INTERPRETATION OF MARGINS OF EXPOSURE FOR GENOTOXIC CARCINOGENS

FINAL REPORT

Project number: T01051

Duration: 01 Jan 2010 - 30 April 2013

Participants:

Imperial College London
Alan R Boobis (PI)
Lesley Rushton
Ehi Idahosa-Taylor (until 19/12/2012)
DEFRA, Central Science Laboratory,
Andy Hart
John-Paul Gosling (until August 2011)
Villie Flari (from September 2009)
University of Durham
Peter Craig

Date of report: 09 August 2013, revised 28 November 2013

1. Introduction

Many chemicals cause cancer by a mode of action involving genotoxicity, in which an early key event is mutation of a critical oncogene. As cancer is a disease of clonal (single cell) origin, it has been argued that a single mutation in a single cell is sufficient to give rise to cancer and, hence, there would be no threshold exposure for such a compound, below which there is no mutation and therefore no cancer. Until relatively recently, risk assessment of such compounds took one of two forms. Either (1) risk assessment stops with identification of the hazard: the mode of action for the carcinogenicity of the compound could reasonably be a consequence of its genotoxicity, or (2) the experimental data on carcinogenicity are used to extrapolate to an exposure associated with a risk considered to be of low or negligible concern, often 1 in 10^5 or 1 in 10^6. This exposure has been termed the "virtually safe dose".

In (1), such a conclusion often results in risk management action of reducing exposure to levels that are "as low as reasonably achievable/practicable" (ALARA/P). However, the conclusion of such a risk assessment does not provide any information on which to decide how urgent or extensive risk management action is necessary. Similarly, it is not possible to prioritise competing hazards for action (EFSA, 2005; O'Brien et al., 2006; Benford et al., 2010).

In (2), agreement is needed on what assumptions should be made regarding the nature of the dose-response relationship below the range of experimental observations. In the most recent update to the Cancer Guidelines of the US EPA, it was concluded that the shape of the dose-response curve at human relevant exposures is not known and, hence, a plausible worst case would be to assume a linear relationship between the point of departure from the experimental data to human exposure levels (EPA, 2005).

Approach (2), sometimes known as low dose linear extrapolation, has been criticised due to the considerable uncertainty about the shape of the dose-response curve at human relevant exposures (Williams et al., 2005; EFSA, 2005) and is not supported by the Committee on Carcinogenicity. It also requires a policy decision on what level of excess risk is considered to be of low or negligible concern at the virtually safe dose, which can lead to difficulties in risk communication. To address these concerns, whilst enabling risk assessors to provide information that would help in risk management decisions regarding the overall level of concern for a carcinogen and in the prioritisation of competing hazards, several authoritative bodies have recommended use of the margin of exposure (MOE). The MOE is the ratio of the point of departure (POD), typically the Benchmark Dose – Lower Confidence Limit ($BMDL_{10}$) for a tumourigenic response in experimental animals, to the estimated human exposure for a genotoxic carcinogen. Exposure estimates often use plausible worst-case assumptions regarding likely routes, e.g. oral, inhalation, from anticipated uses and/or environmental levels. Values of MOE equal to or greater than 10,000 have been considered to indicate "low concern", although different organisations (e.g. European Food Safety Authority; Committee on Carcinogenicity/Committee on Mutagenicity; JEFCA – Joint FAO/WHO Expert Committee on Food Additives) may express the risk level differently.

Interpretation of the MOE requires consideration of the uncertainty and variability that underlie inter- and intra-species differences in carcinogenicity. Additional factors that might need to be taken into account include human variability in cell cycle control and DNA repair and the fact that the point of departure is not a NOAEL (EFSA, 2005). At a joint EFSA/WHO/ILSI conference to discuss the MOE approach for genotoxic carcinogens, there

was some agreement that a factor of 100 for inter- and intra-species uncertainty and variability is scientifically justifiable. However, there was little agreement on the basis of additional factors (Barlow et al, 2006).

Whilst these discussions focused on interpreting the level of concern for an MOE of 10,000 on the basis of safety/uncertainty factors, there was also recognition that the magnitude of the potential risk at "acceptable" exposures needs to be considered. EFSA's proposal that an MOE of 10,000 or above should be considered of low concern received a mixed response. Some had reservations about the rationale for this figure (Barlow, 2006). It was agreed that further discussion was necessary on whether and how to assign levels of concern to particular values or ranges of MOEs. As a follow-up activity, MOEs were calculated for twelve different chemicals, as case studies (Benford et al. 2010). Although these were discussed at a workshop held for this purpose in October 2008, no conclusions could be reached on whether or how to define levels of concern.

The MOE approach clearly has potential advantages, including practicality, avoidance of assumptions needed for low-dose extrapolation, avoidance of providing numerical risk estimates considered of low concern and easier risk communication. MOEs can be useful for ranking the possible risks from different chemicals or exposure situations (COT , 2007). The COC (2007) has also agreed the value of using the MOE approach in a number of situations, whilst emphasising that this does not replace ALARA/P. However, there is still no agreement on the scientific rationale for the derivation and interpretation of the level of concern for an MOE of $\geq$ 10,000 (Benford, 2010).

It would be of considerable practical value to develop a robust scientific rationale for defining levels of concern for the MOE, and to achieve consensus on this. This would improve the usefulness of MOEs to risk managers in helping to decide whether and to what extent action is needed on individual chemicals. This current project was undertaken to try to develop a robust scientific rationale for defining levels of concern associated with given values or ranges of MOEs. This was investigated by systematically reviewing evidence and eliciting expert opinion on dose-response relationships for genotoxic carcinogens, developing a statistical framework to inform the definition of levels of concern, and comparing the results with data on a number of known or potential human carcinogens.

2. Objective 01: Critical review of existing proposals for a level of concern and alternative approaches for interpreting Margins of Exposure for genotoxic carcinogens

The full report for Objective 01 is provided as Annex 1.

A systematic search was undertaken to retrieve publications in both the peer reviewed and "grey" literature that might be relevant to addressing this topic. A search of PubMed identified 31 relevant papers and a search of Google revealed a further 8 relevant sources. The search covered the period until December 2009. Full details of the search strategy and results can be found in the attached report for this Objective.

Some authorities have proposed a single value for the margin of exposure (MOE) for genotoxic carcinogens, to indicate the level of concern (LOC), whilst others have proposed using bands based on a series of defined MOE values, representing differing levels of concern. Both these approaches could be used as a basis for prioritising chemicals for risk management consideration. There have also been instances where case-by-case interpretation of the LOC associated with an MOE has been advocated, to take into account the various sources of uncertainty.

EFSA (2005) published an opinion on interpretation of the MOE for genotoxic carcinogens. It was concluded that an MOE above 10,000, calculated as they proposed, would indicate a low priority for risk management action if the MOE was not associated with an unreasonable degree of uncertainty. EFSA emphasised that this interpretation of the MOE applied only if the MOE was calculated using the lower 95% confidence interval for the benchmark dose ($BMDL_{10}$), corresponding to a 10% response above background in studies in experimental animals, as the point of departure (POD) or reference point (RP).

EFSA provided some discussion for their conclusion that those genotoxic carcinogens with an MOE above 10000 were of low concern. In addition to the normal default values for inter- and intra-species differences, additional factors of 10 for interindividual variability in the carcinogenic process itself and for use of a $BMDL_{10}$ rather than a NOAEL as POD were proposed. However, in a subsequent meeting organised by WHO, EFSA and ILSI Europe, it was concluded that the scientific rationale for these additional factors was weak and whilst a value of 10,000 was endorsed, above which there would be a low level of concern, it was not possible to provide a strong scientific rationale for choice of this value.

The issue of providing advice on substances that are genotoxic and carcinogenic was also addressed at the 64th meeting of the Joint FAO/WHO Expert Committee on Food Additives (JECFA, 2005). Similar conclusions were reached to those of EFSA. Although no explicit recommendations were made regarding an MOE above which there would be low concern, from the assessments provided and conclusions reached, it was implicit that an MOE of above 10,000, when derived from a $BMDL_{10}$ for a carcinogenic response in experimental animals, was considered to be of low concern.

The UK Committee on Carcinogenicity reviewed the recommendations of EFSA and JECFA regarding the use of the MOE as a means of prioritising genotoxic carcinogens for risk management action and for communicating relative levels of concern. The COC concluded that the MOE could be useful for these purposes. The Committee further concluded that risk

communication might be improved by providing textual descriptors for different MOE bands. Thus, MOEs less than 10,000 were considered to be of possible concern, those between10,000 – 1,000,000 were considered unlikely to be of concern and those above 1,000,000 were considered highly unlikely to be of concern. The basis for these bounds was not clearly explained. The COC provided three case studies to illustrate the use of the proposed banding. These were the soil contaminants chromium VI, benzo[a]pyrene (B[a]P) and 1,2-dichloroethane (COC, 2007).

As a follow up activity to the recommendations of EFSA and JECFA, ILSI Europe convened an expert group to explore the application of the MOE approach to 12 case studies. The findings were discussed at a workshop organised for this purpose and then published in a series of articles in *Food and Chemical Toxicology* (2010, Vol. 48, Suppl 1). It was concluded that there are a number of issues that need to be considered in interpreting the MOE, and that the numerical value should be interpreted on a case-by-case basis, taking into account the attendant uncertainties. Amongst the key areas highlighted were: mode of action for the carcinogenic response, selection of tumour data, dose-response modelling, selection of POD and exposure assessment.

The literature search undertaken for this project revealed 9 examples of where the MOE approach had been used in the risk assessment of genotoxic carcinogens. These were published between 2006, following the recommendations by authoritative bodies on its use, and 2009, when the literature search was conducted. Lachneimer et al (2009a, b, c, d) published a series of papers on the MOE for low molecular weight compounds found in alcoholic beverages and other products. In the case of acetaldehyde in mouthwash, the MOE above which it was considered that there was a low level of concern was 30,000, the authors including an additional factor of 3 to allow for possible increased sensitivity of children. In the case of acetaldehyde in alcohol beverages, an additional factor of 10 was included to allow for those who are genetically deficient in the enzyme acetaldehyde dehydrogenase. The resulting MOE for exposure to acetaldehyde from alcoholic beverages (~500) was such that it gave rise to concern. The authors also determined the MOE for ethyl carbamate from consumption of certain alcoholic beverages. The MOE was used to determine how many drinks per day would give rise to concern, i.e. where consumption was such that it would give rise to an MOE less than 10,000. The MOE for furan in commercial baby food was determined. The authors used a published value for the T25 to calculate the MOE. Whilst recognising that interpretation of an MOE based on the T25 rather than the $BMDL_{10}$ introduced greater uncertainty, the authors concluded that the values obtained were such that there was reason for concern, particularly as an additional factor of 10 might be necessary when assessing the risk to infants.

Akpambang et al (2009) used the MOE approach to determine the relative risk from exposure to benzo[a]pyrene through the consumption of different types of smoked fish in Nigeria. It was concluded that the MOE for some types of fish was such that there was need for regulatory limits on PAH levels in commercially smoked fish.

