CC/2019/14

COMMITTEE ON CARCINOGENICITY OF CHEMICALS IN FOOD, CONSUMER PRODUCTS AND THE ENVIRONMENT (COC)

First draft revised Guidance Statement (G05): Defining a Point of Departure and Potency Estimates in Carcinogenic Dose Response

1. The COC has periodically published guidelines for the evaluation of chemicals for carcinogenicity. The first guidance was published in 1982 and has undergone several updates since then to reflect advances in development and validation of methods for assessing risk of chemical carcinogenicity.

2. These updates included the separation of the overall guidance into individual documents during 2012 – 2014 to allow faster revisions to be made in the case of rapidly developing areas. This included a separate document addressing points of departure and potency estimates in carcinogenic dose response (G05).

3. Since publication of the first version of G05, EFSA and WHO have jointly reviewed the use of the threshold of toxicological concern (TTC) approach (EFSA and WHO, 2016) and EFSA has published new guidance on bench-mark dose (BMD) modelling (EFSA, 2017). An updated version of G05 containing the relevant amendments in light of the new BMD guidance was agreed in September 2018. At that time, it was also agreed that a full review of G05 would be undertaken when EFSA published further work on the TTC approach.

4. Updated guidance on the use of the TTC approach was published by EFSA in April 2019. The Scientific Committee confirmed that the approach '*is a pragmatic screening and prioritisation tool for use in food safety assessment'* for use in circumstances where the chemical structure of the substance of interest is known, there is limited specific toxicity data and the exposure can be estimated.

5. This discussion paper presents a first draft revised guidance statement G05. All sections have been reviewed and updated where needed.

Questions for the Committee

6. Members are asked to:

   i. Comment on the structure of the first draft revised document.

   ii. Comment on whether the current level of detail included is appropriate.

iii. Comment on whether the nomenclature of DNA-reactive mutagens and/or carcinogens adopted by EFSA in the TTC guidance should be adopted in a) the TTC section and b) more widely in G05.

iv. Address the questions given throughout the document

v. Comment on whether the Committee recommendations are still appropriate.

NCET at WRc/IEH-C under contract supporting the PHE COC Secretariat
November 2019

Committee on CARCINOGENICITY

Committee on Carcinogenicity of Chemicals in Food, Consumer Products and the Environment (COC)

COC Guidance Statement G05 –First Draft v2-0

Defining a Point of Departure and Potency Estimates in Carcinogenic Dose Response

https://www.gov.uk/government/groups/committee-on-carcinogenicity-of-chemicals-in-food-consumer-products-and-the-environment-coc

COC Secretariat

c/o Public Health England

Centre for Radiation, Chemical and Environmental Hazards

Chilton, Didcot, Oxfordshire OX11 0RQ

Committee on CARCINOGENICITY

COMMITTEE ON CARCINOGENICITY OF CHEMICALS IN FOOD, CONSUMER PRODUCTS AND THE ENVIRONMENT

Defining a Point of Departure and Potency Estimates in Carcinogenic Dose Response

Contents

Note to Members for November 2019 meeting: *The order of sections is mainly based on the previous version with slight amendments. However,*

consideration should be given as to whether a comprehensive re-ordering is required?

1.0 Introduction

1. This guidance statement provides an overview of the various methods for deriving points of departure (POD; also known as a reference point) and potency estimates associated with exposures to chemical carcinogens. It is part of a series of guidance statements by the Committee on Carcinogenicity of Chemicals in Food, Consumer Products and the Environment (COC). It should be read in conjunction with these, in particular G01 on the overall strategy of risk assessment of chemical carcinogenicity, G02 on synthesising epidemiological evidence, G03 on hazard identification and characterisation: conduct and interpretation of animal carcinogenicity studies, and G06 on cancer risk characterisation methods.

2. This guidance document describes how to derive POD such as the No Observed Adverse Effect Level / Lowest Observed Adverse Effect Level (NOAEL/LOAEL; section 2.1) Benchmark Dose (BMD; section 2.2) and potency estimates such as the TD50 (section 2.4) and T25 (section 2.5), and describes how they can be used to estimate the relative potency of carcinogens. This guidance document also details the Threshold of Toxicological Concern (TTC) approach which can help assessment of chemicals for which there are a lack of chemical-specific toxicity data and identify priorities for a more detailed substance-specific risk assessment.

3. Hazard characterisation involves a qualitative description of the nature of the hazard and a quantitative description of the change in effect caused by differing doses of a chemical substance after a certain exposure time, i.e. the dose-response relationship. The purpose of analysing the dose-response relationship is to investigate the magnitude of response (in terms of severity or incidence) within the dose range used in an animal study or within the range of exposures experienced in a human study. This helps to estimate the response and, ultimately, the risk from the levels of exposure to the chemical in the environment, food etc. Environmental levels are usually much lower than doses used in animal studies and often also lower than those to which individuals have been exposed in studies used to characterise effects in humans (e.g. observational epidemiological studies in occupational and non-occupational cohorts). The relationship between dose and response may be used to aid hazard characterisation by allowing a comparison of carcinogenic potency. These estimates give an indication of the dose of a substance administered over a standard animal lifespan that results in a fixed incidence of tumours, such as, 5, 25 or 50%, after correction for the spontaneous background incidence of tumours among controls (Barlow et al., 2006). However, other important factors that can affect this relationship in humans, and should be further considered, are species differences in absorption, distribution, metabolism and excretion (ADME), mode of

action and variability in susceptibility between species (inter-species) and within humans (intra-individual).

4. There are a number of methods for the characterisation of hazard due to the potential of a compound to be a genotoxic or non-genotoxic carcinogen. In all of these, chemicals are classified with regard to tumourigenicity on the basis of potency. In this context, potency is ideally represented by the overall position and shape of the dose-effect or dose-response curve, but the value (dose) at a particular point on the curve is often used as a surrogate. A POD is defined as the dose-response point that marks the beginning of a low-dose extrapolation. This point can be the lower bound on dose for an estimated incidence of a tumour or a change in response level from a dose-response model (BMD), or a NOAEL or LOAEL for an observed incidence, or change in level of response. An example is the dose level associated with a tumour incidence that is 10% above the incidence in the control group. The Committee recognises that where data on tumourigenicity *per se* are lacking, it may be possible to use continuous data as a surrogate measure of response, such as specific DNA damage observed in target organs, for determining a POD.

5. It should be noted that there is no difference in the methodology used for determining PODs for genotoxic and non-genotoxic carcinogens. It is how the dose-response relationship and the POD are used in the final assessment of risk that varies, depending on whether or not a carcinogenic response occurs through a genotoxic or non-genotoxic mode of action (see Guidance Document G06 for further discussion of Cancer Risk Characterisation).

2.0 Points of Departure and Potency Estimates

2.1 The NOAEL (No Observe Adverse Effect Level) approach

6. For the majority of toxicological effects, with the exception of most genotoxic effects or where extensive testing has failed to identify a threshold (e.g. in the case of lead), it is generally assumed that there is an exposure threshold below which no adverse effects occur. The highest administered dose at which no statistically significant adverse difference from the concurrent control group is observed is designated the No Observed Adverse Effect Level (NOAEL) and is used as a POD in risk assessments. This avoids unnecessarily conservative risk estimates as assessment is based on adverse effects rather than on minor or adaptive effects.

7. If a statistically significant adverse effect is observed at all dose levels, the lowest dose used in the study, i.e. the LOAEL (Lowest Observed Adverse Effect Level), may be used as the POD. Typically, the NOAEL (or if one is not available, the corresponding LOAEL) is determined for the most sensitive, relevant effect identified in human epidemiological studies or from sub-chronic or chronic studies in laboratory animals.

8. The NOAEL approach has traditionally been the method of choice for determining a POD for such effects, including carcinogenicity by a non-genotoxic

mode of action. In human risk assessment, the NOAEL has been used to establish health-based guidance values such as ADIs for food additives and pesticide residues, and TDIs or tolerable weekly intakes (TWIs) for contaminants. These health-based guidance values are derived from the highest NOAEL for the.

9. The NOAEL approach has been the standard method for deriving PODs for a long time and it is familiar to most risk assessors (US EPA, 2000). However, the NOAEL approach has a number of limitations, with a major limitation being the constraint for the NOAEL to be one of the experimental doses. The approach does not take into consideration dose spacing, the shape of the dose-response curve, the number of animals per group, or the statistical variation in the response and its measurement. The NOAEL approach tends to give lower health-based guidance values for studies with a higher power to detect adverse effects, which in effect 'penalises' better-designed studies (IPCS, 2009a). It should also be noted that studies with low power (e.g. small group sizes) and/or insensitive methods may only detect relatively large effects, resulting in higher NOAELs.

> ***Note to Members for November 2019 meeting:*** *Should this section contain more detailed information on the approach?*

2.2 Benchmark Dose (BMD) approach

10. The BMD methodology was initially introduced by Crump (1984) as an alternative to the use of NOAELs and Lowest Observed Adverse Effect levels (LOAELs) in dose–response assessment. It was subsequently developed further within the US EPA (US EPA, 1995).

