which lead to a Special Issue of Mutagenesis.

The COM is now in a position to discuss this topic further. An introductory scoping paper has been produced (MUT/2016/07) in order to bring together the current strategies in risk assessment and risk management and the debate on the use of genotoxicity data in quantitative risk assessments, including a discussion of thresholds in genotoxicity endpoints.

The scoping paper lists some key questions which the committee is asked to consider.

To facilitate the discussion Professor George Johnson from Swansea University Medical School, will present some of the work from the ILSI-HESI GTTC and IWGT, as well as touching on the work he is doing with the ICH.

COM Secretariat

October 2016

MUT/2016/07

Quantitative approaches to the assessment of genotoxicity data

Introduction:

1. The COM first considered quantitative approaches to assessing genotoxicity data and how they may be used in chemical risk assessment at its Horizon Scanning exercise in June 2013. Members were aware of work being conducted by ILSI HESI on quantifying genotoxic responses and assessing non-linear dose-response relationships and that the implications of this work may need to be considered. It was also noted that many chemicals present in the environment have not been tested for carcinogenicity; therefore the possibility of developing quantitative (or semi-quantitative) methods for the analysis of dose-response data from *in vivo* genotoxicity studies for use in a margin of exposure (MoE) approach, similar to that utilised in the interpretation of carcinogenicity data, was discussed.

2. A workshop was organised by the ILSI HESI Genetic Toxicology Technical Committee (GTTC), in collaboration with the European EMS/UKEMS, at their meeting in Lancaster in July 2014. The committees' goals, the meetings' proceedings and many of the presentations were recently published in a Special Issue of Mutagenesis (Vol 31 issue 3; White and Johnson 2016). Additionally, the International Workshop of Genotoxicity Testing (IWGT) working group on Quantitative Genetic Toxicology Risk Assessment (the QWG) recently published the outcome of their discussions and consensus views in two publications, which will also be pivotal to the discussion of the feasibility of a quantitative model in genetic toxicology (MacGregor et al 2015 a,b). These publications provide a basis for the COM to now consider this topic.

With regards to the evaluation of threshold dose responses in genetic toxicology, the COM published a Guidance Statement 'Thresholds for in vivo mutagens' in April 2010 (https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/315698/assessment_of_threshold_for_in_vivo_mutagens.pdf)

3. This discussed a number of chemicals that had previously been shown to exhibit in vivo thresholds for their genotoxicity, as examples (aneugens, acting by tubulin inhibition, topisiomerase inhibitors, rapid detoxification eg phenol). The Committee also considered evidence that there are thresholds for the genotoxicity of some alkylating agents as a consequence of the induction of O^6-methylguanine-DNA-methyltransferase (MGMT).

4. This paper is presented as a scoping paper introducing quantitative analyses of genotoxicity responses. Including some recent data on non-linear dose responses, the existence of thresholds for some genotoxins and recent developments and proposals for using genotoxicity data to quantify risk. A general overview of the

present risk assessment strategies used for evaluating genotoxic chemicals and carcinogens is presented to enable comparisons to be drawn with the proposed quantitative assessments of genotoxicity data.

Current risk assessment and risk management strategies

5. Genotoxicity testing strategies were developed to identify a genotoxic or mutagenic hazard and intended to be used only in a qualitative (yes/no) manner and the current risk assessment and management strategies for chemicals, which have been shown to be genotoxic, are well established. The COC approaches to risk characterisation are described in a Guidance Statement COC/G-06 (2012 – https://www.gov.uk/government/publications/carcinogenic-dose-response-defining-a-point-of-departure-and-potency-estimates). These are broadly in accordance with those proposed by EFSA (2005).

6. Essentially, it is assumed that, other than for a few exceptions, there is no exposure level that is without risk. Hence, the ALARA (as-low-as-reasonably-achievable) or ALARP (as-low-as-reasonably- practicable) approach is advised, which suggests that levels of the chemical must be controlled to ensure that intake is as is low as reasonably, or technically, possible (Barlow et al 2006). This is a widely adopted principle used by regulatory authorities in Europe and many other regions. It is widely applied, and is particularly relevant when addressing scenarios such as impurities in products that may enter the food chain (e.g. as pesticide or veterinary medicine residues), environmental contaminants in water or air. It may also be necessary when information on the genotoxicity of a chemical (e.g. a process impurity or contaminant) exists in the absence of a carcinogenicity (and reproductive toxicity) study. It is a purely qualitative risk management approach, and there is no consideration of the genotoxicity or carcinogenicity data in a quantitative manner. As it is based only on the identification of a genotoxic hazard (i.e. a yes/no answer) and as there is no concomitant assessment of potency it will be overly conservative for some chemicals.