Wang et al (2009) used the MOE approach to assess relative risks from exposure to aflatoxin B1 in different parts of China. Based on the MOEs obtained, it was concluded that the level of concern was moderate for those with average exposure and high for those with high exposure.

The MOE approach was used by Schuetze et al (2008) to determine the level of concern for exposure to malachite green through consumption of fresh-caught eels in Germany. The MOEs for worst-case scenarios were all well above 10,000 but the authors emphasised their view that exposure to any genotoxic carcinogen via food should be avoided.

Zeilmaker et al (2010) used the MOE approach to assess the risk from exposure to N-nitrosodimethylamine (dimethylnitrosamine, DMN) via the diet. The assessment was performed for children of 1 year of age and adults of > 25 years. The authors included a detailed consideration of the interpretation of the MOEs obtained. This included evaluation of the relative exposure of young and adult rats, which mirrored that in humans. It was the view of the authors that this could be taken into account by probabilistic modelling. A further consideration was time-to-tumour. This revealed that the differential impact of exposure of children on cancer risk relative to that of adults was negligible. The authors also argued that interspecies differences were low and that the results of their probabilistic modelling could be used rather than a default factor of 10.

The MOE approach was used by Tardiff et al (2009) to assess the risk from dietary exposure to acrylamide. A number of approaches were used to calculate MOEs, including internal dosimetry and mode of action considerations to select inter- and intra-species adjustment factors. "Conventional" MOEs were also calculated, which were interpreted on the basis of a putative threshold-dependant mode of action, such that there would be little concern at MOEs well below 10,000. Indeed MOEs as low as 50 were considered by these authors to be of low concern for acrylamide.

Since the preparation of the report for Objective 01, there have been a number of additional applications of the MOE approach in risk assessment. EFSA have applied the approach in the assessment of residuals in polymer coatings. In the assessment of polyvinylpyrrolidone-vinyl acetate copolymer, both the vinyl monomer and the hydrazine were assessed and, based on the MOEs derived from the maximum levels in the proposed specification, the EFSA Panel on Food Additives and Nutrient Sources Added to Food (ANS Panel) concluded that these were unlikely to be of safety concern. However, in the case of the hydrazine residual, where MOEs were in the range of 23,000 to 140,000, the Panel recommended that it would be prudent to lower the level of hydrazine as far as reasonably achievable, whereas for vinyl monomer, with MOEs above 1,000,000, no further recommendation was made (EFSA 2010). The ANS Panel recently came to a similar conclusion (i.e. no safety concern at an MOE of > 1,000,000) in its assessment of a PEG-PVA copolymer (EFSA, 2013). The ANS Panel also used the MOE approach to conclude on the safety of PAHs in vegetable carbon (EFSA 2012a). The EFSA Panel on Contaminants in the Food Chain (CONTAM Panel) have recently used the MOE approach on several occasions e.g. pyrrolizidine alkaloids (EFSA, 2011). EFSA (2012b) recently issued a statement endorsing the applicability of the MOE approach in assessing impurities in substances added to food or feed, which are genotoxic and carcinogenic. FSA also confirmed its view that a margin of exposure of $\geq$ 10,000 would be of low concern from a public health point of view.

FSA informed the research group that the MOE approach had also been used in the UK in the risk assessment of compounds which are genotoxic and carcinogenic in food incidents since 2009 (FSA, 2011) e.g. morpholine (FSA, 2010). The values obtained have provided a basis for decisions on risk management options and communicating the risk to the public (Gott, personal communication).

Since the closing date of literature retrieval for the Objective 1 review, a number of additional publications have appeared on the use of the margin of exposure approach in assessing the level of concern for genotoxic carcinogens in food and consumer products (as of 31 July 2013). These include aflatoxin in the diet of Brazilians (Andrade et al, 2013), dietary exposure to 3-chloropropane-1,2-diol and related compounds (Andres et al, 2013; Bakhiya et al, 2011), natural occurring genotoxic carcinogens present in botanical preparations, used as pesticides (Hernández-Moreno et al, 2013) or as food supplements (van den Berg et al, 2011a, b), benzene (Lachenmeier et al, 2010) and furan (Lachenmeier et al, 2012a) in infant food, furan in other foodstuffs such as coffee (Moro et al, 2012 ; Waizenegger et al, 2011) ethanol in alcoholic beverages (Lachenmeier et al, 2011) and mouthwash (Lachenmeier et al, 2013), formaldehyde in alcoholic beverages (Monakhova et al, 2012b), 5-hydroxymethylfurfural in alcoholic beverages (Monakhova et al, 2012a) and the triphenylmethane dye Brilliant Green present in paper towels (Oplatowska et al, 2011). Lachenmeier et al (2012b) have published a study on the use of the MOE approach for the comparative assessment of a number of carcinogens present in alcoholic beverages. These were acetaldehyde, acrylamide, aflatoxins, arsenic, benzene, cadmium, ethanol, ethyl carbamate, formaldehyde, furan, lead, 4-methylimidazole, N-nitrosodimethylamine, ochratoxin A and safrole.

This review highlighted a number of issues in the application of the MOE approach. Some of these were identified by the ILSI Europe Expert Group and have been the subject of follow-up work. Others include the need for a robust scientific basis for interpretation of the MOE and the conclusion that values of above 10,000 are of low concern; what factors should be considered when interpreting the MOE, and how to communicate that in some cases an MOE of less than 10,000 might still be of low concern; interpretation of the MOE when applied to children – when based on data from children, when based on data from adults.

3. Objective 02: Examination of the theoretical basis for how assessment factors should be combined in risk assessment

The full report for Objective 02 is provided as Annex 2.

The objective was to examine the theoretical basis for how assessment factors should be combined in risk assessment and in particular to build an over-arching statistical framework as a rational basis for doing so. The core of the framework presented is the idea that assessment factors are applied to a numerical measurement known as the point of departure (POD) in order to arrive at another number, the point of arrival (POA) used as a basis for risk assessment, which is felt to be sufficiently conservative relevant to some target quantity T which is not measured directly. Often all these quantities are doses of a chemical so that the problem is to arrive at a dose that is sufficiently small that risk is acceptable at or below that dose. For carcinogens, the POD might be an animal $BMDL_{10}$ while T might be the dose that corresponds to a human population lifetime excess cancer risk of 1 in 100,000. In this case, the view that an MOE of $\geq$ 10,000 is acceptable can be seen as equivalent to obtaining the POA by dividing the $BMDL_{10}$ by a total assessment factor of 10,000.

The first issue is what is (or should be) represented by an individual assessment factor (AF). It is clear that an AF may address scaling and/or variability and/or uncertainty and also that a precise judgement is not usually being made, the chosen value being seen instead to be "large enough". For a statistical framework, an AF must have a quantitative probabilistic interpretation and the natural formulation is that AF is chosen large enough such that $P[POA > T] < \gamma$ for some (small) probability γ. Writing U=POD/T, the probability becomes

$$P[U > AF] < \gamma \qquad (1)$$

γ can be thought of as a desired upper limit on the probability that the AF is inadequate. In practice, the value of γ has rarely been specified in risk assessments that use AFs. There is an issue as to what kinds of probabilities are involved here. Our conclusion is that these are essentially judgements of an expert (or group of experts) about uncertainty although the judgements may well be informed by frequencies of occurrence of various phenomena.

In situations where multiple AFs are applied, we argue for thinking of them as applying to a "trajectory": a sequence of *m* steps from the POD to the target T via (unobserved) intermediate "way-points" so that $U = U_1 \times U_2 \times \cdots \times U_m$ where each individual U_i is the ratio of two successive values in the trajectory. In the non-carcinogen context considered later, the POD might be an animal subchronic LOAEL and the ratios might then be:

U_1 = (animal subchronic LOAEL) / (animal subchronic NOAEL)
U_2 = (animal subchronic NOAEL) / (animal chronic NOAEL)
U_3 = (animal chronic NOAEL) / (typical human chronic NOAEL)
U_4 = (typical human chronic NOAEL) / (sensitive human chronic NOAEL)

A trajectory often consists of relatively standard steps for which individual assessment factors (each AF_i relevant to U_i) have already been specified. A common practice to obtain an overall AF for U=POD/T is then to multiply the individual AFs: $AF = AF_1 \times AF_2 \times \cdots \times AF_m$. Assuming that each individual assessment factor satisfies $P[U_i > AF_i] < \gamma_i$, the only probabilistic statement we can make with mathematical certainty, about the result of multiplying individual AFs, is that the probability that U exceeds the overall AF is less than or equal to the sum of the individual γ_i values used in determining the individual AFs. This

may be useful **if**: (i) a value of γ_i was stated when arriving at each individual AF_i; and (ii) those individual γ_i values are sufficiently small that the sum $\gamma_1 + \gamma_2 + \cdots + \gamma_m$ is usefully small.

Mathematically, it is clear that summing the individual γ_i values would often be very conservative, in the sense that (1) would be true but would also be true for a much smaller value of γ than $\gamma_1 + \cdots + \gamma_m$ as otherwise there would have to be very extreme probabilistic dependence between the U_i. We do not immediately gain much just by asserting or assuming probabilistic independence of the individual U_i. In fact, to exploit independence to avoid conservatism, we appear to need much more probabilistic detail: the full distribution describing uncertainty for a chemical about each U_i. But specifying a full distribution removes the one-sidedness in (1); instead, for every possible value of AF there is a known probability γ that U>AF.

Cooke (2000) coined the term "random-chemicals" for approaches that attempt to arrive at individual and overall assessment factors by statistical modelling of empirical data from a sample of chemicals. The idea is to compute the empirical value of an individual ratio U_i for a number of chemicals and then model the variation using a probability distribution; the result represents some of the uncertainty about the ratio for a new chemical. If, as Cooke proposes, we also incorporate uncertainty about the parameters of the distribution (or the family of distributions) in a Bayesian manner, the result can represent many, if not all, uncertainties.

Parts of this approach were tried for non-carcinogens in several articles by Kodell and Gaylor, greatest detail being provided in Gaylor and Kodell (2000), and by Slob and co-authors, e.g. Slob and Pieters (1998). The former use a log-normal distribution to model inter-chemical variability in each individual ratio based on data and, neglecting uncertainty, make a variety of calculations about the conservatism of multiplying standard factors of 10 and the sizes of assessment factors needed to obtain a specific overall value of γ; the latter also informally use distributions to model both variability and uncertainty, performing Monte Carlo calculations which are equivalent to a Bayesian synthesis, but they do not use Bayesian methods to construct such distributions based on data.

The log-normal approach has appeal: it permits limited modelling of inter-chemical variability and of some uncertainties; calculations for AFs are straightforward. However, it cannot adequately address parameter uncertainty or deal with ratios where the empirical distribution is not log-normal. The solution is simple in principle. As proposed by Cooke (2000), we should build a Bayesian statistical model of all the steps in a trajectory and use data from relevant chemicals and knowledge elicited from experts to obtain a posterior distribution for parameters in the model. The model can then be used to compute the predictive distribution of U for a new chemical and the overall AF would then be the relevant percentile of that distribution. Application of this solution is not so simple. For any particular context, it requires detailed statistical modelling and Bayesian computation, neither of which is trivial, but the hard part would be revisiting the evidence relating to each individual ratio.