11. The BMD_x is defined as the dose that corresponds to a specific change (x%) in response compared to the (modelled) response in control animals, the benchmark response (BMR) (Crump, 1995). The BMD is determined by fitting a range of mathematical curves to the dose-response data over the range of observable responses from animal studies or human studies (if available), using a selection of different models. From each statistically acceptable modelled dose–response curve, values for the BMD and the lower and upper bound 95% confidence limits (BMDL and BMDU) are obtained. To take experimental uncertainty into account, the lower 95% confidence bound on the benchmark dose ($BMDL_x$) is used as the POD. Figure 1 illustrates the BMD approach.

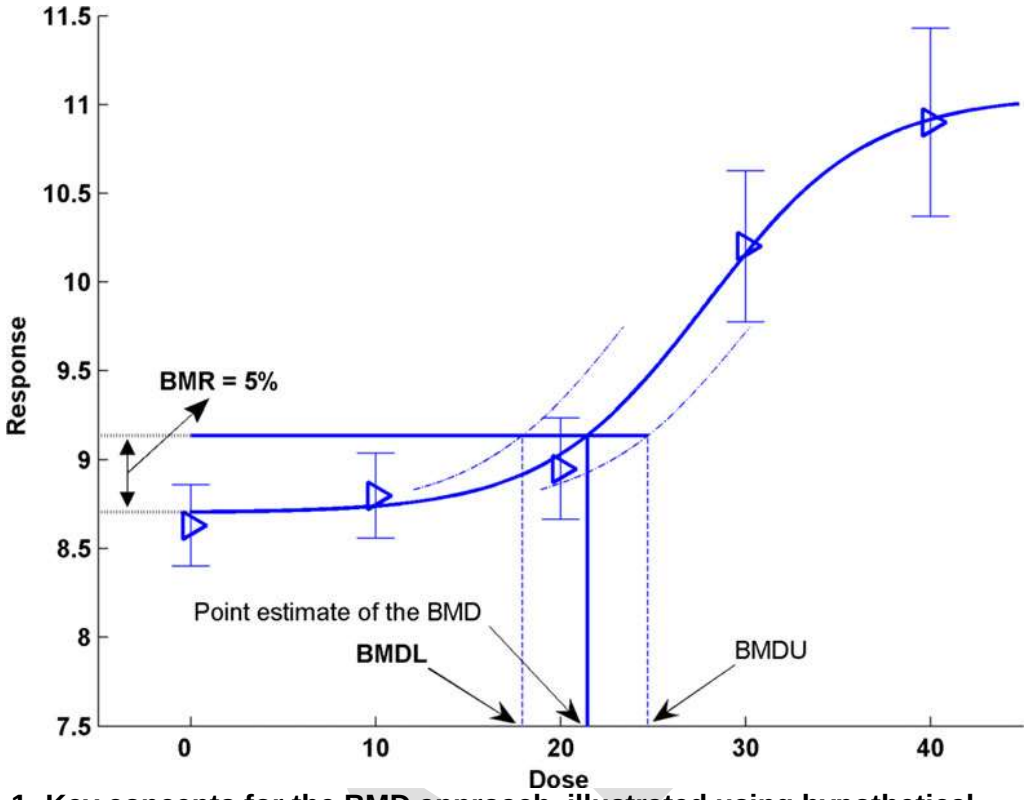

Figure 1: Key concepts for the BMD approach, illustrated using hypothetical continuous data.

Source EFSA (2017). The observed mean responses (triangles) are plotted, together with their confidence intervals. The solid curve is a fitted dose–response model. This curve determines the point estimate of the BMD, which is generally defined as a dose that corresponds to a low but measurable change in response, denoted the benchmark response (BMR). The dashed curves represent, respectively, the upper and lower 95% confidence bounds (one sided)4 for the effect size as a function of dose. Their intersections with the horizontal line are at the lower and upper bounds of the BMD, denoted BMDL and BMDU, respectively. It should be noted that the BMR is not defined as a change with regard to the observed mean background response, but with regard to the background response predicted by the fitted model. This distinction is important because, in general, the fitted curve does not hit the observed background response exactly (so that adding the BMR to the observed background response will in general not provide the correct intersection with the dose–response at the BMD). In the Figure, the BMD corresponds to a 5% change in response relative to background (BMR = 5%). The fitted curve yields an estimated background response of 8.7, and a 5% increase of that equals 9.14 (= 8.7 + 0.05 9 8.7). Thus, the BMD_{05} of 21.50 is obtained from the intersection of the horizontal line, at a response of 9.14, with the fitted dose–response model. In this example, the $BMDL_{05}$ has a value of 18.

12. Both dichotomous data and continuous data from animal and human dose-response studies can be evaluated using the BMD approach (EFSA, 2017). Dichotomous (quantal or incidence) data describe whether an effect has occurred in an individual or not, e.g. presence of tumour, death. The tumour data obtained from animal carcinogenicity studies fall into the dichotomous category. Continuous data are typically quantitative measurements.

13. The human dose-response data is generally more difficult to interpret than animal dose-response data due to the presence of confounders and imprecision in the exposure estimates (EFSA, 2017).

14. For the purpose of this guidance statement on defining a POD in a carcinogenic dose response, only considerations of quantal data will be discussed, though it is acknowledged that for effects such as DNA damage an approach appropriate for continuous data would need to be used.

15. Before a dataset is analysed using the BMD methodology, it is necessary to evaluate all available studies and potential critical effects, ensuring that the datasets meet minimum criteria, as outlined in the EFSA opinion (2017). There are two aspects here, and in part they depend on the BMD approach to be used. One is selection of tumour-response data relevant to risk assessment of a genotoxic carcinogen and the other is ensuring a dataset is suitable for modelling. In the BMD approach, one might model all suitable datasets or combinations thereof (accepting the need for caution in combining data) and then interpret the resulting BMDLs, or one may choose to model only what is considered to be the critical data set, the one likely to give the most conservative outcome from amongst those that are considered relevant.

16. In preparation for BMD modelling, the BMR must also be chosen. For quantal responses, the BMR is expressed in terms of a percent increase in risk[1] of adverse outcome above the modelled background. The BMR is typically set at the lower end of the range of responses that can be detected experimentally, or the observations in epidemiological studies. EFSA, (2017) recommend that a default BMR value of 10% be used for quantal data from a guideline rodent carcinogenicity study, since the modelling of lower responses generally results in greater uncertainty. Based on statistical and toxicological considerations, a modified BMR can be used, for example a BMR of 1% has been used with epidemiological studies of large populations (US EPA, 2012 and EFSA, 2017).

17. Both the WHO/IPCS and EFSA have produced guidance on dose–response modelling, including guidance on cancer dose-response data (IPCS, 2009b and EFSA, 2017). Different models that fit the data equally well, as judged by statistical comparison, can result in different BMDs and BMDLs, reflecting model uncertainty. The selection of the group of models to investigate is dependent on the endpoint being modelled (quantal or continuous) and the experimental design used to generate the data (e.g. number of dose groups utilised and nested study design (Davis et al., 2011)). The US EPA technical guidance document (2012) and the EFSA guidance (2017) both detail the various models that can be used in BMD modelling with existing software. Model selection and model constraints are important considerations in BMD estimation. The main option in model selection for BMD estimation using quantal data is the choice of model classes (Sand et al., 2008).

1 This can be expressed as ‘extra’ risk (the default in BMDS) or as ‘added’ risk. The BMR is calculated differently, depending on which risk type is chosen.

18. Once the selected models have been fitted to the data, a series of scientific judgements must be made to ensure the fitted models adequately describe the data. Different types of statistical testing can be utilised to assess the adequacy of model fit. For model selection, an important criterion is that the selected model should adequately describe the data, especially in the region of the BMR. The earlier EFSA guidance (2009) for model fit involves two principles: deciding which model fits best within a nested family of increasingly complex models, where this is necessary, and then a determination of overall goodness-of-fit. Both principles are based on the likelihood-ratio test and EFSA (2009) recommended a minimum goodness of fit value of p = 0.05 for model acceptance based on log-likelihood. For dichotomous data, the US EPA software employs Pearson's chi-squared goodness of fit test (US EPA, 1995). The US EPA (2012) recommends a minimum goodness of fit p value of p = 0.1 for model acceptance. The Akaike information criterion (AIC) value, which is a measure of the degree of fit weighted by the number of free parameters in the model and/or Pearson's chi-squared goodness of fit test can also be used for selection within a nested series. The latest EFSA guidance recommends that the AIC should be used (instead of log-likelihood) to characterise goodness of fit (EFSA, 2017).