7. One of the difficulties of the ALARA approach is that it provides no indication of the level of concern and hence cannot be used to prioritise the use of resources. For this reason, the margin of exposure (MOE) approach was developed. Thus is generally applied to chemicals that have been shown to be genotoxic and carcinogenic. It takes into account carcinogenic potency and estimated exposure (EFSA 2005; Barlow et al 2006). The MOE is calculated using a suitable point of departure (POD) derived from reliable rodent bioassay data or human epidemiology information, divided by the anticipated/estimated exposure. The resulting value, which is unit less, has been classified by COC as follows (based on MOEs calculated using animal carcinogenicity data):

- may be a concern (<10,000);
- unlikely to be a concern (10,000-1,000,000) or
- highly unlikely to be a concern (>1,000,000)

8. This approach can be used for risk communication for scenarios where there are unavoidable exposures to carcinogens, for risk mitigation prioritisation and to set risk management actions.

9. It should be noted that the MOE approach requires no explicit assumptions about the nature of the dose-response relationship. Whilst this is not a major limitation when ranking chemicals for risk prioritisation, it does cause difficulties when concluding on levels of concern. However, the fact that an MOE of 10,000 is equivalent to a risk level of 1 in 10^5, when one assumes a linear dose-response relationship, provides some confidence in the MOE bands utilised by the COC and others.

10. The methods used to define a POD currently endorsed by COC are detailed in a Guidance Statement COC/G05 - 'Defining a Point of Departure and Potency Estimates in Carcinogenic Dose Responses'. The Benchmark dose (BMD) and T25 approaches are outlined and compared. Briefly, the T25 value is the dose that induces tumours in 25% of animals above control, determined by linear extrapolation. The BMD approach utilises mathematical modelling to calculate the lower 95% confidence interval of a dose causing a defined response, typically a 10% increase in tumours in a cancer bioassay, i.e. the BMDL10. Dybing et al (2008) provide some examples of MOE's calculated from food borne carcinogens using the two approaches. The chemicals examined include: acrylamide, aflatoxin B1, benzo(a)pyrene, and PhIP. The COC consider the BMD approach to be superior to the T25 for defining a POD.

11. There are deliberations surrounding the nature and shape of the dose response curve for chemical carcinogenesis and the assumptions implicit when using the MOE approach, particularly at low, environmentally relevant exposures (Boobis et al 2013). It is considered that the chemical mode of carcinogenic action (MOA), and species differences considerations such as in pharmacokinetics, contribute significant uncertainties to the assumption of linearity of response and therefore the provision of pragmatic advice on relative risk.

12. The threshold of toxicological concern (TTC) is a de minimus approach developed to facilitate the risk management of substances, primarily contaminants in food, for which good (or at least conservative) exposure estimates are possible but when chemical-specific toxicity data, including genotoxicity data, are insufficient for normal risk characterisation (Kroes et al 2004). The history and development of the TTC approach are comprehensively described by Dewhurst and Renwick (2013) Analysis of a dataset of carcinogens, which included those in the Carcinogen Potency Database (CPDB) at the time of the study (a total of 975 chemicals), revealed that a daily lifetime exposure of 0.15 µg/day (ie equivalent to 0.025 µg/kg bw/day) would be sufficiently protective, at a risk level of 1 in 10^6, assuming linearity in the dose-response relationship, even for chemicals with structural alerts for genotoxicity, other than for three high potency groups of chemicals. These were aflatoxin-like, azoxy and N-nitroso compounds, the so-called cohort of concern. Chemicals in these structural classes are excluded from the TTC approach (Kroes et al 2004). This approach was strengthened and supported following a recent workshop co-organised by EFSA and WHO (EFSA 2016).

13. The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) have proposed that where exposure to potentially genotoxic impurities is less than lifetime, higher TTC values can be used. Based on this, Felter et al (2009) have proposed the TTC approach for chemicals in general be refined - that for those with structural alerts, but for which there is less than 1 year's exposure, exposure could also be set at 1.5ug/day.

Non-linear dose responses and thresholds for *in vivo* mutagens

14. In risk assessment, the default assumption for chemicals that are mutagenic *in vivo* is that there is a linear dose response and no response threshold. In such cases, regulatory frameworks generally require the application of the ALARA or ALARP principle when exposure is unavoidable. However there are examples of chemicals that are genotoxic in a standard battery of tests but for which a threshold-mechanism has been defined. These are generally determined on a case-by-case basis. The COM has previously considered these and this is detailed in a Guidance Statement – Thresholds for *in vivo* mutagens (https://www.gov.uk/government/publications/assessment-of-thresholds-for-in-vitro-mutagens). Threshold terms have been defined (ie true threshold, threshold dose, practical threshold, biologically meaningful threshold, threshold mode of action) and the chemicals considered by COM to have MOA's for which a threshold mechanism has been demonstrated are described. These include aneugens acting by tubulin inhibition (eg benzimidazole group), topoisomerase inhibitors, and those that are rapidly detoxified following oral exposure (eg phenol).

15. The possibility that there is a threshold in mutagenic response due to DNA repair of adducts formed by some low molecular weight alkylating agents was also considered (eg ENU, MNU, EMS, MMS). The disparity between the frequency of DNA adducts and of mutations suggests that there is factor involved in the conversion of adduct to mutation which may exhibit a threshold (Jenkins et al 2005; Doak et al 2007). Therefore it is possible that an organism has the ability to be subjected to a low level of DNA damage which is repaired, and it is only when repair mechanisms are exhausted or overwhelmed that a mutation occurs.