Annex 2 includes a substantial numerical study of what happens in a simple 4-step trajectory for non-carcinogen risk assessment based on a variety of distributional assumptions, models for dependence and values of γ for individual assessment factors. This extends the work of Gaylor and Kodell (2000), who assumed log-normal variation with no treatment of uncertainty. It shows that the size of overall assessment factor needed varies substantially depending on: (i) the nature of dependence; (ii) the amount to which individual factors

represent scaling rather than pure variability/uncertainty; (iii) distribution shape. Indirectly these show that uncertainty should not be neglected in the random-chemicals approach.

Annex 2 closes with a small case-study, in the context of non-carcinogen risk assessment, based on the example used by Gaylor and Kodell (2000). They looked at the product of some or all assessment factors covering inter-species extrapolation, variability in human sensitivity, prediction of chronic effects from subchronic studies and the ratio of the LOAEL to the NOAEL, when the former is used in place of the latter. They model variability underlying each factor using a log-normal distribution for which the parameters are assumed to be known precisely. The case-study reconsiders their modelling, proposing some different choices, and examines some of the uncertainties involved in their example. It explores how those uncertainties affect: (i) the size of overall assessment factor needed to achieve a given level of credibility γ in the coverage; and (ii) the level of credibility achieved by multiplying individual standard factors of 10.

The case study shows that it is possible to carry out the Cooke approach by careful modelling and thereby to allow for uncertainty about distribution parameters. It shows that there is considerable uncertainty about the overall assessment factor required to cover 95% of chemicals in the specific context considered by Gaylor and Kodell (2000) and confirms their findings, namely that while individual assessment factors of 10 may well not cover 95% of chemicals, the overall factor of 10000 has a high probability of covering 95% of chemicals and that it is unlikely that it covers less than 90%.

4. Objective 03

4.1 Part 1: Critical evaluation of evidence concerning the empirical form of dose-response relationships at low dose exposure to genotoxic carcinogens in animals

The full report for Objective 03, part 1 is provided as Annex 3.

Interpretation of the margin of exposure for genotoxic carcinogens for prioritisation of risks and for assessing the level of concern above and below an MOE of 10,000 requires explicit or implicit appreciation of the nature of the dose-response relationship at exposures below the POD. In some publications in which the level of concern for an MOE of 10,000 has been addressed it has been argued that as the excess risk is 1 in 10^5 when linearity is assumed, this provides some justification for the use of this MOE to establish a low level of concern. Hence, in providing a more robust scientific basis for interpreting the MOE, it would be of value to consider the strength of the evidence for the nature of dose-response relationships below the POD. The primary purpose of this review was therefore to evaluate available evidence on the empirical form of the dose-response relationship for genotoxic carcinogens at low, human-relevant, exposures in experimental studies.

A comprehensive search of the literature was undertaken to retrieve all relevant publications. A number of search strategies were explored, using key papers to test for their accuracy in retrieving the relevant literature. Citation-searching was undertaken within the references from all relevant publications until no further papers were being identified. In addition, a number of experts in the field were consulted and asked to suggest relevant publications. Almost 4,000 papers were identified and screened in this process, from which only 17 relevant publications, covering 10 different chemicals, were identified (full details in accompanying report). It was noted that there were several papers covering different aspects of the same experimental studies, respectively.

The key studies identified were already known to the investigators. These were the ED01 study of 2-AAF in mice, the "mega-rat" study of nitrosamines, overlapping dose-response studies of 2-AAF and DEN in rat by Williams et al, the study in trout of dibenzo[a,l]pyrene and a series of dose-response studies on intermediate endpoints for EMS and ENU.

The ED01, or "mega-mouse", study was conducted by the US National Centre for Toxicological Research. In this study, over 24,000 mice were used to investigate the dose-response relationship for bladder and liver tumours induced by 2-acetylaminofluorene (2-AAF). The study was powered to detect a 1% increase in tumour incidence. Whilst this was 10-fold lower than that normally detectable in a cancer bioassay, it is still 3-4 orders of magnitude above the incidence level considered to be of low concern in humans (1 in 10^5 – 1 in 10^6). The report of the study concluded that the dose-response was essentially linear over the observable range for liver, whereas for bladder the incidence dropped off sharply with dose. It was concluded that in view of the results with liver, it is necessary to use linear extrapolation for the risk assessment of genotoxic carcinogens to ensure sufficient conservatism (Gaylor, 1980).

The data from this study have been re-analysed on a number of occasions. A task force of the US Society of Toxicology (SOT, 1981) concluded that there was evidence for a threshold in

the dose-response for bladder tumours. Purchase & Auton (1995) pointed out that the appearance of a threshold is very dependent on how the data are plotted. They used the data from the ED01 study to emphasise the difficulty of demonstrating thresholds empirically, due to inherent limitations in the statistical power of any feasible study, hence the importance of considering mechanism. In a series of publications, Waddell (2003a) re-plotted the data from the ED01 study using the Rozman scale for dose (molecules/kg bw per day, plotted on a logarithmic scale). This gave a strong appearance for a threshold for both liver and bladder tumours. However, these analyses have been heavily criticised (Crump & Clewell, 2003; Andersen et al, 2003; Lutz, 2003; Haseman, 2003) for a number of reasons, such as the inability to include data from the concurrent controls, the appropriateness of the dose metric, the introduction of compression artefacts and the inability to include all dose groups in the plots. Waddell (2003b) has rejected these criticisms as unfounded.

Gaylor et al (1985) undertook additional pathological evaluation of bladder samples from the ED01 study and performed further statistical analyses of the data. They reaffirmed their view that there was no threshold in the dose-response relationship for bladder tumours induced by 2-AAF but acknowledged that "the shape of the low dose response curve for bladder carcinomas remains uncertain". Poirier et al (1991) conducted DNA analysis of tissue samples from a 28-day study conducted using the same doses as in the ED01 study. They compared levels of dG-C8-AF adducts in liver and bladder with the tumour response in these tissues in the ED01 study after 24 months. Whereas in the liver, there was a linear relationship between adduct levels and tumour response, in the bladder the relationship was non-linear. The authors concluded that this indicated the need for only one "hit" in the liver to produce tumours, but more than one "hit" or event was needed in the bladder.

In the mega-rat study, Peto et al (1991a, b) studied the tumourigenic effects of N-nitrosodimethylamine (DMN) in the liver and N-nitrosodiethylamine (DEN) in the liver and oesophagus in approx. 4000 rats. The background incidence of liver tumours in the rats used in this study was relatively high and precluded any conclusions on the nature of the dose-response at low doses, where the incidence was close to that of background. In order to provide group sizes sufficiently large to enable reliable statistical estimates, data from males and females for both compounds (DMN and DEN) were pooled for the analysis of tumours. With this caveat, over the dose range 0.1 - 1 ppm, the relationship between dose and response in the liver appeared to be linear. The background incidence of tumours in the oesophagus was zero and no such tumours were observed in any animal in the first several dose groups. As a consequence, the relationship between dose and response in the oesophagus, again for the pooled data, appeared to be highly non-linear.

Because of the unfeasibly large number of animals that are required to study dose-response relationships in rodents at response levels below 1% of background, Bailey et al (2009) conducted a mega cancer study in trout, where it was possible to study responses at an order of magnitude lower than in any of the rodent studies conducted to date. In this study, known as the ED001 study, the effects of dibenzo[a,l]pyrene on the incidence of stomach and liver cancer were investigated in over 40,000 rainbow trout (*Oncorhyncucs mykiss*). The observable range of responses was $\geq 0.02\%$ i.e. from 2 in 10^4. However, this is still at least 1-2 orders of magnitude greater than the incidence considered to be low concern in the human population.

Model fitting strongly suggested that neither the liver nor the stomach tumour data were consistent with a linear dose-response relationship. The authors highlighted the fact that the

virtually safe dose (VSD) calculated by linear extrapolation from the BMD_{10} was more conservative by 2-3 orders of magnitude than the VSD calculated by linear extrapolation from the unmodelled ED0002 (response of 0.02%).

Williams et al (1993, 1998, 1999, 2004) have conducted a series of studies exploring the inter-relationships in the dose-response for a number of pre-neoplastic effects in the liver associated with the tumourigenic response to the genotoxic carcinogens DEN and 2-AAF. A number of these pre-neoplastic effects showed non-linear dose-response relationships and for several, no significant difference from background was discernible at low doses of the compounds. It was concluded that linear extrapolation could lead to overestimation of risk and that such extrapolations should be supported by mechanistic data.

Doak et al (2007) and Gocke and Muller (2009) conducted a series of studies on the nature of the dose-response relationships for the genotoxic effects of a number of alkylating agents, methylmethane sulfonate (MMS), methylnitrosourea (MNU), ethylmethane sulfonate (EMS) and ethylnitrosourea (ENU). Studies were conducted both *in vitro* and *in vivo*, in some cases over a wide dose/concentration range. The dose response relationship for MNU and ENU appeared to be linear. In contrast, the response to EMS and MMS was non-linear and there was evidence for a threshold.

The review of the available literature suggests that evidence for the nature of the dose-response relationship for genotoxic carcinogens needs to be evaluated on a case-by-case basis. For some compounds there is reasonable evidence for non-linearity, and that the assumption of linearity will lead to an over-estimation of risk. In other cases, it is not possible on the basis of the available information to dismiss linearity. However, no study had sufficient power to explore dose-response relationships at incidence levels that would be considered to be of low concern in the human population, ≤ 1 in 10^5. Such a study would not be feasible in a vertebrate species, due to the number of animals that would be required and the background incidence of most tumour types. Several groups have evaluated the published data sets for evidence of the existence of threshold in the dose-response relationship for genotoxic carcinogens. In general, it is not possible to establish the presence or absence of a threshold on the basis of empirical observation. Mechanistic studies, in which the dose-response of necessary intermediate events in the mode of action for the carcinogenic response is characterised, are more promising for this purpose. A series of such studies on alkylating agents provided sufficient evidence for a threshold in the dose-response relationship for EMS that this was accepted by regulatory authorities, such as the EMA (2008). However, few other compounds have been studied in such detail.

4.2 Part 2: Expert elicitation of knowledge regarding the form of the dose-response curve for genotoxic carcinogens at low exposures and implications for a level of concern for Margins of Exposure

The full report for Objective 03, part 2 is provided as Annex 4 (published report from the first workshop on expert elicitation) and Annex 5 (report of the second workshop on expert elicitation). It should be noted that further detailed analysis of the findings of the Second Expert Workshop was still underway at the time of submission of this report (28/11/13).