19. Although the current international guidelines for study design have been developed with the NOAEL approach in mind, they offer no obstacle to the application of the BMD approach. The current guidelines may, however, not be optimal given that the BMD approach allows for more freedom in balancing between number of dose groups and group sizes (Slob, 2014). As these guidelines are revised, e.g. within the OECD Test Guidelines Programme, the possibility to recommend study designs that tend to result in better dose–response information (e.g. more dose levels with the same total number of animals) should be taken into account.

20. It is often the case that a number of models will adequately fit the data, as judged on statistical considerations. One option then is to select the model with the lowest AIC value from all statistically acceptable models. However, applying this approach may lead to models being excluded, which would otherwise provide higher, or lower, risk estimates. A second option is to select the model that leads to the highest extra risk or lowest BMDL on the basis that this selection is likely to be more conservative. This option was recommended by EFSA (2009). A third option is to report a range of risk estimates from those models that provide an acceptable fit to the observed data. A fourth option is to average risk estimates/BMDLs based on the support for each model provided by the data ('Model Averaging') (Wheeler and Bailer, 2007). It should be noted that this is not the simple averaging of the individual BMDL estimates, but a pooled analysis of the data. This approach better characterises the uncertainty in the value of the BMDL that derives from ignorance of the true dose response (Wheeler and Bailer, 2007), and hence is expected to be numerically higher than the lowest BMDL value resulting from applying a suite of models (Benford et al., 2010). Revised EFSA guidance now recommends Model Averaging as the preferred approach, combining results from each of the fitted models to establish a final BMD confidence interval. However, selection/rejection of

models can be considered as a sub-optimal alternative in situations where Model Averaging tools are not available (EFSA, 2017).

21. Different software programs are currently available for BMD analysis. The US EPA developed the Benchmark Dose Software (BMDS – available from: https://www.epa.gov/bmds, accessed 29/09/2019). PROAST is another BMD software package (available from: https://www.rivm.nl/en/Documents_and_publications/Scientific/Models/PROAST, accessed 29/09/2019), developed by the Dutch National Institute for Public Health and the Environment, and the basis on which EFSA provide a web-based platform for performing BMD analysis. Both these software packages are suitable for dose-response analysis and deriving a BMDL from the dose-response data. The Dutch National Institute of Public Health and the Environment and EPA collaborate to achieve consistency between the BMDS and PROAST software. At the current time, BMDS does not incorporate model averaging tools.

22. Once the BMDL is derived as the POD, the assessment moves to the risk characterisation stage which brings together hazard identification and hazard characterisation and the exposure assessment process (see Risk Characterisation Guidance Statement G06). The BMDL can be used for setting regulatory levels such as acceptable or tolerable daily intakes (ADIs or TDIs) or reference doses/concentrations (RfD/RfCs) for effects for which it is assumed there is a threshold. Both the European Food Safety Authority (EFSA) and the World Health Organization (WHO) recommend the BMD approach for deriving a POD to be used as a starting point for human health risk assessment, for all endpoints, including carcinogenicity by a genotoxic mode of action.

> ***Note to Members for November 2019 meeting:*** *There is a lot of detail in this section compared to others in the document. What are the key points to put across on the BMD and what level of detail is required?*

2.3 Comparing NOAEL and BMD methodologies for use in risk assessment

23. The BMD approach has a number of advantages over the NOAEL approach in that it makes more complete use of the available dose–response data, takes into account the shape of the dose-response curve more explicitly and is less dependent on dose spacing. BMD also enables quantification of the uncertainties in the dose-response data using statistical methodology (EFSA, 2017).

24. Despite the adoption of the BMD approach as an alternative to the NOAEL in determining a POD, there continues to be a need for the NOAEL/LOAEL approach. Not all data sets are amenable to BMD modelling, such as those resulting from incomplete data availability or from a lack of models that can describe a dataset adequately (US EPA, 2012) and the NOAEL approach can be used in this instance. A typical situation where the NOAEL approach is applicable, whereas the BMD approach is not, is when there is a response only in the highest dose group.

2.4 The TD50 approach

25. The TD50 (Peto et al., 1984) is defined as the chronic dose rate which would induce tumours in a given target site(s) in 50% of the test animals at the end of a standard lifespan for the species, provided that there were no tumours in control animals. However, since the tumour(s) of interest often do occur in control animals, the TD50 is more precisely defined as the daily dose rate required to halve the probability of remaining without tumours at the end of a standard life span. TD_{50} values have been estimated for chemicals listed in the Carcinogenic Potency Database (CPDB) developed by Gold and Zeigler (https://toxnet.nlm.nih.gov/cpdb/cpdb.html, accessed 29/09/19) (Gold et al., 1984, 1997).

26. The TD50 concept is based on the assumption that there is linearity between dose and hazard until tumour onset, which may be complicated by premature deaths from causes other than tumour formation. The concept also depends on the assumption that tumour onset times are observable prior to mortality and, as a result, the approach relies heavily on careful observation of the animals. Tumours that are discovered after death within the study period may cause confounding between mortality and tumour onset and would ultimately result in a biased TD50 estimate. Alternatively, tumours that do not significantly alter survival and remain undiscovered until death would result in the TD50 value relating to the 'rate of death with tumour', rather than the tumour incidence rate. This undermines the objective of the carcinogenicity study, which is to evaluate tumour incidence. A description of the TD50 methodology and the complex statistical analysis involved in its derivation is provided at http://toxnet.nlm.nih.gov/cpdb/td50.html (accessed 29/09/19).

27. The Committee reiterates its previous position that the TD50 is a practical quantitative estimate of carcinogenic potency for the ranking of genotoxic carcinogens, but not for deriving a POD.

> ***Note to Members for November 2019 meeting:*** *Should this section contain more detailed information on the approach?*

2.5 The T25 approach

28. Although primarily used in carcinogenic potency estimates, the T25 approach can also be used to derive a POD. For example, although the European Chemicals Agency (ECHA) prefers $BMDL_x$ as a starting point, if the data do not permit BMD analysis, ECHA suggests that the T25 can be used. The T25 is defined as the dose eliciting a 25% increase in the incidence of a specific tumour above the background level within the standard lifespan of that species. It was originally proposed by Dybing et al. (1997) and further developed by Sanner et al. (2001). The methodology does not require elaborate statistical methods. The T25 is determined by simple linear interpolation or, in some cases, extrapolation beyond the data points. According to Dybing et al. (1997) the data used for calculating a T25 should preferentially be from long-term carcinogenicity bioassays. The estimation of T25 is

dependent on the incidence of tumours at a selected site at a single dose level. The minimum data requirements to calculate a T25 are one incidence level significantly greater than the controls (Gillespie et al., 2011). The T25 is influenced by the quality of the bioassay information (e.g. design and evaluation of studies) and factors such as time to first tumour, the influence of toxicity on tumour induction and mortality, and the approach taken regarding statistical analysis of tumour data.

29. The European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) has evaluated the use of T25 estimates for regulatory risk assessment of non-threshold carcinogens (ECETOC, 2002). There may also be uncertainties regarding the application of the T25 for potency ranking, particularly with regard to selection of the most sensitive site relevant for humans, the relevance of rodent tumours for humans, and different cancer susceptibilities between rodent species (ECETOC, 2002). The T25 is also the method used by the EU to assess relative potency for the setting of specific concentration limits of preparations and mixtures (EC, 1999). Using the T25 method, Sanner and Dybing (2005) found a good correlation between the values based on human epidemiological data and those based on animal experiments, although the data available for such comparison were limited. Previously, the T25 approach has been used in risk assessment for regulation of non-food, genotoxic carcinogenic chemicals in the EU (EFSA, 2005).

> ***Note to Members for November 2019 meeting:*** *Should this section contain more detailed information on the approach?*

2.6 Comparing BMD and T25 methodology for use in risk assessment

30. T25 and the BMD methodology differ in that the T25 is calculated from one data point on the dose-response curve whereas the BMD is derived from dose-response modelling of all available data on the dose-response curve (EFSA, 2005).

31. Dybing et al. (2008) compared the Margin of Exposure (MOE), the numerical value obtained by dividing a POD on the dose-response curve by estimated human exposure to the chemical, for 6 substances obtained using either the $BMDL_{10}$ or the T25. They found that MOEs obtained using the T25 as the POD were on average around 2.35 times higher than those derived using the $BMDL_{10}$ as the POD (Dybing et al., 2008). Benford et al. (2010) compared MOEs for 12 substances in food that are genotoxic and carcinogenic (5 of which were the same as those examined by Dybing et al., 2008) and found that the ratio of MOEs derived from a T25 value varied from those using a $BMDL_{10}$ value by between 0.9 and 4.6, with a mean of 2.9 and a median of 2.6. These results were in line with the expected ratio of 2.5 to account for the 25% vs. 10% risk, assuming linearity in the dose-response relationship, when comparing the T25 with the $BMDL_{10}$ (Benford et al., 2010).