16. A detailed evaluation and human risk assessment of a chemical known to be a genotoxic carcinogen was undertaken following the discovery of EMS as an impurity in tablets of Viracept (nelfinavir mesilate), an HIV protease inhibitor, in 2007 (Walker et al 2009; Muller et al 2009). It was present at up to 1068 ppm in contaminated drug batches – the estimation of the maximal daily dose of EMS to patients receiving Viracept at that time was 0.045 mg/kg bw/day (daily drug dosage of 2.92 g/day). Following discussions with the EMEA, the company (Roche) went on to perform a comprehensive quantitative risk assessment of EMS (EMEA 2007) and a threshold and 'safe level' were determined.

17. The COM considered the evidence and agreed that a threshold had been demonstrated for EMS and estimated the margin of exposure between the NOAEL

for mutagenicity and the maximal estimated intake (COM)[1]. More recently Jenkins et al (2010) examined the role of DNA repair enzyme polymorphisms and highlighted the impact that this may have on susceptibility to genotoxic damage in exposed humans.

18. Since the Viracept incident, the use of qualitative evaluation of data in assessing the risk of impurities in human pharmaceuticals has been established in ICH Guideline Human pharmaceuticals M7 - Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk provides a framework for classifying and resultant controls for different mutagenic or carcinogenic substances. The framework includes controlling known mutagenic carcinogens at or below compound-specific acceptable limits or application of the TTC for impurities that are known mutagens with unknown carcinogenic potential (e.g. bacterial mutagenicity positive, no rodent carcinogenicity data).

Quantitative risk assessment of genotoxicity data

19. Genotoxicity testing strategies were developed to identify a genotoxic or mutagenic hazard and intended to be used only in a qualitative (yes/no) manner. Most assays are optimised to enable the detection of genotoxins with enhanced sensitivity (e.g Ames tester strains). The risk assessment process determines potential public exposure to genotoxic [and therefore potentially carcinogenic] agents. However, it is considered by some that this is an overly conservative approach and in the process, potentially valuable chemicals are being screened out and discarded, or that expensive strategies to remove agents from the environment or food have to be undertaken (Pottenger and Gollapudi 2010).

20. A move towards providing quantitative assessment will require entirely new approaches to evaluating the data. This may take two forms –

- Establishing biological thresholds for a response, determining a ‘hockey stick’ shaped curve and deriving NOAELs and POD’s as for other toxicity endpoints to determine acceptable exposure levels – this has already been achieved for some genotoxic carcinogens (see above)
- Establishing the relationship between the measured endpoint (mutation, chromosome aberration, DNA strand breaks) and carcinogenicity (or other relevant endpoint), derive a [biologically relevant] POD and apply appropriate risk management approaches such as MOE, to determine levels likely to be of low concern.

Alkylating agents and thresholds

21. There are a number of publications that have examined comprehensive datasets for MMS, MNU, ENU and MNU to examine the hypothesis that DNA reactive, alkylating agents do not necessarily have linear dose responses and that thresholds can be determined.

[1] Estimated safety factor, based on the threshold dose level in mice and the estimated maximal daily dose of EMS in patients, was calculated as 454 (i.e. 25 mg/kg/day / 0.055 mg/kg/day).

22. A report from the IWGT Quantitative Analysis Working group (QAW, - a workshop organised by ILSI/HESI, comprising scientists from government, academia and industry) describes a series of analyses using dose response data to derive POD's using the no observed genotoxicity level (NOGEL), the threshold effect level (TdL) and the BMD approach (Gollapudi et al 2013). The focus of the study was EMS and MMS - both *in vitro* and *in vivo* data were analysed. *In vitro* NOGELs were of a similar order of magnitude for gene mutation and MN endpoints for both chemicals. The TdL values were similar to the NOGEL's for all data sets with the exception of *in vivo* MN for MMS which was 3-fold higher. The BMD and TdL approaches are recommended as the estimates have a quantified level of uncertainty. It was also concluded that these analyses provided support for the use of the lower confidence limit for a 10% response.

23. A similar study compared POD estimates from BMD analyses using PROAST and the EPA BMDS software, NOGEL's and breakpoint dose-response modelling and smoothing software. A variety of in vitro and in vivo genotoxicity endpoints from studies investigating MNU and ENU were modelled (Johnson et al 2014). BMDL10 values were generally lower than the other POD estimates.