4.2.1 First Expert workshop on low-dose extrapolation of genotoxic carcinogens 23rd – 24th May 2011

Most bodies advocating quantitative risk assessment, to identify a virtually safe dose of a genotoxic carcinogen, now recommend the use of low dose linear extrapolation (for example: USA, NL); the provision of a quantitative risk estimate is one major advantage of low dose linear extrapolation (European Commission, 2009). However, this approach has been criticised due to the considerable uncertainty about the shape of the dose–response curve at human relevant exposures. In addition, the need to choose a risk level considered acceptable at the virtually safe dose (a policy decision) can lead to difficulties in risk communication. To overcome these concerns, several bodies have proposed use of the MOE approach as an alternative. In interpreting the level of concern associated with the margin of exposure, a number of issues can be considered, for example (a) that the point of departure is not equivalent to the NAEL (no adverse effect level), (b) uncertainties about human variability in cell cycle control and DNA repair, and (c) uncertainties about the shape of the dose-response curve below the BMD and the dose level below which the cancer incidence is not increased (European Commission, 2009).

As part of this project on the interpretation of the MOE, we organised a workshop to elicit expert knowledge regarding the form of the dose-response curve for genotoxic carcinogens at human relevant exposures with the view to analysing this information for any implications for a level of concern for the MOE. The questions we desired to answer via the structured expert elicitation exercises in this workshop were:

- What are the known and unknown factors underlying the different dose-response models for genotoxic carcinogens?
- What is the rigorous scientific rationale to support the choice among the different dose-response models and assumptions for genotoxic carcinogens?

The expert elicitation scheme was designed to capture information about the nature of the dose-response curve at human relevant exposures and was conducted over two phases. During the first phase expert judgments were elicited remotely (via a structured online questionnaire) with a view to discussing and refining the elicited judgments in a follow-up experts' workshop (the second phase). The experts' inputs elicited during the first phase were essential because they resulted in the identification of two approaches for interpreting the shape of the dose-response curve, either after or before translating it from available experimental animal data to humans (i.e. approaches 1 and 2 respectively in Figure 1).

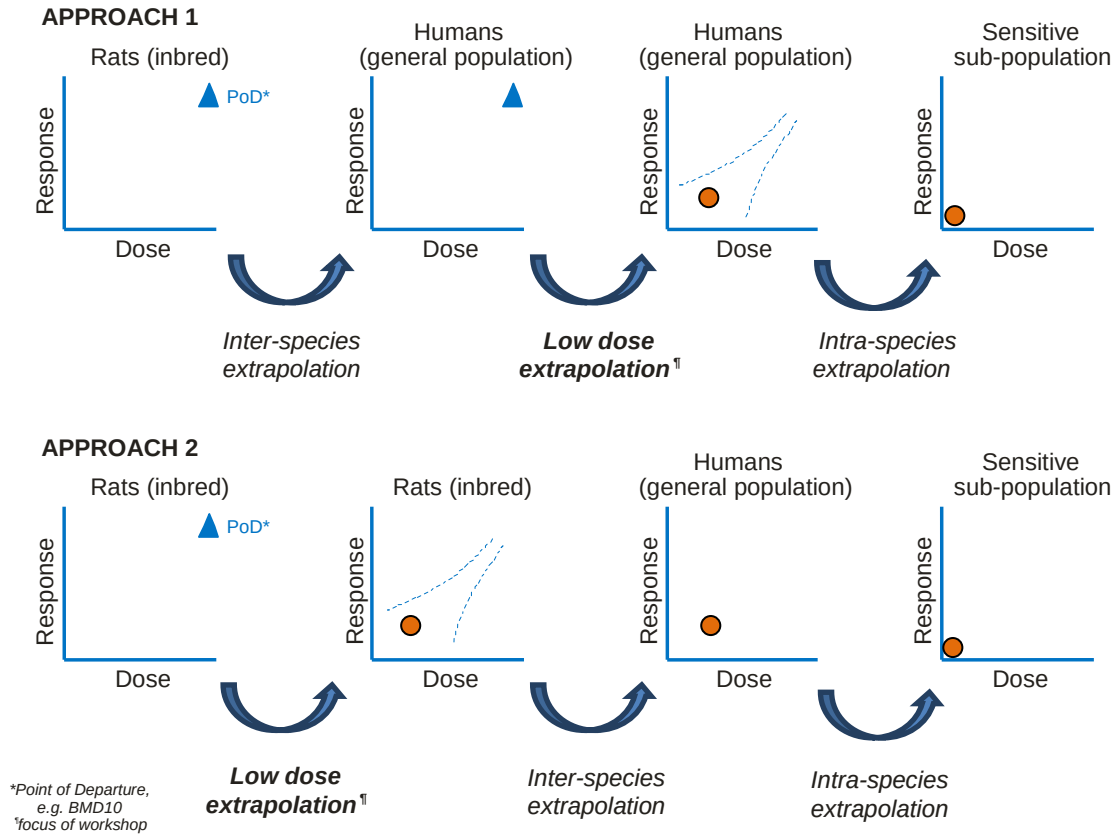

Figure 1 Illustration of possible sequence of steps when assessing risk of genotoxic carcinogens. In particular, one could think of interpreting the shape of the dose-response curve either before or after an inter-species extrapolation is performed. When designing the experts' workshop it was hypothesised that the choice of the approach chosen could affect the final assessment.

A one-and-a-half days experts' workshop was organised at the Food and Environment Research Agency (Fera), Sand Hutton, York (23rd – 24th of May 2011), in which 11 experts from Germany, Italy, Switzerland, UK and USA participated. Two of these experts were also participants in the first phase of this study[1].

One particularly useful outcome of the exercise was the variability within and between experts' quantitative judgments (Figure 2[2]): why do experts' judgments differ so much?

The results of our study indicated that:

- Experts are reluctant to express quantitative views on the dose-response curve for genotoxic carcinogens due to numerous uncertainties burdening the underlying processes.

[1] Experts who participated in May 2011 workshop: Prof. Alan Boobis; Dr Philip Carthew; Dr Rory Conolly; Prof. Corrado Galli; Dr Helmut Griem; Dr Werner Lutz; Barry Maycock; Dr Franz Oesch; Dr Lorenz Rhomberg; Dr Lesley Rushton; Dr Rita Schoney. Full details on elicitation protocol are included in Boobis et al, 2013 (Annex 4).
[2] Full results are included in Boobis et al, 2013 (Annex 4).

- Extrapolation from the point of departure from experimental animals to humans can be performed either first from high to low dose in animals and then to humans, or from animal to humans and then from high to low dose.

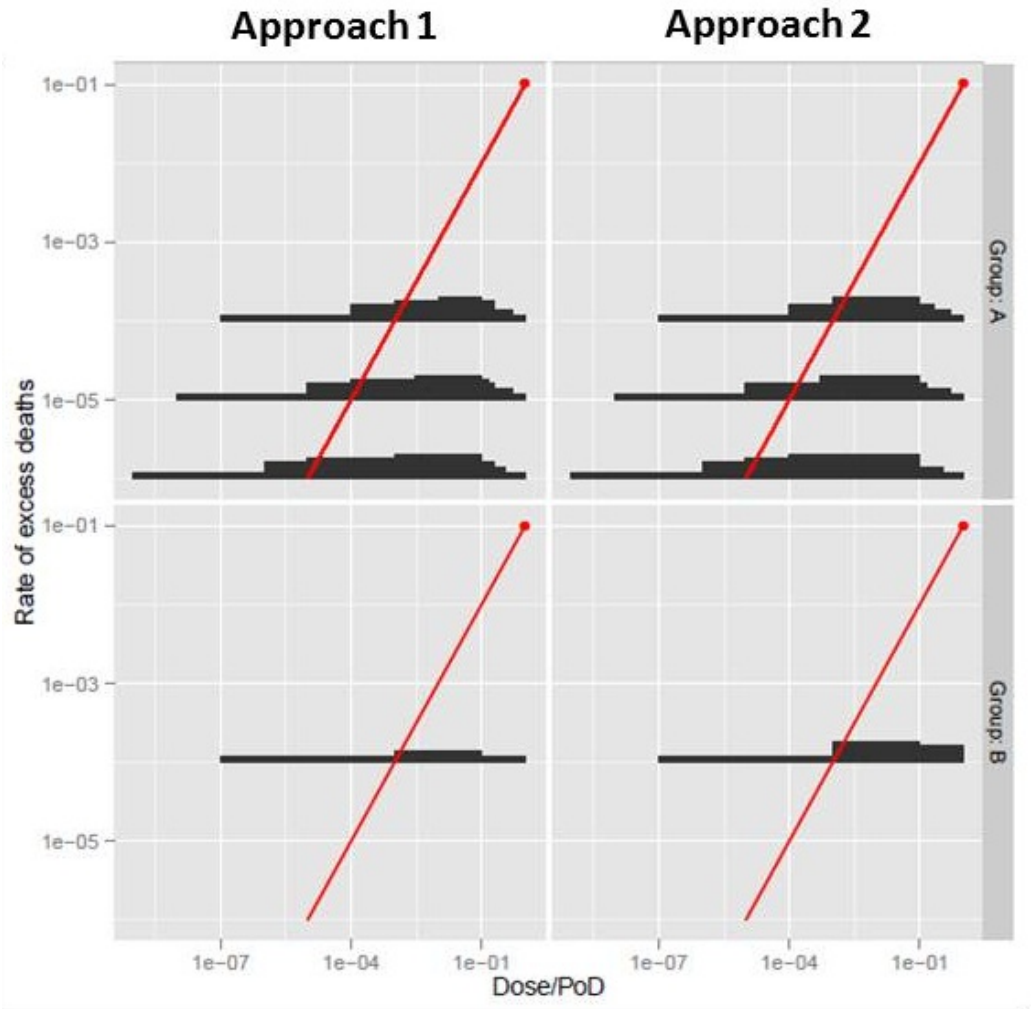

Figure 2: Summarised elicited quantitative estimates of exposure (horizontal black bars) that could lead to a specified number of additional cancer cases in human (approach 1) or inbred rat (approach 2). Diagonal red line indicates the linear extrapolation from point of departure (POD; red dot). Histograms' height is indicative of the number of experts who believed that a particular exposure could lead to a specified number of additional cancer cases (shown on Y-axis). The lowest height shown in the graph indicates the contribution from one expert.

- Expert judgement was that the dose-response curve is highly non-linear at human relevant exposures.
- The majority of experts considered worst-case would be linear, but varied from sub-linear to supra-linear.
- Stochastic events and the distribution of susceptibilities (two different predetermined philosophical perspectives) will contribute to the dose-response relationship, albeit to an unknown extent.

- Most importantly, experts agreed that the level of concern when extrapolating from the point of departure from experimental animals to humans requires consideration of mode of action, species differences and inter-individual variability.

Is the level of concern conservative? How close do experts think that it reflects true risk? The experts considered it was more likely than not that the dose-response curve at exposure levels of concern was non-linear. However, all found it difficult to reach generic conclusions on the contribution of each factor (e.g. cell repair mechanisms, inter-individual differences, differences in mode of action, tissue/s targets, etc) in the cancer process that might contribute to the nature of the dose-response relationship at such human relevant exposure levels. This inevitably has consequences on the interpretation and use of MOE values in the risk assessment of genotoxic carcinogens.