32. In the case where dose-response data are inadequate for deriving an estimate of the BMD_{10} or $BMDL_{10}$, EFSA (2005) recommended the use of the T25 as a means of deriving a POD. However, use of this approach when it was not possible to derive a $BMDL_{10}$ was questioned by Benford et al. (2010). As the BMD

methodology uses all the available data, if a dataset does not allow derivation of a BMDL, even at a dose rate higher than 10%, e.g. a $BMDL_{25}$, the dataset may not be suitable for derivation of a meaningful POD at all. An example would be if there is a very high incidence of tumours at all dose levels, in which case it would not be feasible to derive a BMDL but a T25 could still be calculated by dividing the lowest dose level by the ratio of the percentage response to 25%. However, the resultant T25 value would be meaningless.

33. In the Committee's discussion of the MOE approach for G06, the guidance document on cancer risk characterisation methods, the Committee considered the use of the BMD approach as a means of deriving a POD to be superior to that of the T25.Therefore, in the event that it is not possible to derive a $BMDL_{10}$, the Committee does not recommend the routine use of the T25 for risk characterisation.

3.0 Potency Ranking of Genotoxic Carcinogens

> ***Note to Members for November 2019 meeting:*** *Should this section include genotoxic carcinogens and potent non-genotoxic carcinogens?*

34. Relative potency estimates could have some pragmatic use in carcinogenic risk assessment as an aid in the prioritisation of genotoxic carcinogenic substances but are not considered adequate for quantifying cancer risks. The uncertainties inherent in potency ranking mean that relative potencies should not be over-interpreted. For example, it is unclear whether the relative ranking identified in the observed dose range would be maintained at low doses, and whether the relative potency in animal studies would be applicable to humans.

35. Data from animal bioassays can be used to rank carcinogenic potency without reference to human intake. Carcinogenic potency estimates, as described in paragraph 3, make use of the available dose-response data, and points of departure can be derived from TD50, T25 or BMD approaches for use in potency ranking. For example, in a series of publications Gold et al. tabulated data on a large number of compounds allowing their carcinogenic potencies to be expressed as the TD50 (Gold et al., 1997). These values can be used to indicate the relative potencies of a series of compounds.

36. Potency Equivalence Factors (PEFs) have been suggested in circumstances where there is a good surrogate compound for comparison, e.g. inhalation of polycyclic aromatic hydrocarbons (PAHs) (Collins, 1998; Pufulete et al., 2004). Pufulete et al. (2004) suggested that an approach based on PEFs could be developed to include highly potent PAHs provided an appropriate reference data set for relevant PAHs using a route acceptable for inhalation risk assessment is selected. The US EPA (2010) also developed an approach to assessing cancer risk for PAH mixtures using relative potency factors (RPFs), which estimates the cancer risk of individual PAHs relative to that of benzo[a]pyrene (BaP). The US EPA

suggests that these RPFs are applicable to all routes of exposure but acknowledges that there is appreciable uncertainty in doing this. The COC notes that PHE has adopted a surrogate marker approach rather than the use of PEFs for assessment of the public health risk of PAHs in contaminated land (PHE,2017).

Note to Members for November 2019 meeting: *Should PEFs be included in G05?*

37. Comparing the TD50 and T25 approaches for estimating potency, the TD50 has an advantage in that it takes account of effects of chemicals on survival, however it requires specific software to undertake its derivation. In contrast, the T25 is quick and easy to calculate. There is evidence of a good correlation between rank order produced by TD50 and T25 (Dybing, 1997). In 2006, the COC compared the TD50 with the T25 in an attempt to develop an approach for potency ranking of genotoxic carcinogens for single exposure. Very limited data were available for this purpose and little correlation was found among those substances for which it was possible to obtain chronic TD50 and T25 values, compared to acute T25 values (COC, 2006).

38. The Committee acknowledges that the T25 approach can be used in potency ranking of genotoxic carcinogens but is of the view that the statistics should not be over-interpreted. The reason for this is that there are a number of basic uncertainties, such as whether the relative ranking identified in the observed dose range would be maintained at low doses, and whether the relative potency in animal studies would be applicable to humans. Currently, there is no need to use the T25 to rank non-genotoxic carcinogens, for which tolerable exposure levels can be derived using an approach based on knowledge of mode of action, identification of a NOAEL, and the use of uncertainty factors.

4.0 The Threshold of Toxicological Concern (TTC)

Note to Members for November 2019 meeting: *The following section outlining the TTC approach has been drafted on the basis that the Committee is in agreement with the latest EFSA Guidance. This includes use of the nomenclature 'DNA reactive mutagenicity and/or carcinogenicity' as cited by EFSA.*

39. The TTC approach is used to screen and prioritise the risk assessment of substances with a known chemical structure but no, or little, specific toxicity data. Application of the TTC approach has been most widely applied to the oral route of exposure and, as such, the following sections focus on that route. Application of the TTC approach to inhalation and dermal exposure routes is not as widely applied but has been considered by SCCS/SCHER/SCENIHR (2012). For the TTC approach to

be applied the estimated exposure of humans to the substance via the oral route should be low (EFSA, 2019).

Note to Members for November 2019 meeting: *It should be noted that 'low' has a value judgment, see discussion in paper CC/2019/16 on combined exposures*

4.1 Development of the TTC

40. The use of *de minimis* exposure values as a means of identifying substances of low concern was first proposed by Frawley (1967). This was further developed by the US Food and Drug Administration (FDA) (Rulis, 1986) for application to substances that do not contain a structural alert for potential DNA reactive mutagenicity and/or carcinogenicity. An analysis of the 500 carcinogens then present in the Cancer Potency Database (CPDB) allowed derivation of virtually safe doses (VSDs) (for a 1 in 10^6 excess cancer risk) from their TD_{50}s. This led to the adoption by the FDA (1995) of a Threshold of Regulation of 0.5 µg/kg of diet (equivalent to an intake of 1.5 µg/person/day) for all substances used in food contact materials. At this level, it was intended that consumers would be protected 'with reasonable certainty of no harm', even if that substance was later shown to be a carcinogen.

41. Cheeseman et al. (1999) later analysed and validated the approach using the expanded CPDB containing information on 700 chemicals (Gold et al., 1997). Cheeseman et al. (1999) identified that the Threshold of Regulation (ToR) of 0.5 µg/kg of diet would not be sufficiently protective for certain categories of potent carcinogens. . These were azoxy compounds, benzidines, N-nitrosamines and aflatoxin-like compounds. A number of other groups were excluded, but these were non genotoxins and potency was estimated using linear extrapolation from the TD50, which would result in an appreciable overestimate. With the exclusion of these structural classes, it was considered unlikely that an unstudied compound would be both carcinogenic and have a potency far greater than the typical potency of studied compounds.

42. Subsequently, Kroes et al. (2004) re-evaluated the distribution of VSDs for carcinogens, grouped into structural classes, e.g. aromatic amines, benzidines. They concluded that, with a few exceptions, adequate protection would be provided (i.e. there was low probability that the excess cancer risk from an untested chemical would be greater than 1 in 10^6), even from compounds that were potential DNA reactive mutagens and/or carcinogens, using a TTC value of 0.15 µg/person/day. Groups of compounds that would not be covered by this value were aflatoxin-like, azoxy and N-nitroso-compounds, as well as steroids and dioxins, because of their very high potencies, thus largely confirming the conclusions of Cheeseman et al. (1999). In their joint opinion published in 2012, the SCCS, SCHER and SCENIHR committees (Scientific Committee on Consumer Safety, Scientific Committee on Health and Environmental Risks and the Scientific Committee on Emerging and Newly Identified Health Risks respectively) added benzidines and hydrazines to this exclusion list. In their latest Guidance, EFSA (2019) no longer exclude hydrazines as only 4% (2 of 57) would be expected to exceed a cancer risk of 1 in 10^6 at the TTC

value for potential DNA-reactive mutagens or carcinogens (i.e. 0.0025 µg/kg bw). However, organosilicon substances are newly excluded based on the lack of representation of these in the toxicity database of Munro et al. (1996).

43. When considered by the COC for the 2004 version of the guidelines, the application of the TTC to carcinogens was a relatively new approach and the Committee concluded that:

> "*careful consideration was needed of the biological, analytical and mathematical issues as well as a much wider database for validation. The Committee consider that it should not currently be used as a generic approach, as the proposed exclusions covered some important classes of genotoxic carcinogens (such as aflatoxin-like compounds, azoxy compounds and N-nitroso compounds) and a number of classes of other carcinogens, such as heavy metals and TCDD (Kroes* et al., *2004). However, as it is based on ranking by theoretical risk and exposure the Committee agree that it could be used, along with hazard identification and characterisation data, for prioritisation of chemicals, particularly for chemicals that are not subject to regulatory approval schemes."* (COC, 2004).