24. Zeller et al (2016) investigated detailed dose response data for a number of genotoxic effects of MMS, including mutations in the pig-A assay, in rats. Rats received oral doses of 0, 1.25, 2.5, 5, 10, 15 or 30mg/kg/day for 28 days followed by a treatment free period of 33 days (days 29-62). Blood was taken for analysis in the pig-A assay at day -1, day 15, 28, 55 and 65 and for Comet on day 22. The same rats also received doses 48, 24 and 3 hours prior to necropsy (ie on days 63-65) when liver was taken for Comet assay and bone marrow for analysis of micronucleus induction. □-H2AX staining was undertaken in liver tissue as an indicator of DNA double strand breaks. POD's were derived using the DRSMOOTH package (which contains a bilinear model) for each endpoint. LOGEL's and NOGEL's were determined . A response quotient (95% tolerance interval of assay historical control range) was converted to ratio-based CES (critical effect size, analogous to a benchmark response). Dose responses were constructed at each time point. The results showed clearly that MMS acts primarily as a clastogen and therefore its potency as a gene mutagen is lower (PODs calculated from ratio-based CES for chromosomal damage were in the range 0.6–3.3 mg/kg/day, those for the Pig-a assay ranged from 2.8-8.1mg/kg/day). The authors suggest that ratio-based CES method, which is based on an endpoint-specific choice of an appropriate CES or BMR, provides a pragmatic approach for obtaining PODs from genotoxicity datasets.

25. Also from the recent Mutagenesis issue, Ji et al (2106) describe detailed analyses of dose-responses for a variety of endpoints for MNU and MMS. Male F344 rats were administered 4 daily oral doses of MMS at 0, 0.5, 1, 5, 25 or 50 mg/kg/day or MNU at 0, 0.01, 0.1, 1, 5, 10, 25 or 50 mg/kg/day. Twenty four hours after the final dose the following endpoints were examined: Hb adducts (MMS only), DNA adducts (N7MeG and O6MeG) in blood and liver; MN in peripheral reticulocytes and gene expression, by microarray analysis (Agilent 4x44K whole rat genome oligo), in liver. A linear dose response was apparent only for Hb adducts, an exposure biomarker. . Following MMS exposure, DNA adducts (N7MeG) were not increased above control level at 0.5 and 1 mg/kg/day in blood and at 0.5 mg/kg/day in liver, but increased with dose at higher doses. No increase in O6MeG adducts

was detected at any dose level. MN frequency was statistically significantly increased from control, but only at 25 and 50 mg/kg/day. Following MNU exposure there were no increases in N7MeG and O6MeG adducts at 0.01, 0.1 or 1 mg/kg/day but dose-related increases were apparent at 5-50 mg/kg/day. A similar pattern of response was seen in the MN assay.

26. The authors examined gene expression with a view to gaining some understanding of MOA. Of note was a dose-related increase in the number of genes altered by MMS and MNU - similar patterns of gene expression changes were apparent for MNU and MMS though *Mgmt* expression (=O^6-methylguanine DNA methyltransferase) was elevated only following MNU and not with MMS. Furthermore, the elevation appeared to occur only at doses greater than 5 mg/kg/day when effects were apparent at lower doses in the genotoxicity assaysThe authors conclude that overall the data provide evidence for a MOA that has a threshold, bilinear response. Accordingly a PoD can be calculated for MNU and MMS in a similar way to that derived for EMS using breakpoint dose-response type modelling.

Comparing genotoxic and carcinogenic potencies

27. A preliminary comparison of carcinogenic and *in vivo* genotoxic potencies was undertaken by Sanner and Dybing (2005). They proposed a framework that could be used in regulatory settings when a chemical is considered to be mutagenic but for which carcinogenicity studies are not available or of poor quality. The investigation identified the lowest effective dose (LED) in an *in vivo* genotoxicity study (dose giving a positive response) and the T25 from a carcinogenicity study (identified from IARC). Forty four chemicals with long term studies suitable to calculate a T25 and an appropriate *in vivo* genotoxicity study were identified. Subsequently 2 were excluded from further consideration as their carcinogenic MOA did not involve genotoxicity and either was unique to the compound (TCDD) or was not relevant to humans (atrazine). The authors considered the IARC evaluations of a further 8 chemicals did not elicit their carcinogenicity by a genotoxic mode of action (acetamide, o-anisidine, chloroform, diazepam, DEHP, 1,4-dioxane, doxylamine succinate, methyluracil). Results from a range of *in vivo* genotoxicity tests were utilised (MN, DNA breaks, CA, Comet) and studies using rats and mice were considered together (Figure 1).

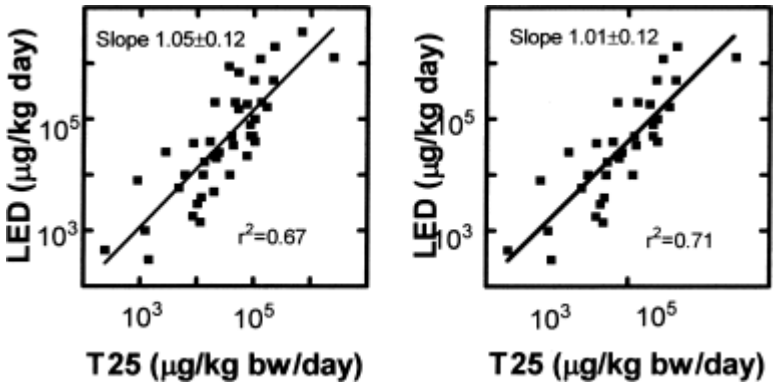

Figure 1 Logarithmic plots of LED vs T25 for all compounds (left) and for the 34 compounds that were genotoxic (right) Reference Sanner and Dybing (2005)

29. It was noted that slope and correlation coefficient was lower when MN was the endpoint used to determine the LED (0.63 ± 0.21 compared to 1.02 ± 0.17 for DNA breaks). Differences were noted between species (mice n=19, slope = 0.79 ± 0.17; compared to rats n=16 slope = 1.05 ± 0.14). The authors concluded that whilst the carcinogens evaluated came from diverse chemical classes with different MOA's , there was a correlation between genotoxic and carcinogenic potency and it was suggested, that on the basis of these findings, the LED could be used to derive a tolerable risk level for compounds without carcinogenicity data.