During the workshop, experts indicated that one might need to consider interpretation of the MOE of such chemicals on a case-by-case basis. The follow up workshop aimed therefore to elicit expert judgment on how to classify genotoxic carcinogens by taking into account these factors, and which case studies would best represent classes/categories of genotoxic carcinogens to be evaluated further.

4.2.2 Second Expert workshop on differences in low-dose response relationship between various classes of genotoxic carcinogens 21st – 22nd March 2013

The March 2013 workshop[3] built upon the conclusions of the May 2011 workshop, which employed expert elicitation to capture expert opinion on the general nature of low dose-response relationships for genotoxic carcinogens (see summary of first expert workshop in this report; Boobis et al, 2013 [Annex 4]). The decision on the need to adopt either a linear low-dose response extrapolation (US practice) or an MOE approach (EU practice) for a substance depends heavily on information available on its mode of action (MOA). In particular the interpretation of such information dictates whether the low-dose response of a substance would be expected to have a threshold or could be linear. However, one would need to consider that most MOAs are poorly understood. The question then becomes whether one can use MOA information, in parallel with information on inter-individual and intra-species factors, to reliably group carcinogens and further, to estimate the level of concern more accurately either for each class and/or for each carcinogen, at such low doses.

A two-day expert workshop was held in Fera, Sand Hutton (21st - 22nd March 2013[4]) with the aim of exploring differences in the form of the dose-response curve at human relevant exposures for genotoxic carcinogens where there may be dissimilar modes of action or different modifying factors. Four case studies were selected, to represent different classes of genotoxic carcinogens, in particular: aflatoxin B1, benzo[a]pyrene, ethylmethanesulfonate (EMS) and ethylnitrosourea (ENU). Experts were divided into two breakout groups and each

[3] Results are being analysed and a peer-reviewed publication is being prepared; detailed description of the methodologies applied, results and discussion are provided in a separate annex (Annex 5).

[4] Participants of March 2013 workshop: Prof. Alan Boobis; Dr Sue Barlow; Dr Philip Carthew; Dr Eduardo Cemeli; Dr Kenny Crump; Dr Shareen Doak; Dr Helmut Griem; Dr Gareth Jenkins; Prof Daniel Krewski; Dr John Larsen; Dr David Lovell; Barry Maycock; Dr Franz Oesch; Dr Frances Pollitt; Dr Lorenz Rhomberg; Dr Lesley Rushton; Dr Paul Schlosser; Dr Benjamin Smith;

group assessed two case studies[5]. Initially, experts were asked to build conceptual models (i.e. flowcharts of causal steps and factors influencing each step, including MOA) for the carcinogenic response in each case study (an example of the conceptual model for aflatoxin B1 and benzo[*a*]pyrene is shown in Annex 5, Figure 1). These models served as a coherent collective platform and facilitated the elicitation of quantitative estimates informing the shape of the dose response curve at human relevant exposures that followed (Figure 3).

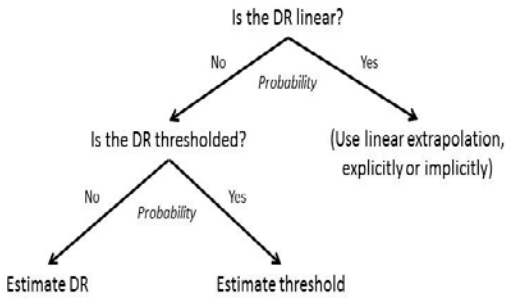

Figure 3: Individual probabilities were elicited of whether (a) the dose-response is linear at human relevant exposures, and (b) there is a threshold for each compound assessed. Thereafter, experts' individual quantitative estimates were elicited of (a) exposures that would cause additional cancers in the human population (i.e. 1 in 10,000; 1 in 100,000; 1 in 1,000,000), and (b) a threshold. Experts disseminated the rationale underlying their estimates.

Following the elicitation of quantitative estimates, experts assessed relevant uncertainties, whether each compound discussed could be a reliable representative of a particular class of genotoxic carcinogen, and whether more classes, and representative case studies, should be included in order to address coherently all genotoxic carcinogens.

Research gaps and uncertainties were flagged during group discussions held in the breakout groups[6], and experts emphasised that quantitative estimates elicited reflect mainly one or a small number of key events and considerations. The experts also emphasised that further consideration of all available evidence may lead to revised estimates, and therefore the elicited quantitative estimates should be perceived rather as "starting points" than final points. Linearity of dose response at human relevant exposures was questioned in all case studies (Figure 4). For aflatoxin B1 there was agreement that the dose-response would likely deviate sub-linearly from the POD to human relevant exposures, and for benzo[*a*]pyrene experts had a mixed view: experts agreed that there would be differences depending on the particular route of exposure (i.e. oral or inhalation) and that there would be some deviation from linearity; however, there was uncertainty as to how much (Figure 4). Experts in this group thought that most probably there is no threshold either for aflatoxin B1 or for benzo[*a*]pyrene. Experts in the group that assessed EMS and ENU had more diverse views regarding the linearity of dose-response at human relevant exposures (Figure 4), and they thought that it is more probable that EMS is thresholded than EMU. Regardless of the

[5] Breakout group 1 assessed: aflatoxin B1 and benzo[*a*]pyrene; breakout group 2 assessed: ethylmethanesulfonate (EMS) and ethylnitrosourea (ENU).

[6] E.g. Aflatoxin B1: Not all of relevant data were available to the group for systematic review; lack of a good concordance study that would address (a) cytotoxicity, (b) genotoxicity, (c) carcinogenicity (i.e. endpoints should be assessed in animal studies; for example, what is POD for cytotoxicity for aflatoxin B1?); Kinetics by different routes of exposure (target tissue exposure); Quantitative information on species differences and critical metabolic routes. Benzo[a]pyrene: Data : What enzymes are induced, in which tissue, at what time and at what dose? Clear definition of (a) time, (b) dose, (c) route dependency of effects? Lack of high quality dose-response data at low doses for both oral and inhalation routes? Physiology: Kinetics in target tissue unknown; what is the contribution of other mechanisms beyond genotoxicity? High dose problem – but what is the impact on POD –what is implication for low dose extrapolation? What effects are apparent below POD? What effects are not apparent below POD?

compound concerned, there was significant variation amongst experts in their quantitative estimates (Figure 2 of Annex 5).

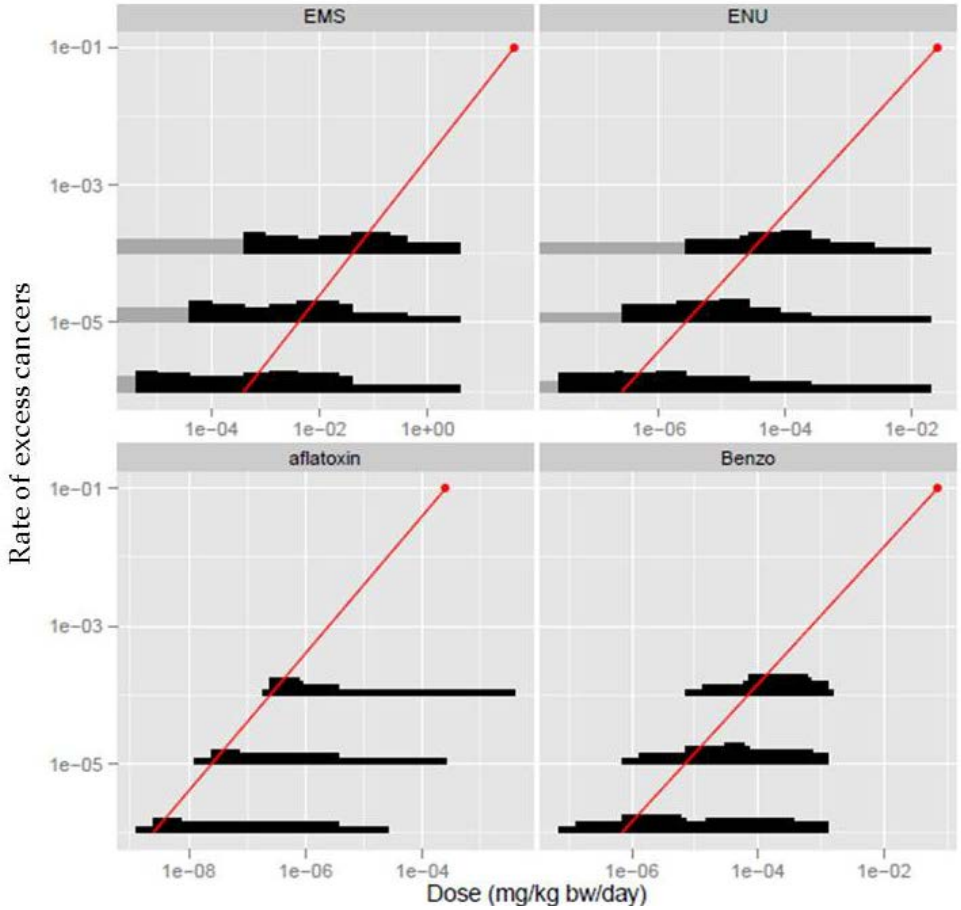

Figure 4: Summary of 90% individual expert ranges (width of horizontal black bars) for the case studies assessed at the 2nd workshop. Diagonal red line indicates the linear extrapolation from POD (red dot). Histograms' height is indicative of the number of experts who believed that a particular exposure could lead to a specified number of additional cancer cases (shown on the Y-axis). The lowest height shown in the graph indicates the contribution from one expert. The horizontal grey shading shown for EMS and ENU denote instances where an expert (or multiple experts) gave 0 as the lower end of their 90% range for the dose leading to the specified excess risk. The extent of agreement amongst the experts varied depending on the compound concerned, and their individual estimates differed.

Experts agreed that there is something more than just DNA reactivity that drives carcinogenicity, and rephrased the question "Representativeness of selected case studies?" to "Should the same MOE (possibly different than 10,000) be used for all genotoxic carcinogens or are they too diverse for this?" It was suggested to take a range of genotoxic carcinogens for which there was evidence for deviation from linearity below the POD and determine the minimum deviation from linearity, if the difference between compounds was not too great, to determine a suitable MOE for such compounds. Otherwise one would have to create a number of different groups of genotoxic carcinogens or treat each compound individually on

a case-by-case basis. It appeared that each case study was unique, and that there is a lot of uncertainty on "how far off" from linearity and chemical specific information is needed for each to decide on what level of MOE would raise concern – an open question from the workshop was whether the carcinogenic process is too complex to group genotoxic carcinogens or whether it would be possible to create small groups of genotoxic carcinogens.