Since 2004, experience on the application of the TTC approach has increased, and the approach itself has been refined, including proposals for use both for inhalation (Carthew et al., 2009; Escher et al., 2010, Tluczkiewicz et al., 2016) and dermal (Safford et al., 2008; Safford et al., 2011; Safford et al., 2015; Roberts et al., 2015) exposure. The TTC approach has been reviewed by EU committees (SCCS/SCHER/SCENIHR, 2012; EFSA, 2012b; EFSA/WHO, 2016; EFSA, 2019). A paper describing the development of the TTC concept since its introduction in 1995 and the respective EU committee opinions was presented at COC in 2012, and a further update was given to the Committee in 2017 (Paper CC/2012/18, available from: http://webarchive.nationalarchives.gov.uk/20140506122048/http://www.iacoc.org.uk/papers/index.htm and CC/2017/21, available from: https://www.gov.uk/government/groups/committee-on-carcinogenicity-of-chemicals-in-food-consumer-products-and-the-environment-coc#meetings).

44. In their analysis, Munro et al. (1996) utilised a dataset comprising repeat dose oral toxicity data for 613 organic chemicals with 2941 associated No Observed Effect Level (NOEL) values derived from a variety of non-cancer endpoints from sub-chronic, chronic, reproductive and developmental toxicity studies carried out in rodents and rabbits. NOELs for sub-chronic studies were adjusted to chronic exposure equivalents using a factor of 3. The 5th percentiles of the NOELs, grouped according to their respective Cramer et al. (1978) class (i.e. Class 1, 2 or 3), were used to derive TTC values by multiplying by 60 (assuming an average individual weighs 60 kg) and then dividing by a safety factor of 100, mirroring the ADI approach. This resulted in TTC values of 1800 µg/person/day (30 µg/kg bw per day) for Class I chemicals, 540 µg/person/day (9 µg/kg bw per day) for Class II chemicals and 90 µg/person/day (1.5 µg/kg bw per day) for Class III chemicals, respectively.

The Cramer decision tree (1978) distinguishes compounds with structures indicating low oral toxicity (Class I) from those where the structure allows no presumption of safety to be made or indicates significant toxicity (Class III). Class II is an intermediate class where compounds do not have the characteristics of Class III, but neither are they innocuous (Cramer et al., 1978). EFSA has recommended that the TTC values should be expressed per kg body weight so that they are applicable to different age groups, differing in body weight. It is considered that at oral lifetime exposures below the respective TTC value there is a low probability of any risk, even for a chemical with little or no toxicological data. The EU Joint Research Centre has developed a free software package, ToxTree, to enable Cramer classification of a chemical (http://toxtree.sourceforge.net/, accessed 23/09/2019).

4.2 Initial considerations prior to applying the TTC decision tree

45. Prior to its use, it is important to confirm that the substance of interest is suitable for application of the TTC approach. Literature searches are required to evaluate the level of data available (including using read-across) to perform a risk-assessment. If the group of chemicals within which the substance sits has well-established toxicity data, then the TTC approach should not be used. In addition, substances falling under certain regulations, e.g. EU food/feed legislation, are excluded from use of the TTC where they require submission of toxicity data for approval.

46. During development of the TTC approach, a number of substances were excluded, the rationale for which has been previously well documented (Cramer et al., 1978; Kroes et al., 2004; EFSA Scientific Committee, 2012b). As detailed in paragraph 38, in their most recent guidance, EFSA no longer exclude hydrazines however, organosilicon substances are newly excluded (EFSA and WHO, 2016; EFSA, 2019).

47. Current exclusion categories are: groups of potent genotoxic carcinogens discussed above (for example, aflatoxin-like, azoxy- or *N*-nitroso substances and benzidines), metals in elemental, ionic or organic form, metal-containing compounds, other inorganic compounds, substances known or predicted to bioaccumulate (for example, polyhalogenated-dibenzodioxins, -dibenzofurans and -biphenyls), proteins, substances with a steroid structure, nanomaterials, radioactive substances and organosilicons (EFSA, 2019).

48. The application of the TTC approach to mixtures requires evaluation on a case-by-case basis. Where all components are known, EFSA recommend a tiered approach to risk assessment, with the assumption of dose addition as a starting point. In the case of mixtures that are not dully defined, the TTC approach may be used provided that analysis has shown that excluded compounds are not present. Under these circumstances, the unknown compounds are considered to be potentially DNA reactive and the sum of the mixture components is evaluated against the lowest TTC value (0.0025 µg/kg bw/day).In circumstances where there are no excluded compounds present, there is no concern for unknown components

with regards to DNA reactivity and no organophosphates or carbamates present, the mixture is classed as Cramer Class III.

4.3 Estimates of exposure

49. EFSA recommend that chronic exposure is estimated using the upper end of the distribution range from dietary exposure assessments; where this is unavailable, use of the maximum reported level is suggested. Consideration should be given to subgroups of the population whose dietary exposure may be higher (for example infants and children). In cases of acute exposure (i.e. < 24 h) EFSA advices, where data is available, to use the highest percentile levels in conjunction with high percentile food consumption. If data is unavailable then, as previously, the maximum reported level should be used.

4.4 Application of the TTC decision tree

50. Kroes et al. (2004) developed a decision tree for application of the TTC approach to chemicals in food through combined considerations of structural alerts for genotoxicity with the approach developed by Munro et al. (1996) for *de minimis* exposure values for non-cancer endpoints, based on the structural classification scheme of Cramer et al. (1978). . This scheme was proposed for use by EFSA in 2012 and a revised scheme recommended by EFSA and WHO (2016). The latest version of the EFSA/WHO decision tree, given as part of the most recent EFSA guidance (EFSA, 2019), is shown in Figure 2.

51. In Step 1, the substance is assessed against the exclusion criteria (see paragraphs 38-39). If the TTC approach is valid to use then in Step 2, the potential for DNA-reactive mutagenicity or carcinogenicity is assessed. This is carried out using a weight or evidence approach collating any available experimental data, read across data from structurally-similar substances and the use of structural alerts or (Q)SAR models. These models have been reviewed by the COM.

52. If the weight of evidence derived in Step 2 indicates the substance of interest is a potential DNA-reactive mutagen or carcinogen, then in Step 3 the estimated exposures are applied. If exposure/intake is below the TTC value for DNA-reactive mutagens or carcinogens (i.e. 0.0025 µg/kg bw/day) then it can be ***concluded that there is a low probability of adverse health effects occurring*.** In the case that the TTC value is exceeded, a non-TTC approach (for example, substance-specific risk assessment) is recommended to allow a full assessment of potential adverse health effects.

53. If there is no alert for DNA-reactive mutagenicity or carcinogenicity, in Steps 4 and 5 the substance of interest is evaluated to see if it is an organophosphate or carbamate. If so, and the estimated exposure is below the TTC value of 0.3 µg/kg bw/day then it can be ***concluded that there is a low probability of adverse health effects occurring*.** In the case that the TTC value is exceeded, a non-TTC approach (for example, substance-specific risk assessment) is recommended to allow a full assessment of potential adverse health effects.