30. Hernandez et al (2011 –*ANNEX 1*) published an evaluation based on the principle that 'if genotoxic potency was found to correlate with carcinogenic potency, genotoxicity tests might be useful in more effectively deciding which compounds need a carcinogenicity study, and which ones could be waived'. The vision is that this approach will eventually alter the regulatory framework, thus making the use of animals in testing more effective. Their study used the BMD approach for dose response analysis and studies selected for comparison were standardized as far as possible – analysis of 24 h and 48 h sampling times for MN from bone marrow and peripheral blood respectively, but it is noted that different exposure routes were used. Their criteria for assessing the comet or transgenic mutation assays are not given. Chemicals selected were either IARC class 1 or 2A. BMD_{10} was calculated for a total of 18 chemicals (ie dose to induce a 10% increase in tumour incidence) using tumours in the most sensitive tissues – the PROAST model was used to determine BMD10s.

31. Dose-response analysis scheme. For the comet assay 27 BMD_{10} values were derived for 4 compounds from a range of tissues, for the MN assay, 54 BMD_{10} values for a range of tissues were derived for 15 compounds, for the TG 44 BMD_{10} values for a range of tissues were derived for 15 compounds, and 224 tumour BMD_{10} values for a range of tissues were derived for 18 compounds. The results were depicted in Figure 2 but further details of the individual compound dose responses were not illustrated.

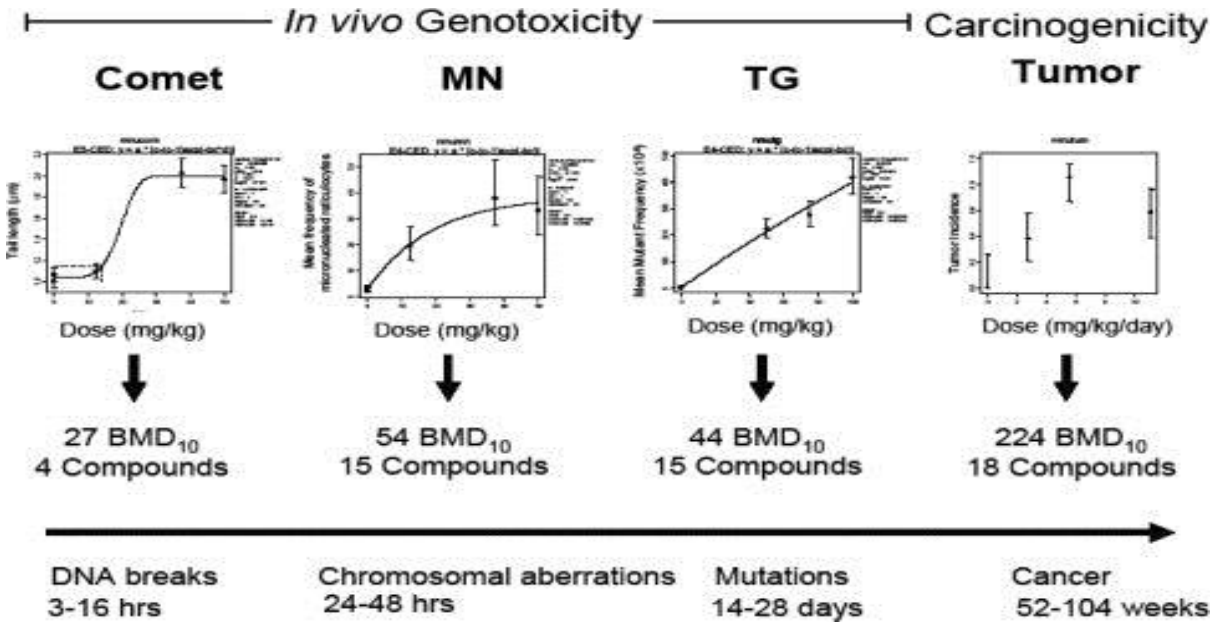

Figure 2 - Quantitative relationships were investigated by comparing BMD10 values derived from short-term *in vivo* genotoxicity tests (comet, MN, and TG) to BMD10 values derived from carcinogenicity studies (Reference Hernandez et al 2001)

32. It was pointed out that in 10/18 carcinogenicity studies only two test dose levels were used and that the genotoxicity studies rarely used more than two dose levels. It was also stated that no further analyses were undertaken for comet because only 4 studies were available. The authors suggest that the use of the BMD, using BMDU and BMDL confidence limits, takes some account of uncertainty and therefore is the better approach. Overall this study provides a preliminary evaluation of the use of BMD_{10} as the POD for comparison, although it appears that limitations in the available data do not permit definitive conclusions.