A number of key issues were identified to take into account for a future strategy: (a) several processes are ongoing in carcinogenicity – is the POD a poor index to represent them all? (b) compare POD in the *in vivo* studies to that in *in vitro* studies to find a scaling factor that would enable account to be taken of key factors in order to modify the MOE; (c) how will it be possible to avoid the possibility of being very under (or over) predictive?; (d) nature and determinants of dose-response at low doses is largely unknown; there are many factors that might cause an increase or decrease in response; (e) candidate key processes to group genotoxic carcinogens, e.g. repair, saturation of metabolic steps, autoinduction?; (f) impact of these key processes on deviation from linearity; (g) biomarkers of key events; (h) metabolism – PBPK principles.

5. Objective 04

5.1 Part 1: Use of existing epidemiological data to analyse the relationship between Margins of Exposure and human cancers

The full report for Objective 04, part 1 is provided in Annex 6.

The aim of this part of the project was to compare risk estimates obtained from carcinogenicity data in experimental animals, using the MOE approach, with the measured risk in exposed subjects, obtained from epidemiological studies. Chemicals were sought for which there was sound evidence of carcinogenicity in humans, reasonably assumed to have arisen by a genotoxic mode of action, starting with IARC group 1 carcinogens. A thorough evaluation of chemicals in this group revealed that in many cases, suitable data for this exercise were not available. Either there was inadequate information on exposure-response relationships in human subjects, or no suitable information was available for determination of a POD for carcinogenicity in experimental animals. The compounds for which adequate data could be retrieved were aflatoxin B1, benzidine, chromium VI and vinyl chloride monomer.

Relevant data for humans, including estimates of the excess cancer incidence in exposed populations, were collated and reviewed. An estimate of the exposure that would be expected to be associated with an excess cancer incidence of 1 in 10^5 was determined, based on the excess cancer incidence attributed to a defined exposure to the chemical and the background tumour incidence for the relevant population (sex, geographical area, etc).

Key animal studies were located and $BMDL_{10}$s were determined for all substances *de novo*, to ensure consistency in the modelling. The most sensitive tumour endpoint, for which data suitable for dose-response modelling were available, was selected. Where there was more than one potentially suitable endpoint, all of these were modelled. Preference was given to studies in which the same route of exposure as that primarily involved in human cancer causation, where this was possible. Where this was not possible, and there was a difference in exposure route between animals and humans, exposures were converted to equivalent oral doses, using default values for physiological parameters such as air breathed in 24 h and body weight. No attempt was made to correct for any difference in bioavailability by the different routes. Doses were adjusted for less than 24 per day or less than 7 days per week administration. Dose was also corrected for the molecular weight of any salt, to express dose in substance equivalents. Dietary and drinking water concentrations were converted to intakes using conventional defaults for food and water consumption, respectively. In general, site concordance was not a primary consideration, unless there was good mechanistic data to justify this (which was the case with aflatoxin B1 and vinyl chloride monomer. In the case of benzidine, the difference in target site, bladder in humans and liver in rodents, could be explained mechanistically).

Dose-response modelling was performed with BMDS version 2.4 (US EPA), using default constraints for the models, as appropriate (EFSA, 2009).

The following models were fitted to all data sets:

Gamma
Logistic

LogLogistic
LogProbit
Multistage (2 nested models)
Multistage-Cancer (3 nested models)
Probit
Weibull
Quantal-Linear

Models were rejected if $P<0.05$ (EFSA, 2009), unless otherwise indicated.

In the case of nested models (multistage and multistage-cancer), the model with the lowest AIC and/or chi-squared value was used where there was a large difference in the parameters. Where the difference was small, the model with the fewest parameters was used (EFSA, 2009).

The lowest acceptable $BMDL_{10}$ value was used to estimate the exposure associated with a 1 in 10^5 risk (i.e. that associated with an MOE of 10,000), assuming a linear relationship between exposure and response ($BMDL_{10}$/10,000). Where few if any models were acceptable, even when the statistical criterion was relaxed to $P>0.01$, model fits were inspected visually and those that were judged reasonable were used to identify a suitable POD.

The dose associated with an excess cancer incidence of 1 in 10^5 calculated from the data in experimental animals was compared with the exposure in humans similarly associated with an increase in cancer incidence of 1 in 10^5 above the background. Where possible, quantitative information on uncertainty was obtained, for example 95% confidence intervals on risk estimates in humans. Other sources of uncertainty were identified and described.

Table 1 Summary of epidemiological data

Compound	**Tumour**	**RR (95% CI)**	**Exposure** [**]	**Background (per 10^5)**	**Excess cancers**[+] **(per 10^5)**	**Dose associated with 1 in 10^5 increased incidence**[++]
Aflatoxin B1	HCC[*]	2.5 (1.08-4.86)	3 ng/kg/d	4.2	6.3	0.48 ng/kg/d
Benzidine	Bladder	13 (4.79-28.4)	24.9[@] µg/kg/d 5.2[@@] µg/kg/d	12.7	152.4	0.16 µg/kg/d[@] 0.034 µg/kg/d[@@]
Chromium VI	Lung	2.09 (1.08-3.65)	2.17 µg/kg/d	28.1	30.6	0.071 µg/kg/d
Vinyl chloride	Liver, mainly angiosarcomas	1.19 (0.25-3.47)	63.9 mg/m^3	4.2	0.8	79.9 mg/m^3 Equiv to 22.8 mg/kg/d

[*]HCC = Hepatocellular carcinoma
[**]Converted to oral equivalent as necessary
[@]High exposure estimate
[@@]Mean exposure estimate
[+]Background x (RR-1)
[++] Exposure/Excess cancers

Table 2 Summary of data from experimental animals

Compound	Tumour	$BMDL_{10}$	BMD_{10}	Dose associated with 1 in 10^5 increased incidence[+]
Aflatoxin B1	HCC[*]	0.22 µg/kg	0.33 µg/kg	0.022 ng/kg/d
Benzidine	HCC	0.38 mg/kg	0.48 mg/kg	0.038 µg/k/d
Chromium VI	Small intestine	0.84 mg/kg	1.33 mg/kg	0.084 µg/kg/d
Vinyl chloride	Angiosarcomas	102 mg/m^3	138 mg/m^3	10.2 $\mu g/m^3$ Equiv to 6.5 µg/kg/d

[*]HCC = Hepatocellular carcinoma
[+]$BMDL_{10}$/10,000

Table 3 Conclusions

Compound	Dose resulting in 1 in 10^5 in cancer incidence		Ratio (Human/animal
	Human-based	**Animal-based**	
Aflatoxin B1	0.48 ng/kg/d	0.022 ng/kg/d	22
Benzidine	0.16 µg/kg/d[*] 0.034 µg/kg bw/d[**]	0.038 µg/kg/d	4[*] 1[**]
Chromium VI	0.071 µg/kg/d	0.084 µg/kg/d	1
Vinyl chloride	79.9 mg/m^3 (22.8 mg/kg/d)	10.2 $\mu g/m^3$ (6.5 µg/kg/d)	7800 (3500)

[*]Assume high exposure
[**]Assume mean exposure

The dose associated with a 1 in 10^5 increase in cancer risk based on extrapolation of animal data was much less for aflatoxin B1 (20 times) and VCM (8000 times), slightly less for benzidine but similar for chromium, to the dose associated with this risk level, determined in exposed humans. Another way of expressing this is that the potency of aflatoxin B1 and vinyl chloride was substantially over-predicted from the animal data relative to the potency actually determined in humans, that of benzidine was slightly over-predicted whilst that of chromium VI was as predicted. These results generally provide some confidence that the MOE cut-off of 10,000 would be adequately protective for the case carcinogens, given the level of concern considered acceptable by risk managers. Uncertainties in the data, e.g. from poor exposure assessment in human studies, heterogeneity in both animal and human studies, differences in routes of exposure and target organs between animal and humans are unlikely to greatly affect this conclusion for aflatoxin B1 and vinyl chloride, given the margin in potency estimates. However, given the margins for benzidine and chromium, there would seem to be merit in assessing the MOE for these compounds in more detail.

5.2 Part 2: Use of existing epidemiological data to estimate upper bounds for the incidences of unrecognised chemical-induced cancers

The full report for Objective 04, part 2 is provided in Annex 6.

The aim of this part of the project was to compare risk estimates obtained from carcinogenicity data in experimental animals, using the MOE approach, with upper bound estimates of the risk in exposed subjects, obtained from epidemiological studies. Carcinogens were sought from IARC groups 1, 2a and 2b for which there was good evidence that carcinogenicity in experimental animals was by a genotoxic mode of action, but for which there was no reported association with any increase over the background incidence of cancer in exposed subjects.

A thorough evaluation of chemicals in this group revealed that in many cases, suitable data for this exercise were not available. Either there were no reported studies in exposed subjects or no suitable information was available for determination of a POD for carcinogenicity in experimental animals. The compounds for which adequate data could be retrieved were acrylamide, ethylene oxide and tamoxifen.

As site concordance cannot be assumed for a genotoxic carcinogen, in the absence of mechanistic information, choice of tumour type in humans was not immediately obvious. This has been discussed on a case-by-case basis for the three case chemicals in the main report (Annex 6). A starting point was that the greatest possible risk (in terms of population attributable risk) would be for that cancer type with the highest background incidence. This was the case for acrylamide, where data on breast cancer were used. In the case of ethylene oxide, the main focus has been on leukemia and whilst no consistent association with exposure has been demonstrated, this was selected as the critical endpoint. Information on other tumour types in humans was scant to non-existent. Tamoxifen can cause cancer in experimental animals by both genotoxic (liver) and non-genotoxic (endocrine) modes of action. For the purpose of this exercise, only liver cancer data were investigated in humans. However, there is little evidence that tamoxifen is carcinogenic at other sites in humans, apart from the uterus, where the endocrine mode of action is responsible.

Existing epidemiological data were retrieved and used to estimate upper bound estimates for the incidences of unrecognised chemical-induced cancers at the sites discussed above. The upper 95% confidence interval was used to estimate the lower bound of an exposure that could be associated with an increase over the background incidence of cancer of 1 in 10^5, based on the excess cancer incidence potentially attributable to a defined exposure to the chemical and the background tumour incidence for the relevant population (sex, geographical area, etc).

Key animal studies were located and $BMDL_{10}$s were determined for all substances *de novo*, as described above under Objective 4.1. Using an MOE of 10,000, the dose predicted to be associated with an excess risk of 1 in 10^5 (i.e. $BMDL_{10}$/10,000) was calculated. The two estimates of exposure, animal and human, associated with an excess cancer incidence of 1 in 10^5 were compared. Where possible, quantitative information on uncertainty was obtained, for example 95% confidence intervals on risk estimates in humans. Other sources of uncertainty were identified and described.