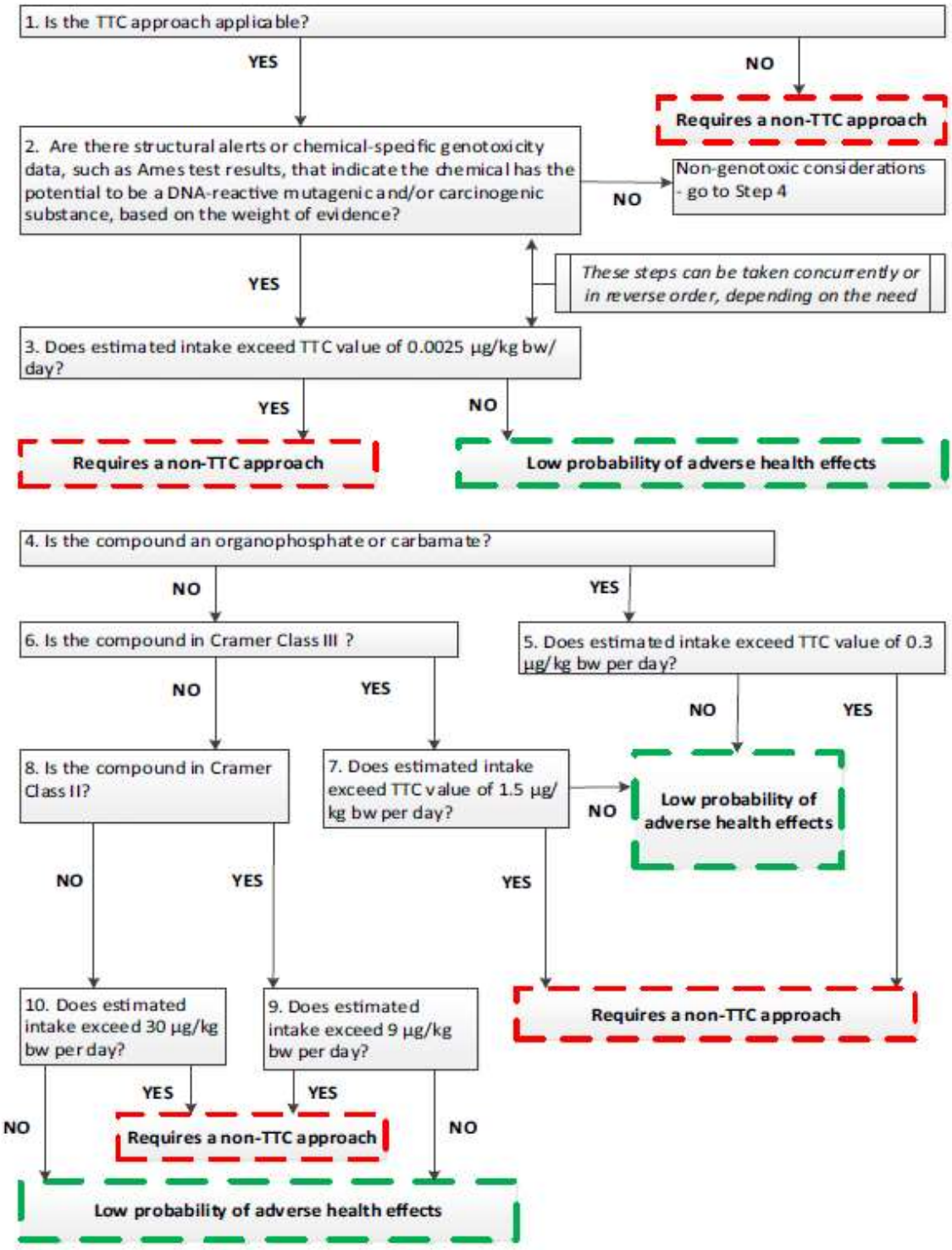

Figure 2: EFSA TTC Decision Tree (EFSA, 2019)

54. If the chemical of interest is not an organophosphate or carbamate then in steps 6 and 7, then it is evaluated against Cramer Class III criteria (see para 37). If the substance belongs to Cramer Class III then the estimated exposures are compared with the TTC value for that Class. If the exposure is below a TTC value of

1.5 µg/kg bw/day then it can be ***concluded that there is a low probability of adverse health effects occurring*.** In the case that the TTC value is exceeded, a non-TTC approach (for example, substance-specific risk assessment) is recommended to allow a full assessment of potential adverse health effects.

55. If substances do not belong to Crammer Class III then they are evaluated against Crammer Class II criteria (see para 37) in steps 8 and 9. If the estimated exposure is below the TTC value of 9 µg/kg bw/day it can be ***concluded that there is a low probability of adverse health effects occurring*.** In the case that the TTC value is exceeded, a non-TTC approach (for example, substance-specific risk assessment) is recommended to allow a full assessment of potential adverse health effects.

56. Remaining substances can be classed as belonging to Crammer Class I and in step 10 the estimated exposures are compared to the TTC value for that Class. If the TTC value of 30 µg/kg bw/day is not exceeded, then it can be ***concluded that there is a low probability of adverse health effects occurring*.** In the case that the TTC value is exceeded, a non-TTC approach (for example, substance-specific risk assessment) is recommended to allow a full assessment of potential adverse health effects.**Error! Hyperlink reference not valid.**

4.5 Special considerations in applying the TTC decision tree

57. Exposure estimates in infants under the age of 16 weeks require additional considerations to be applied and these have been discussed fully by EFSA (EFSA, 2017). In addition, differences in dietary exposure and reaction to certain substances in the diet between infants, children and adults are possible and have also been discussed fully by EFSA (EFSA, 2019).

4.6 Regulatory use of the TTC approach

58. In 2009, Felter et al. proposed further refinements to the TTC decision tree, including consideration for chemicals that have structural alerts for genotoxicity but negative data from genotoxicity tests. They proposed using a higher threshold value of 1.5 µg/person/day as an appropriate TTC exposure limit in such cases. This was based on the work by Cheeseman et al. (1999) on carcinogenic potency and results from Ames tests. This paper also suggested that in circumstances where exposures were unlikely to be over a lifetime, a value of 1.5 µg/person/day may be appropriate for exposures which will not be longer than 1 year (Felter et al., 2009). The concept of a staged TTC was proposed by Müller et al. (2006) and takes into account the fact that duration of exposure is a key factor impacting on the probability of a carcinogenic response. In 2010, the European Medicines Agency (EMA) agreed to the use of a staged TTC approach during clinical development of medicines for a less than lifetime exposure and recommended limits for daily intake of genotoxic impurities (GTIs) of 1.5, 5, 10, 20 and 60 µg/day for greater than 12 months, 6-12 months, 3-6 months, 1-3 months and less than 1 month, respectively. For single doses, an intake of 120 µg/day was agreed to be acceptable (EMA, 2010).

59. TTC values derived from the Cramer et al. classes are used by EFSA and the Joint FAO/WHO Expert Committee on Food Additives (JECFA) for assessing flavouring substances in food (EFSA CEF Panel, 2010). Other uses by EFSA across their remit have included assessments of: impurities, metabolites and degradation products of food additives (EFSA ANS Panel, 2012); pharmacologically active substances present in food of animal origin (EFSA CONTAM Panel, 2018); metabolites and degradation products of plant protection products in the context of residue definition for risk assessment (EFSA PPR Panel, 2016); derivation of 'maximum acceptable feed concentrations' for flavouring additives based on default values for feed consumption (EFSA FEEDAP Panel, 2017); development of the criteria for the safety evaluation of mechanical processes to produce recycled poly(ethylene terephthalate) (PET) intended to be used for manufacture of materials and articles in contact with food (EFSA CEF Panel, 2011).

60. A TTC of 1.5 µg/day is used for different reasons as part of a staged assessment for the acceptability of known genotoxic impurities present in pharmaceuticals (EMA, 2006). This value (1.5 µg/day) is considered appropriate under such circumstances, as a risk of 1 in 10^5 (assuming linear extrapolation) is considered acceptable for human medicines. The use of a TTC of 1.5 µg/day by the EMA applies even to compounds that show evidence of genotoxicity in *in vitro* tests. A similar approach is used for genotoxic constituents of herbal medicinal products/preparations (EMA, 2008).

61. The TTC approach has also been proposed for use with assessing household and personal care products (Blackburn et al., 2005), skin sensitising substances (Safford, 2008) and for industrial chemicals assessed under REACH (ECHA, 2008).

4.7 TTC endorsement by sister committees

62. The COM published a statement in April 2012 on the genotoxicity testing and hazard assessment of impurities. As part of this, Members agreed that the TTC was a useful concept in identifying impurities requiring genotoxicity assessment, although reference needed to be made to the excluded classes of most concern, e.g. aflatoxin-like, azoxy and N-nitroso compounds, which are potent genotoxic carcinogens (COM, 2012).

63. The COC endorses the views of the COM and the views of EFSA and the SCCS, SCHER and SCENIHR Committees on the TTC.

> ***Note to Members for November 2019 meeting:*** *are these endorsements still valid?*
> *Is there a need for the current level of detail on TTC to be included in G05?*

5.0 New fields and Developments in Deriving Points of Departure

5.1 The Signal-to-Noise Crossover Dose (SNCD) approach

64. Sand et al. (2011) developed a new approach for derivation of a POD based on the concept of a signal-to-noise crossover dose (SNCD) and compared it with other methods for deriving the POD. The SNCD provides an estimate of the lowest dose that can be derived as a POD for risk assessment without low-dose extrapolation. It is defined as the dose at which the additional risk equals the 'background noise' or a specified fraction thereof. Background noise is defined as the difference between the upper and lower bounds of the two-sided 90% confidence interval (CI) on absolute risk. Sand et al. (2011) concluded their comparison of the different methods by noting that, if the standard BMD approach is used, then the $BMDL_{10}$ is the most appropriate POD and that the SNCD should be developed further. Responding to the new SNCD approach, Chiu et al. (2012) proposed augmenting the statistical approach for human risk assessment by additional steps so that inter- and intra-species differences and other biological considerations relating to the key end points are addressed.

65. The SNCD approach gives equivalence with the $BMDL_{10}$ approach using a default uncertainty factor of 100. The SNCD-based exposure guideline was derived by linear extrapolation from the upper bound on extra risk at the SNCD (UERSNCD) down to a target risk of 1 in 10^3. However, it should be noted that for a genotoxic carcinogen it is likely that target risk values would be appreciably lower than this (typically 1 in 10^5 or 1 in 10^6). The Committee will continue to keep a watching brief on the developments of the SNCD approach as an alternative approach to deriving a POD but notes for this update that there have been no further publications on this methodology since 2011.