33. In the recent Mutagenesis Issue this study was extended using similar methodologies (Soeteman-Hernandez et al 2016 – ANNEX 1). A total of 48 chemicals were evaluated. It is noted that the chemicals from the earlier evaluation were included, other than PhIP (it is not clear why this was omitted). This study calculated BMD_{05} from MN studies and the log 10 of this figure was plotted against log10 of cancer $BMD_{10.}$ BMD_{05} were calculated for individual datasets (ie if there were several studies available, the data were not pooled) and PROAST was used for dose response modelling for both sets of data.

34. The plotted data exhibited a wide scatter but it was concluded that there was a positive correlation between MN and carcinogenic potencies. The authors comment that there are large deviations and speculate that these could be due to the individual chemicals' characteristics and MOA. It is stated that the carcinogenicity study designs also varied considerably (eg group size ranged from n=7-999 (in the mega-mouse study with 2-AAF) and number of dose groups from 3 to 8).

35. It is not clear what criteria were used to select the datasets for the calculation or why BMD_{10} was used previously, rather than BMD_{05}, as here. Different carcinogenic endpoints reflecting different stages of the carcinogenic process were

examined and a 'lesion category' was assigned – e.g. pre-neoplastic foci, adenomas, carcinomas – it was suggested that this could have contributed to the large scatter.

36. This is currently the most up to date comparison of *in vivo* genotoxicity and carcinogenicity potencies. The authors suggest that the genotoxicity data might provide a prediction of tumour potency with an uncertainty range spanning a factor of 100.

General publications on the development of quantitative approaches

37. The conclusions of an IWGT working group on quantitative approaches to genetic toxicology risk assessment (QWG) is summarised in two publications. (MacGregor et al 2016 a, b –ANNEX 1). In the first publication methods to analyse dose responses and define the POD were examined; the second publication examined how the POD's can be used to establish acceptable exposure limits and to assess risk to humans.

38. *This QWG addressed the following topics:*

1. *The need for quantitative dose–response analysis of genetic toxicology data*
2. *The existence and appropriate evaluation of threshold responses*
3. *Methods to analyze exposure–response relationships & derive points of departure (PoDs) for extrapolation to low-dose exposure levels*
4. *Approaches to define exposure-related risks*
5. *Empirical relationships between genetic damage (mutation) and cancer*
6. *Extrapolation across test systems and species.*

39. In the first publication, the move towards identifying and accepting that all genotoxic chemicals will have an exposure level below which the risk is considered acceptable was discussed. As mentioned previously these explorations have arisen predominantly when examining genotoxic impurities in drug substances (ICH, EMA, FDA), where extensive testing is possible. It is noted that there is a diversity of products and potential exposure scenarios and therefore it is unlikely that one criterion will be suitable for all risk management situations. The discussion therefore aimed for a general framework.

40. The methods, NOGEL, BMD and breakpoint dose/bilinear (hockey stick) responses were detailed, with the group selecting BMD as the preferred POD. It was also agreed that the NOGEL was suitable but that there were reservations about the applicability of the breakpoint dose-response methodology. The group provide characteristics and recommendations for data acquisition for the methods and listed that advantages and disadvantages of each modelling method considered.

Advantages and disadvantages of methods considered by the QWG.			
Method	**BMD**	**NOGEL**	**Breakpoint dose determination using a bi-linear model**
Measurement	Benchmark dose – dose associated with a specific	Highest dose with no statistically significant	Estimate of threshold

	benchmark response (BMR)	response	
Advantages	Uses data efficiently and takes account of the shape of the dose response. (Fits a model to data) Currently used by many regulatory agencies	Is easy to apply Does not require dose response modeling	Sparse data tends to yield a lower PoD May be appropriate when mechanistic information supports threshold expectation
Disadvantages	Requires consensus on appropriate biologically relevant benchmark response (BMR) Continuous and quantal data are modeled differently	Sparse data tends to yield a higher PoD Statistical assumptions must be met	Based on specific assumption that data is described by one line of zero slope and another of finite slope is highly model dependent (other models will fit just as well but predict very different PoD) is not robust (PoD often cannot be determined for sparse data sets)

42. Critical to these approaches is establishing what constitutes a 'small risk' and whether it is masked within the usual background noise, and the normal distribution of genotoxicity endpoints.

43. The second publication reviews the approaches used to apply the PoD to define an acceptable exposure and evaluate human risk. It is clear that an understanding of MOA is important for determining which PoD is appropriate to use and more importantly how it will be used. Extrapolation below the PoD requires pragmatic use of uncertainty factors, which should take into account MOA, and the publication notes the importance of understanding TK/ADME characteristics. DNA or Hb adducts are considered appropriate exposure biomarkers which can also be used in risk assessment calculations (for example Viracept).