Table 4 Summary of epidemiological data

Compound	Tumour	RR (95% CI)	Exposure*	Background (10^5)	Excess cancers+ (10^5)	Dose associated with 1 in 10^5 increased incidence++
Acrylamide	Breast	1.5 (0.6-**3.6**) per 10 µg per person	0.167 µg/kg/d**	62.8	163.3	1.02 ng/kg/d
Ethylene oxide	Leukaemia	0.95 (0.64-**1.35**)	13.3 mg/m^3	9.9	3.47	3.83 mg/m^3 Equiv to 1.09 mg/kg/d
Tamoxifen	Liver	3.3 (0.92-**12.1**) at a dose of 40 mg/day	0.67 mg/kg/d**	2.2	24.4	27.3 µg/kg/d

*Converted to oral equivalent as necessary
**Assuming a body weight of 60 kg
+Background x (95% CI RR-1)
++ Exposure/Excess cancers

Table 5 Summary of data from experimental animals

Compound	Tumour	$BMDL_{10}$	BMD_{10}	Dose associated with 1 in 10^5 increased incidence+
Acrylamide	Mammary gland	0.307 mg/kg/d	0.544 mg/kg/d	30.7 ng/kg/d
Ethylene oxide	Lung	16.1 mg/m^3	23.3 mg/m^3	0.29 $\mu g/m^3$ Equiv to 0.338 µg/kg/d
Tamoxifen	HCC*	1.48 mg/kg/d	2.71 mg/kg/d	0.15 µg/kg/d

*HCC = hepatocellular carcinoma
+$BMDL_{10}$/10,000

Table 6 Conclusions

Compound	Dose resulting in 1 in 10^5 in cancer incidence		Ratio (Human/animal
	Human-based	Animal-based	
Acrylamide	1.02 ng/kg/d	30.7 ng/kg/d	0.033
Ethylene oxide	3.83 mg/m^3 Equiv to 1.09 mg/kg/d	0.29 $\mu g/m^3$ Equiv to 0.338 µg/kg/d	13,200 (3,200)
Tamoxifen	27.3 µg/kg/d	0.15 µg/kg/d	180

For ethylene oxide and tamoxifen the dose associated with a 1 in 100,000 excess cancer risk based on extrapolation from animal data was much less than the upper bound estimate based on data from exposed human subjects, approximately 180 times for tamoxifen and 1300 times for ethylene oxide. That is, the animal data over-predicts the upper bound estimate of potency in humans. This provides confidence that the MOE cut-off of 10,000 would be adequately protective for both of these compounds and shows that the predicted excess risk in humans is much greater than the actual risk.

However for acylamide the dose associated with a 1 in 100,000 excess cancer risk based on extrapolation from animal data is 30 times that giving the upper bound estimate based on human data. It is therefore not possible to determine, from the available data, whether an MOE of 10,000 would be adequately protective of humans for acrylamide. This, together with the very low margins of exposure reported by Bolger et al (2010) for exposure to acrylamide in high and average consumers (40 and 160, respectively), emphasises the uncertainty in the possible risk to humans from exposure to acrylamide.

6. Objective 05: Integration the lines of evidence from Objectives 01-04: Conclusions and recommendations on levels of concern for interpreting Margins of Exposure for genotoxic carcinogens

6.1 Introduction

The current project was undertaken to try to develop a robust scientific rationale for defining levels of concern associated with given values or ranges of MOEs for genotoxic carcinogens. This was investigated in several different ways. After first reviewing current practice in the use and interpretation of MOEs (Objective 1), the project conducted a critical review of different types of evidence about responses at low doses and their implications for the interpretation of MOEs. The main types of evidence considered were: studies concerning the empirical form of dose-response relationships at low dose exposure to genotoxic carcinogens in experimental animals (Objective 3.1); elicitation of expert opinion on the general form of the dose-response relationship for genotoxic carcinogens in general, and for a number of specific chemicals, including consideration of mode of action (Objective 3.2); analysis of the relationship between potency predicted on the basis of the margin of exposure from studies in experimental animals with that determined in studies in humans using existing epidemiological data on cancer (Objective 4.1); and use of existing epidemiological data to compare upper bound estimates for the incidences of unrecognised chemical-induced cancers with estimates based on potency predicted using the margin of exposure from studies in experimental animals. In addition, a statistical framework was developed for combining multiple assessment factors, when these are used for determining levels of concern (Objective 2).

In Objective 5, the various lines of evidence obtained through Objectives 1-4 were combined and integrated, to reach conclusions on the level of concern associated with an MOE of 10,000 and to provide recommendations for further work in this area.

6.2 Overview of key evidence

A graphical approach was developed to provide a concise synthesis of all of the key evidence reviewed in the project (Figure 5). The format of the graph is designed to facilitate examination of the evidence in relation to major questions that are relevant for the interpretation of MOEs:

1. What level of excess cancer risk is associated with an MOE of 10,000, the value proposed by EFSA and some other authorities as an indicator of low concern?
2. What is the form of the dose response for genotoxic carcinogens at human-relevant exposures?
3. Is assuming linearity of response based on extrapolation from the point of departure in experimental animals protective for humans?

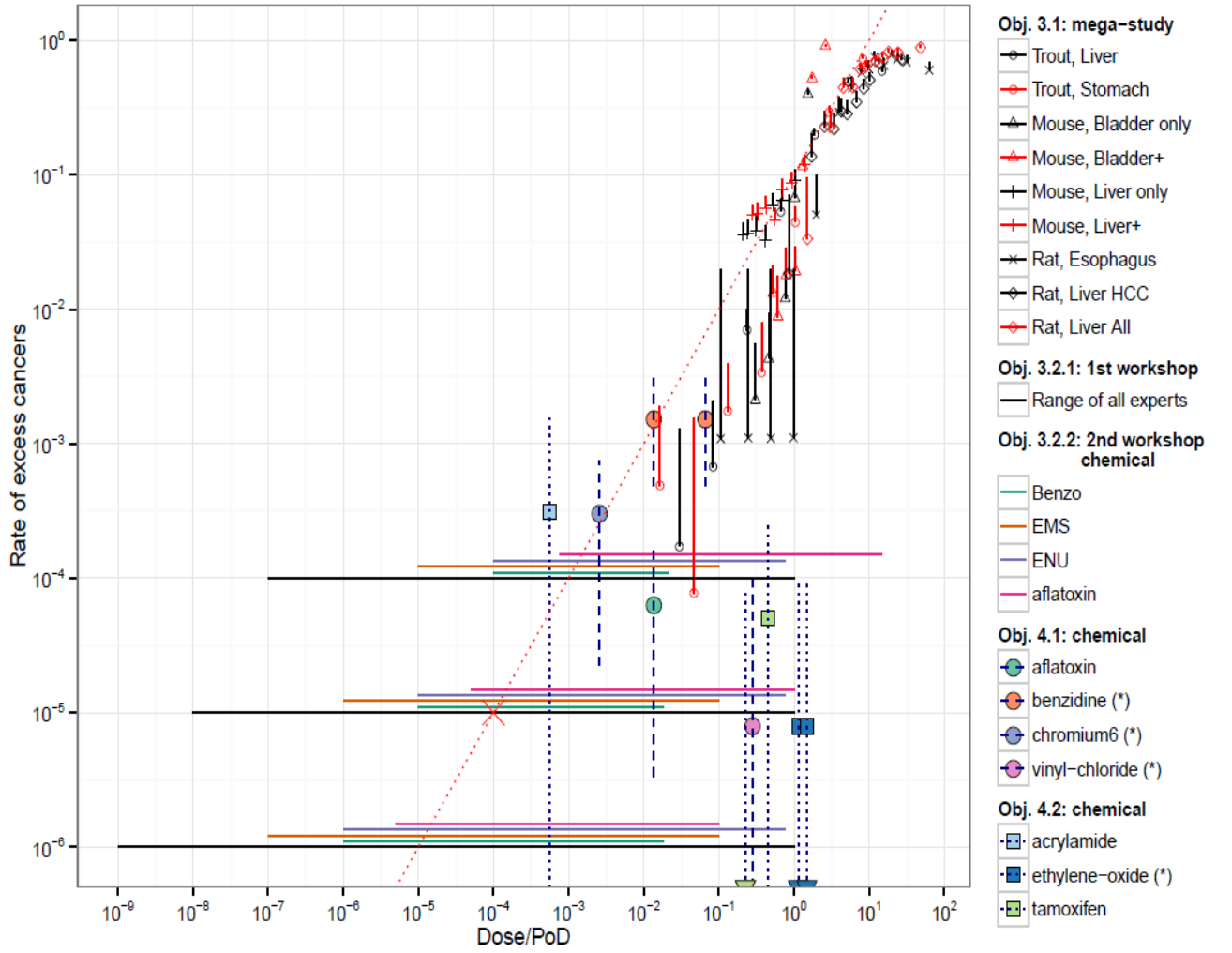

Figure 5. Graphical summary of the key evidence reviewed in this project regarding the form of the dose-response for genotoxic carcinogens for a wide range of dose levels including human-relevant exposures. Different types of evidence were obtained from large-scale animal studies (mega studies), human data on specific chemicals, and judgements from expert workshops, as indicated by the key. Estimates of excess cancer risk from each source were plotted against dose, expressed as a fraction of the Point of Departure from studies in experimental animals (Dose/PoD). The vertical lines provide quantitative estimates of uncertainty around the central estimates. See text for further explanation. Note that in the case of the 2nd expert elicitation exercise (coloured horizontal lines) these have been displaced vertically to increase visibility. The actual risk level in each case is the same as that shown for the 1st expert elicitation exercise (black horizontal lines). The diagonal red dotted line shows the dose response assuming linearity, and the large red cross (X) shows the level of excess cancer risk at a Margin of Exposure of 10,000, also assuming linearity. *Exposure in humans (benzidine and chromium VI) or in both humans and experimental animals (vinyl chloride and ethylene oxide) was by the inhalation route. All other cases relate to dietary exposure.

Figure 5 addresses these questions by plotting all the key evidence together on a single dose-response graph. In order to facilitate comparisons between datasets, the horizontal (dose) axis is normalised by expressing each dose as a fraction of the Point of Departure (POD). Hence all estimates can be plotted on the same horizontal axis in Figure 5, labelled 'Dose/POD' (a unit-less ratio). For the same reason, the incidences of cancer in each dataset were expressed as rate of excess cancers (compared to background or control levels), so that all the responses could be plotted on the same vertical axis, labelled 'Rate of excess cancers' in Figure 5.