> ***Note to Members for November 2019 meeting:*** *No further use of this in the literature has been identified – can Members provide examples or references of its use or should this section be removed?*

6.0 Summary

66. The Committee recommends the use of the BMDL as the POD for all carcinogens. For genotoxic carcinogens, the likeliest use of the BMDL would be to calculate a MOE as outlined in Guidance Statement G06. For non-genotoxic carcinogens, the BMDL can be used to establish guideline values such as TDI/ADI using uncertainty factors, if carcinogenicity is the critical endpoint. If a BMDL cannot be set for a chemical, the Committee agrees that, although it might be possible to derive a T25 from the dataset, this is not recommended. Instead a NOAEL can be adopted for non-genotoxic compounds, and even for genotoxic compounds noting that this should be used in a way that doesn't imply the existence of a threshold for effect.

67. The Committee is of the view that potency estimates can be of pragmatic use in the risk assessment of carcinogenicity as an aid to prioritising carcinogenic substances (e.g. for risk re-evaluation) but considers that such potency estimates do not provide a quantitative estimate of risk. Although potency estimates can be used to rank chemicals within a particular group (such as structurally related groups of putative genotoxic chemicals), extrapolating from high to low dose and from animals to humans introduces sources of uncertainty.

68. The Committee recognises that the TTC approach provides a pragmatic means of assessing whether exposure to a chemical is of low concern or whether further testing is required. However, the Committee reiterates that the TTC is not a replacement for data on any chemical under consideration but could be used where data are lacking or insufficient, to help in reaching a decision.

COC Guidance Statement G05 v2.0 (first draft)
Date TBC

7.0 References

Ashby J, Tennant RW (1991). Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the U.S. NTP. Mutat Res. 257: 229–306.

Barlow S, Renwick AG, Kleiner J, Bridges JW, Busk L, Dybing E, Edler L, Eisenbrand G, Fink-Gremmels J, Knaap A, Kroes R, Liem D, Müller DJG, Page S, Rolland V, Schlatter J, Tritscher A, Tueting W, Würtzen G, (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. 44: 1636–1650.

Benford D, Bolger PM, Carthew P, Coulet M, DiNovi M, Leblanc J-C, Renwick AG, Setzer RW, Schlatter J, Smith B, Slob W, Williams G, Wildemann T (2010). Application of the Margin of Exposure (MOE) approach to substances in food that are genotoxic and carcinogenic. Food Chem Toxicol. 48 Suppl 1: S2–S24.

Blackburn K, Stickney JA, Carlson-Lynch SL, McGinnis PM, Chappell L, Felter SP (2005). Application of the threshold of toxic concern approach to ingredients in personal and household care products. Reg Toxicol Pharmacol. 43: 249-259.

Carthew P, Clapp C, Gutsell S (2009). Exposure based waiving: the application of the toxicological threshold of concern (TTC) to inhalation exposure for aerosol ingredients in consumer products. Food Chem Toxicol. 47: 1287-1295.

Cheeseman MA, Machuga EJ, Bailey AB (1999). A tiered approach to threshold of regulation. Food Chem Toxicol. 37: 387-412.

Chiu WA, Guyton KZ, Hogan K, Jinot J (2012). Approaches to human health risk assessment based on the signal-to-noise crossover dose. Environ Health Perspect. 120: A264.

COC (2004). Guidance on a Strategy for the Risk Assessment of Chemical Carcinogens. Available from: http://webarchive.nationalarchives.gov.uk/20140506122055/http://www.iacoc.org.uk/publications/index.htm (accessed 29/09/2019)

COC (2006). Acute T25 – Possible approach to potency ranking of single exposure genotoxic carcinogens. In COT, COM and COC Joint Annual Report 2006, p279. Available from: https://cot.food.gov.uk/cotreports/cotcomcocrep2006 (accessed 29/09/2018).

COM (2012). Interim Guidance on a Strategy for Genotoxicity Testing and Mutagenic Hazard Assessment of Impurities in Chemical Substances. Available from https://www.gov.uk/government/publications/genotoxicity-assessment-of-impurities-in-chemical-substances (accessed 29/09/2019).

Collins JF, Brown JP, Alexeeff GV, Salmon AG (1998). Potency equivalency factors for some polycyclic aromatic hydrocarbons and polycyclic aromatic hydrocarbon derivatives. Regul Toxicol Pharmacol. 28: 45-54.

Cramer GM, Ford RA, Hall RL (1978). Estimation of toxic hazard – a decision tree approach. Food Cosmet Toxicol. 16: 255-276.

Crump KS (1984). An improved procedure for low-dose carcinogenic risk assessment from animal data. J Environ Pathol Toxicol Oncol. 5: 339–348.

Crump K, Viren J, Silvers A, Clewell H, Gearhart J, Shipp A (1995). Reanalysis of dose-response data from the Iraqi methylmercury poisoning episode. Risk Anal. 15: 523–532.

Davis JA, Gift JS, Zhao QJ (2011). Introduction to benchmark dose methods and U.S. EPA's benchmark dose software (BMDS) version 2.1.1. Toxicol Appl Pharmacol. 254:181-191.

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: 110–117.

Dybing E, Sanner T, Roelfzema H, Kroese D, Tennant RW (1997). T25: a simplified carcinogenic potency index: description of the system and study of correlations between carcinogenic potency and species/site specificity and mutagenicity. Pharmacol Toxicol. 80: 272-279.

EC (1999). Commission Working Group on the classification and labelling of dangerous substances. Guidelines for setting specific concentration limits for carcinogens in Annex 1 of Directive 67/548/EEC. Inclusion of Potency Considerations. http://ec.europa.eu/environment/archives/dansub/pdfs/potency.pdf (accessed 29/09/18).

ECETOC (2002). The Use of T25 Estimates and Alternative Methods in the Regulatory Risk Assessment of Non-threshold Carcinogens in the European Union. Technical Report No. 83, ISSN-0773-8072-83, Brussels, January 2002. http://www.ecetoc.org/wp-content/uploads/2014/08/ECETOC-TR-083.pdf (accessed 29/09/18).

ECHA (2008). Guidance on information requirements and chemical safety assessment; Endpoint specific guidance. Appendix R.7-1 Threshold of Toxicological Concern (TTC) – a concept in toxicological and environmental risk assessment. Chapter R.7.

EFSA (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. EFSA Journal 282, 1–31.

EFSA (2009). Guidance of the Scientific Committee on use of the benchmark dose approach in risk assessment. The EFSA Journal 1150, 40-47.

EFSA (2011). Use of BMDS and PROAST software packages by EFSA Scientific Panels and Units for applying the Benchmark Dose (BMD) approach in risk assessment. EN-113. [190 pp].

EFSA (2012a). Scientific Opinion on Evaluation of the Toxicological Relevance of Pesticide Metabolites for Dietary Risk Assessment. EFSA Panel on Plant Protection Products and their Residues (PPR). EFSA Journal 10, 2799.

EFSA (2012b). Scientific Opinion on Exploring options for providing advice about possible human health risks based on the concept of Threshold of Toxicological Concern (TTC). EFSA Journal 10 (7): 2750 [103 pp].

EFSA (2017). EFSA Scientific Committee, Hardy A, Benford D, Halldorsson T, et al. Update: Guidance on the use of the benchmark dose approach in risk assessment. EFSA Journal; 15(1): 4658, 41 pp.

EFSA (2019). Scientific Committee, Moore SJ, Bampidis V, Benford D et al. Guidance on the use of the Threshold of Toxicological Concern approach in food safety assessment. EFSA Journal; 17(6):5708, 17 pp.

EFSA (European Food Safety Authority) and WHO (World Health Organization) (2016). Review of the Threshold of Toxicological Concern (TTC) approach and development of new TTC decision tree. EFSA supporting publication 2016: EN-1006. 50 pp. https://www.efsa.europa.eu/en/supporting/pub/1006e (accessed 29/09/19).

EFSA ANS Panel (EFSA Panel on Food Additives and Nutrient Sources added to Food), 2012. Guidance for submission for food additive evaluations. EFSA Journal 2012;10(7):2760, 60 pp.

EFSA CEF Panel (EFSA Panel on Food Contact Materials E, Flavourings and Processing Aids), 2010. Guidance on the data required for the risk assessment of flavourings to be used in or on foods. EFSA Journal 2010;8 (6):1623, 38 pp.

EFSA CEF Panel (EFSA Panel on Food Contact Materials E, Flavourings and Processing Aids), 2011. Scientific Opinion on the criteria to be used for safety evaluation of a mechanical recycling process to produce recycled PET intended to be used for manufacture of materials and articles in contact with food. EFSA Journal 2011;9 (7):2184, 25 pp.

EFSA CONTAM Panel (EFSA Panel on Contaminants in the Food Chain), Knutsen HK, Alexander J, Barregard L, et al. 2018. Update: methodological principles and scientific methods to be taken into account when establishing Reference Points for Action (RPAs) for nonallowed pharmacologically active substances present in food of animal origin. EFSA Journal 2018;16(7):5332, 25 pp.