44. The working group considered the importance of quantitative correlations between genotoxic and carcinogenic potency in two contexts – that where the *in vivo* genotoxic potency showed a correlation with and without restrictions such as species, strain, sex or target tissue. The group considered the quantitative analysis by Hernandez et al (2011) to exhibit a high degree of correlation for the 18 compounds examined, when there were no restrictions on tissue, species etc. They conclude that despite not taking into account differences in metabolism, dose route or sex, etc that *'these results suggest, in the absence of carcinogenicity data, an estimate of probable cancer potency can be derived from in vivo genotoxicity studies'*. However there are only limited analyses of genotoxicity and carcinogenicity studies where direct correlations are possible, as the majority of the data are from [mouse] bone marrow MN and not target organ specific studies such as the transgenic assay - a worked example with MeIQx is presented. The working group were not able to generate a generalized scheme as there were insufficient chemicals with the required correlating datasets.

45. The possibility of using *in vitro* data to estimate *in vivo* risk was also considered by the working group. They suggest that there is potential for PBPK modelling to be used to predict ADME and therefore exposure but uncertainty factors, which represented the relevance and utility of the models, would have to be

used. For example, from a comparison of bacterial mutagenicity and carcinogenic potency a qualitative relationship could be established but a quantitative association is less well defined. Attention was drawn to the importance of appropriate metabolic activation in *in vitro* studies.

46. Overall the group supports the use of *in vivo* genotoxicity dose response data to determine PODs and that they can be used to establish regulatory exposure limits when combined with exposure data and pragmatic uncertainty factors. The importance of understanding the relationship between the selected genotoxicity endpoint and the chemical's MOA ie mutagenicity, clastogenicity or aneugenicity as a key event was emphasised.

47. Benford (2016 –ANNEX 1) discusses the use of the MOE approach in the risk management of substances in food which are genotoxic and carcinogenic and whether there is potential for BMD's from genotoxicity studies to be used instead of those from carcinogenicity studies. Attention is drawn to the importance of considering factors such as study design and quality, strain and species and chemical MOA when using carcinogenicity data and that these factors would also be critical if genotoxicity data are used. It is noted that a comparison of potency in carcinogenicity and genotoxicity assays is necessary using a broad range of carcinogen classes and MOAs.

48. Comparison of in vivo vs in vitro endpoints has also been considered (Soeteman-Hernandez et al 2015; Bemis et al 2016) and this can be considered in a further review .

Questions to the Committee.

49. This paper is presented as an introduction to the topic and provides examples of how data are being considered with view to using genotoxicity data in risk assessment scenarios in a quantitative manner.

1. Firstly, has anything fundamentally changed to allow genotoxicity data to be used QUANTITATIVELY rather than only qualitatively?

2. Has anything essentially changed in the interpretation of BMD values and if so how does this improve risk assessment?

3. What evaluations are needed before it is possible to move away from the yes/no answer basis of genotoxicity testing?

50. Further consideration of the following issues

1. How important is the MOA by which genotoxicity occurs and the sensitivity of assay to the chemical?

2. What are appropriate POD's for genotoxicity endpoints?

3. How to establish that non-threshold approach is appropriate? Are some genotoxic MOA's endpoints more likely to show a threshold than others? How should factors such as target organ exposure, route of exposure, metabolism, species differences be considered?

4. Is the 1:1 ratio really plausible for the relationship of genotoxicity to carcinogenicity? (eg what are Members opinion of the Sanner and Dybing calculations?)

5. How would quantitative information on genotoxicity be used: Derivation of MOEs for risk ranking, risk management prioritisation, safety assurance, etc. And is there a need to determine which regulatory frameworks would use these strategies?

Finally, the COM is asked to consider how they would like to take this area of work forward.

References

Barlow S), Renwick AG, Kleiner J, Bridges JW, et al (2006) Risk assessment of substances that are both genotoxic and carcinogenic report of an International Conference organized by EFSA and WHO with support of ILSI Europe. Food Chem Toxicol.:1636-50.

Bemis, J.; Wills, J; Bryce, S.; Torous, D; Dertinger, S.; Slob, W (2016) Comparison of in vitro and in vivo clastogenic potency based on benchmark dose analysis of flow cytometric micronucleus data Mutagenesis 31(3) pp. 277– 285

Benford DJ (2016) The use of dose-response data in a margin of exposure approach to carcinogenic risk assessment for genotoxic chemicals in food Mutagenesis 31(3) pp. 329–331.

Boobis A, Flari V, Gosling JP, Hart A, Craig P, Rushton L, Idahosa-Taylor E. (2013) Interpretation of the margin of exposure for genotoxic carcinogens - elicitation of expert knowledge about the form of the dose response curve at human relevant exposures. Food Chem Toxicol. 57:106-18.

Clewell, Rebecca A.; Andersen, Melvin E. (2016) Approaches for characterizing threshold dose–response relationships for DNA-damage pathways involved in carcinogenicity in vivo and micronuclei formation in vitro Mutagenesis 31(3) pp. 333–340

Dewhurst I, Renwick AG. (2013) Evaluation of the Threshold of Toxicological Concern (TTC)--challenges and
approaches. Regul Toxicol Pharmacol.;65(1):168-77

Doak SH, Jenkins GJ, Johnson GE, Quick E, Parry EM, Parry JM. (2007) Mechanistic influences for mutation induction curves after exposure to DNA-reactive carcinogens. Cancer Res. 67(8):3904-11.