EFSA FEEDAP Panel (EFSA Panel on Additives and Products or Substances used in Animal Feed), Rychen G, Aquilina G, Azimonti G, et al. 2017. Guidance on the assessment of the safety of feed additives for the target species. EFSA Journal 2017;15(10):5021, 19 pp.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2016. Guidance on the establishment of the residue definition for dietary risk assessment. EFSA Journal 2016;14(12):4549, 129 pp.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), Ockleford C, Adriaanse P, Hougaard Bennekou S, et al. 2018. Scientific opinion on pesticides in foods for infants and young children. EFSA Journal 2018;16(6):5286, 75 pp.

EMA (2006). Committee for Medicinal Products for Human Use. Guideline on the Limits of Genotoxic Impurities, CPMP/SWP/5199/02, EMEA/CHMP/QWP/251344/2006. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002903.pdf (accessed 11/05/17).

EMA (2008). Committee on Herbal Medicinal Products (HMPC). Guideline on the assessment of genotoxicity of herbal substances/preparations. EMEA/HMPC/107079/2007.

EMA (2010). Committee for Medicinal Products for Human Use (CHMP) Safety Working Party (SWP). Questions and answers on the 'Guideline on the limits of genotoxic impurities'. EMA/CHMP/SWP/431994/2007. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002907.pdf (accessed 11/05/17).

Escher SE, Tluczkiewicz I, Batke M, Bitsch A, Melber C, Kroese ED, Buist HE, Mangelsdorf I (2010). Evaluation of inhalation TTC values with the database RepDose. Regul Toxicol Pharmacol. 58: 259-274.

FDA (1995). Food Additives: Threshold of Regulation of Substances Used in Food-Contact Articles; Final Rule. Federal Register, 60, 36582–36596.

Felter S, Lane RW, Latulippe ME, Llewellyn GC, Olin SS, Scimeca JA, Trautman TD (2009). Refining the threshold of toxicological concern (TTC) for risk prioritization of trace chemicals in food. Food Chem Toxicol. 47: 2236-2245.

Frawley JP (1967). Scientific evidence and common sense as a basis for food-packaging regulations. Food Cosmet Toxicol. 5: 293-308.

Gillespie Z, Pulido O, Vavasour E (2011). Risk Assessment Approaches for Carcinogenic Food Contaminants. International Food Risk Analysis Journal 1: 1-18.

Gold LS, Sawyer CB, Magaw R, Backman GM, de Veciana M, Levinson R, Hooper NK, Havender WR, Bernstein L, Peto R, Pike MC, Ames BN (1984). A carcinogenic potency database of the standardized results of animal bioassays. Environ Health Perspect, 58: 9-319.

Gold LS, Slone TH, Manley NB, Garfinkel GB, Rohrbach L, Ames BN (1997). Carcinogenic Potency Database. In: *Handbook of Carcinogenic Potency and*

Genotoxicity Databases (L. S. Gold & E. Zeiger, eds.), Boca Raton, FL: CRC Press, pp. 1-605.

Haber LT, Dourson ML, Allen BC, Hertzberg RC, Parker A, Vincent MJ, Maier A, Boobis AR (2018) Benchmark dose (BMD) modeling: current practice, issues and challenges. Crit Rev Toxicol 48: 387-415.

IPCS (2009a). Principles and Methods for the Risk Assessment of Chemicals in Food. Geneva, World health Organization, International Programme on Chemical Safety (Environmental Health Criteria 240). http://www.inchem.org/documents/ehc/ehc/ehc240_index.htm (accessed 08/05/17).

IPCS (2009b). Principles for Modelling Dose-Response for the Risk Assessment of Chemicals. Geneva, World Health Organization, International Programme on Chemical Safety (Environmental Health Criteria 239).

Kroes R, Renwick AG, Cheeseman M, Kleiner J, Mangelsdorf I, Piersma A, Schilter B, Schlatter J, van Schothorst F, Vos JG, Würtzen G; European branch of the International Life Sciences Institute (2004). Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. Food Chem Toxicol. 42: 65-83.

Müller L, Mauthe RJ, Riley CM, Andino MM, De Antonis D, Beels C, DeGeorge J, De Knaep AKM, Ellison D, Fagerland JM, Frank R, Fritschel B, Galloway S, Harpur E, Humfrey CDN, Jacks AS, Jagota N, MacKinnon J, Mohan G, Ness DK et al. (2006). A rationale for determining, testing, and controlling specific impurities in pharmaceuticals that possess potential for genotoxicity. Regul Toxicol Pharmacol. 44: 198-211.

Munro IC, Ford RA, Kennepohl E, Sprenger JG (1996). Correlation of a structural class with no-observed-effect levels: a proposal for establishing a threshold of concern. Food Chem Toxicol. 34: 829–867.

Peto R, Pike MC, Bernstein L., Gold LS, Ames BN (1984). The TD50: a proposed general convention for the numerical description of the carcinogenic potency of chemicals in chronic-exposure animal experiments. Environ. Health Perspect. 58: 1-8.

PHE (2017). Contaminated Land Information Sheet. Risk Assessment Approaches for Polycyclic Aromatic Hydrocarbons (PAHs). https://www.gov.uk/government/publications/contaminated-land-risk-assessment-approaches-for-pahs (accessed 25/10/19).

Pufulete M, Battershill J, Boobis A, Fielder R (2004). Approaches to carcinogenic risk assessment for polycyclic aromatic hydrocarbons: a UK perspective. Regul Toxicol Pharmacol. 40: 54-66.

Roberts DW, Api AM, Safford RJ, Lalko JF (2015). Principles for identification of High Potency Category Chemicals for which the Dermal Sensitisation Threshold (DST) approach should not be applied. Regul Toxicol Pharmacol. 72: 683-93.

Rulis AM (1986). *De minimis* and the threshold of regulation. C.W. Felix (Ed.), Food Protection Technology, Chelsea/Lewis, Michigan (1986), pp. 29–37.

Safford RJ (2008). The Dermal Sensitisation Threshold - A TTC approach for allergic contact dermatitis. Regul Toxicol Pharmacol 51: 195-200.

Safford RJ, Aptula AO, Gilmour N (2011). Refinement of the Dermal Sensitisation Threshold (DST) approach using a larger dataset and incorporating mechanistic chemistry domains. Regul Toxicol Pharmacol 60: 218-224.

Safford RJ, Api AM, Roberts DW, Lalko JF (2015). Extension of the Dermal Sensitisation Threshold (DST) approach to incorporate chemicals classified as reactive. Regul Toxicol Pharmacol. 72: 694-701.

Sand S, Portier CJ Krewsk D (2011). A signal-to-noise crossover dose as the point of departure for health risk assessment. Environ Health Perspect. 119: 1766–1774.

Sand S, Victorin K, Filipsson AF (2008). The current state of knowledge on the use of the benchmark dose concept in risk assessment. J Appl Toxicol 28: 405–421.

Sanner T, Dybing E *(2005).* Comparison of carcinogen hazard characterisation based on animal studies and epidemiology. *Basic Clin Pharmacol Toxicol* **96***:* 66–70*.*

Sanner T, Dybing E, Willems MI, Kroese ED (2001). A simple method for quantitative risk assessment of non-threshold carcinogens based on the dose descriptor T25. Pharmacol Toxicol. 88: 331–341.

SCCS/SCHER/SCENIHR (2012). Joint Opinion on the Use of the Threshold of Toxicological Concern (TTC) Approach for Human Safety Assessment of Chemical Substances with focus on Cosmetics and Consumer Products. http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_o_092.pdf (accessed 11/05/17).

Tluczkiewicz I, Kühne R, Ebert RU, Batke M, Schüürmann G, Mangelsdorf I, Escher SE (2016) Inhalation TTC values: A new integrative grouping approach considering structural, toxicological and mechanistic features. Regul Toxicol Pharmacol. 78: 8-23.

US EPA (1995). The use of the benchmark dose approach in health risk assessment. EPA/ 630/R-94/007. Risk Assessment Forum, Washington DC.

US EPA (2000). Benchmark Dose Technical Guidance Document (BMDTG). External review draft. EPA/630/R-00/001. Risk Assessment Forum, U.S. Environmental Protection Agency: Washington, DC.

US EPA (2010). Development of a Relative Potency Factor (RPF) Approach for Polycyclic Aromatic Hydrocarbon (PAH) Mixtures (External Review Draft). U.S. Environmental Protection Agency, Washington, DC, EPA/635/R-08/012A.

US EPA (2012). Benchmark Dose Technical Guidance Document (BMD TG). External review draft. EPA/100/R-12/100. Risk Assessment Forum, U.S. Environmental Protection Agency: Washington, DC. https://www.epa.gov/risk/benchmark-dose-technical-guidance (accessed 08/05/17).

Wheeler MW, Bailer AJ (2007). Properties of model-averaged BMDLs: a study of model averaging in dichotomous response risk estimation. Risk Anal 27: 659–670.

Wheeler MW, Bailer AJ, (2008). Model averaging software for dichotomous dose–response risk estimation. Journal of Statistical Software Vol 26 Issue 5: 1–15.