Dybing E, O'Brien J, Renwick AG, Sanner T.(2008) Risk assessment of dietary exposures to compounds that are genotoxic and carcinogenic--an overview. Toxicol Lett. ;180(2):110-7.

EFSA (European Food Safety Authority). 2005. Opinion of the Scientific Committee on a request from EFSA related to a harmonized approach for risk assessment of substances which are
both genotoxic and carcinogenic. The EFSA Journal 282: 1-31.

EFSA (2016) Review of the Threshold of Toxicological Concern (TTC) approach and development of new TTC decision tree EFSA Supporting Publication 13(3) EN-1006

Felter S(1), Lane RW, Latulippe ME, et al (2009) Refining the threshold of toxicological concern (TTC) for risk prioritization of trace chemicals in food. Food Chem Toxicol.47(9):2236-45

Fukushima, Shoji; Gi, Min; Kakehashi, Anna; Wanibuchi, Hideki; Matsumoto, Michiharu (2016)
Qualitative and quantitative approaches in the dose–response assessment of genotoxic carcinogens
Mutagenesis 31(3) pp. 341–346

Gollapudi BB, Johnson GE, Hernandez LG, Pottenger LH et al (2013) Quantitative approaches for assessing dose-response relationships in genetic toxicology studies. Environ Mol Mutagen.54(1):8-18.

Hernández LG, Slob W, van Steeg H, van Benthem J. (2011) Can carcinogenic potency be predicted from in vivo genotoxicity data?: a meta-analysis of historical data. Environ Mol Mutagen. 52(7):518-28.

Ji, Zhiying; LeBaron, Matthew J.; Schisler, Melissa R.; Zhang, Fagen; Bartels, Michael J.; Gollapudi, B. Bhaskar; Pottenger, Lynn (2106) Dose–Response for Multiple Biomarkers of Exposure and Genotoxic Effect Following Repeated Treatment of Rats with the Alkylating Agents, MMS and MNU
Mutagenesis 31(3) pp. 297–308 H.

Jenkins GJ, Doak SH, Johnson GE, Quick E, Waters EM, Parry JM. 2005 Do dose response thresholds exist for genotoxic alkylating agents? Mutagenesis. 20(6):389-98. Epub 2005 Aug 31.

Johnson GE, Soeteman-Hernández LG, Gollapudi BB, et al 2014 Derivation of point of departure (PoD) estimates in genetic toxicology studies and their potential applications in risk assessment. Environ Mol Mutagen. ;55(8):609-23.

MacGregor JT(, Frötschl R), White PA(3), et al 2015 IWGT report on quantitative approaches to genotoxicity risk assessment I. Methods and metrics for defining exposure-response relationships and points of departure (PoDs). Mutat Res Genet Toxicol Environ Mutagen. 2015 May 1;783:55-65.

MacGregor JT, Frötschl R, White PA(3), et al 2015 IWGT report on quantitative approaches to genotoxicity risk assessment II. Use of point-of-departure (PoD) metrics in defining acceptable exposure limits and assessing human risk. Mutat Res Genet Toxicol Environ Mutagen. 2015 May 1;783:66-78.

Müller L, Gocke E. (2009) Considerations regarding a permitted daily exposure calculation for ethyl methanesulfonate. Toxicol Lett. 190(3):330-3

Pottenger LH, Gollapudi BB. (2010) Genotoxicity testing: moving beyond qualitative "screen and bin" approach towards characterization of dose-response and thresholds. Environ Mol Mutagen. 51(8-9):792-9

Sanner T, Dybing E. (2005) Comparison of carcinogenic and in vivo genotoxic potency estimates. Basic Clin Pharmacol Toxicol. 96(2):131-9.

Soeteman-Hernández LG(1), Fellows MD(2), Johnson GE(3), Slob W(1). (2015) Correlation of In Vivo Versus In Vitro Benchmark Doses (BMDs) Derived From Micronucleus Test Data: A Proof of Concept Study. Toxicol Sci. 148(2):355-67.

Thompson CM, Bichteler A, Rager JE, et al (2016) Comparison of in vivo genotoxic and carcinogenic potency to augment mode of action analysis: Case study with hexavalent chromium. Mutat Res 800-801:28-34.

Walker VE(1), Casciano DA, Tweats DJ. (2009) The Viracept-EMS case: impact and outlook. Toxicol Lett. 190(3):333-9.

White PA , Johnson GE (2016) Genetic toxicology at the crossroads—from qualitative hazard evaluation to quantitative risk assessment Mutagenesis 31 (3) pp. 233–237

Zeller, A; Tang, L; Dertinger, s Funk, J; Duran-Pacheco, G Guérard, M 2016
A proposal for a novel rationale for critical effect size in dose–response analysis based on a multi-endpoint in vivo study with methyl methanesulfonate Mutagenesis 31(3) pp. 239– 253