Thus the scaling of the axes in Figure 5 allows data on dose response relationships from different types of evidence to be plotted on a single graph:

- **Mega studies**. Dose-response data from large-scale studies in experimental animals (mega-studies) appear in the top right of Figure 5. A $BMDL_{10}$ was obtained for each study by benchmark dose modelling carried out in accordance with current guidance (EFSA, 2009). For each study, doses administered were expressed as a fraction of the $BMDL_{10}$ (Dose/POD) and the cancer incidences were expressed as excess over control level. These data were then plotted in Figure 5, using different symbols and colours to distinguish different studies and cancer types. Symbols are plotted at the median estimate of excess cancers for each dose level, with a vertical line showing the range from the median estimate to the upper 95% confidence limit.
- **Expert workshops.** Expert judgements about dose-responses in humans, taking account of the POD in experimental animals and what is known about modes of action of genotoxic carcinogens, are shown as horizontal lines in the bottom half of Figure 5. Experts at two workshops were asked to provide estimates for the human dose, expressed as a function of the POD from studies in experimental animal, at which particular levels of excess risk would occur in humans (1 in 10^4, 1 in 10^5 and 1 in 10^6). The black lines summarise the range of expert estimates for genotoxic carcinogens in general (Objective 3.2.1) and the coloured lines summarise estimates for specific examples of genotoxic carcinogens, taking account of specific information on factors that might influence the response (Objective 3.2.2). The width of each line reflects both uncertainty expressed by individual experts and variation of opinion between them. The estimates for the specific examples have been displaced vertically to improve the visibility of the graph. The actual risk estimates are the same as those shown respectively for genotoxic carcinogens in general, i.e. the black horizontal lines.
- **Human data.** Empirical estimates of additional cancer rates in humans from epidemiological studies are plotted in the central area of Figure 5. Circles show results for four IARC Class 1 carcinogens (Objective 4.1), while squares show results for three animal carcinogens for which there is no convincing epidemiological evidence of carcinogenicity in humans (Objective 4.2). Each additional cancer risk estimate[7] is plotted against an estimate of the level of human exposure in the epidemiological study from which the risk estimate was derived, expressed as a fraction of the lowest appropriate $BMDL_{10}$ from studies in experimental animals for the same chemical, calculated in accordance with current guidance (EFSA, 2009). Vertical dotted or dashed lines show the uncertainty in the excess cancer rate, derived from the confidence interval reported in the epidemiological study.

When comparing these different types of evidence in Figure 5, it must be borne in mind that the data from mega studies relate to excess cancer rates in animals, whereas the expert judgements and human data relate to excess cancer rates in humans.

[7] Results for Objectives 4.1 and 4.2 are plotted as additional risks, not excess risks. For example, if the non-exposed (baseline) group rate of cancers is 5% and the exposed group rate is 15%, the additional risk is simply (15-5)% = 10% whereas the excess risk is (15-5)/(100-5)= 10.5%. For the studies presented in Figure 5, the difference between the two measures is negligible, and does not affect comparisons or interpretation.

Also plotted in Figure 5 are:

- **Dose response assuming linearity**. The diagonal red dotted line shows the dose response relationship that results from taking the $BMDL_{10}$ as the Point of Departure and assuming linearity. At a dose equal to the $BMDL_{10}$, Dose/POD=1 (shown on the horizontal axis as 100) and the excess cancer rate is assumed to be 10% (shown on the vertical axis as 10^{-1}). When linearity is assumed, dose and response vary together in the same proportions, for example, at a dose equal to one tenth of the $BMDL_{10}$ (Dose/POD = 0.1 or 10^{-1}), the expected rate of excess cancers is 1% (10^{-2}).
- **Excess cancer rate for MOE of 10,000 assuming linearity**. This is plotted in Figure 5 as a large red 'X'. An MOE of 10,000 is obtained when human exposure is a factor of 10,000 below the POD, i.e. Dose/POD = 10^{-4}. Assuming 10% excess cancers at the $BMDL_{10}$ and linear extrapolation to lower doses, the expected excess cancer rate at Dose/POD is 10^{-5}, or 1 in 100,000. When plotted in Figure 5, this falls on the red dotted line towards the bottom of the graph.

The five different types of evidence plotted in Figure 5, as described above, are used together to inform consideration of the three principal questions posed earlier.

1. **What level of excess cancer risk is associated with an MOE of 10,000, the value proposed by EFSA and some other authorities as an indicator of low concern?** It can be seen from Figure 5 that the animal and human data provide no direct evidence on this question, because they all relate to doses higher than that associated with an MOE of 10,000. Most of the data are at doses 2 or more orders of magnitude higher. Therefore, answers to this question must be based primarily on the expert judgements from Objective 3.2, supported by extrapolation (if justified) from the animal and human data at much higher exposures.
2. **What is the form of the dose response for genotoxic carcinogens at human-relevant exposures?** Consideration of this question may be informed by examining trends in the distribution of results in Figure 5, remembering that the data from mega studies relate to responses in experimental animals while the other data relate to responses in humans. A sub-question of particular interest in the context of risk assessment is: are the data indicative of linearity, supra-linearity, or sub-linearity? Results clustered close to the diagonal red dotted line would be consistent with linearity, results above the line might suggest supra-linearity, and results below the line would suggest sub-linearity. The availability evidence shown in Figure 5 indicates a trend towards sub-linearity at lower exposures although, as discussed below, the implications of this need appropriate consideration of uncertainty.
3. **Is assuming linearity of response based on extrapolation from the point of departure in animals protective for humans?** This question is closely related to the preceding question, but focussed on the assumption of linearity. Assuming linearity will be protective if the actual rate of excess cancers in humans is lower than the rate expected assuming linearity: i.e. if the actual rate falls below the diagonal red dotted line in Figure 5. Some of the evidence in Figure 5 lies above the line, while other evidence lies well below it, so it becomes necessary to consider whether assuming linearity is sufficiently protective, given the evidence. This requires consideration of uncertainties affecting the evidence (see below).

In considering answers to each of these questions, and in any other inferences based on Figure 5, it is important to consider uncertainties affecting the evidence it contains. Some types of uncertainty are shown in the graph: vertical solid lines show confidence intervals for the excess risks estimated from mega-studies and vertical dashed or dotted lines show confidence intervals for excess risks estimated from epidemiological studies, while horizontal solid lines show variation and uncertainty in expert judgements about dose-response relationships. However, each type of evidence in Figure 5 is also affected by other potential sources of uncertainty. These imply additional uncertainty about the proper location of each piece of evidence on the graph. Some uncertainties act on the estimation of Dose or POD, and would therefore add uncertainty to the position of points on the horizontal axis. Other uncertainties act on the estimation of excess cancer rates, and would therefore add uncertainty to the position of points on the vertical axis. Ideally, one would quantify and aggregate these additional uncertainties and show them by extending or adding to the confidence intervals already shown in Figure 5. However, quantifying the additional uncertainties would be very difficult and is outside the scope of the present project. Instead, therefore, the impact of additional uncertainties is considered qualitatively in the following sections, and taken into account when answering the questions above and forming overall conclusions.

The following sections discuss each of the main types of evidence in Figure 5 in turn, followed by overall conclusions for the project as a whole.

6.3 Findings from mega-studies (Objective 3.1)

The review of the available literature on studies in experimental animal in Objective 3.1 suggests that evidence for the nature of the dose-response relationship for genotoxic carcinogens needs to be evaluated on a case-by-case basis. For some compounds and tumour types there is reasonable evidence for non-linearity, and that the assumption of linearity will lead to an over-estimation of risk. In other cases, it is not possible on the basis of the available information, to dismiss linearity. However, no study had sufficient power to explore dose-response relationships at incidence levels that would be considered to be of low concern in the human population, ≤ 1 in 10^5.

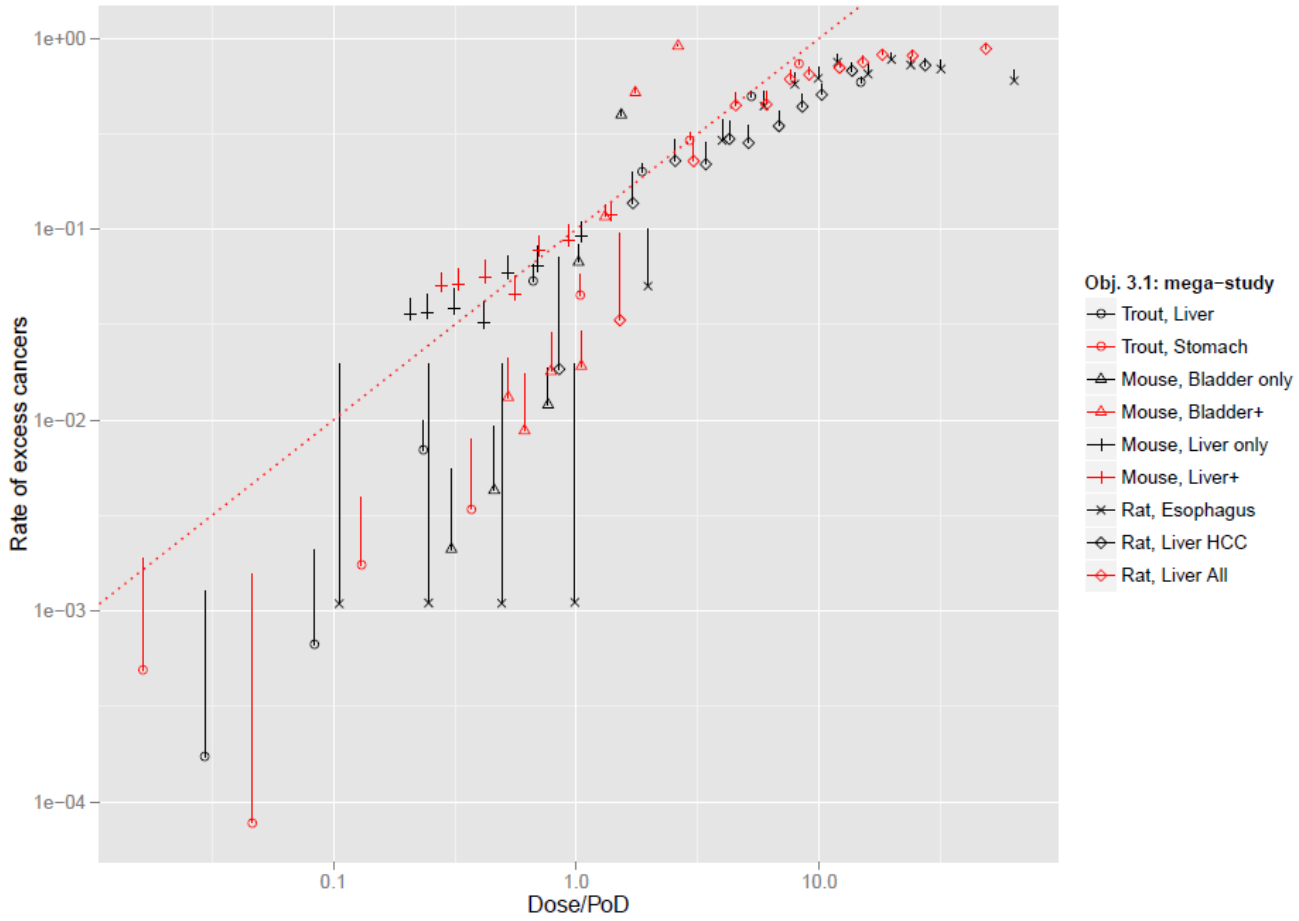

Figure 6. Graphical summary of the results from large studies in experimental animals (mega studies) reviewed in Objective 3.1, plotted in the same format as Figure 5. Estimates of excess cancer risk for each dose group in each study are plotted against dose, expressed as a fraction of the $BMDL_{10}$ for the same study (Dose/POD). The vertical line attached to each point shows the range from the median estimate of excess risk to the upper 95% confidence interval limit. The diagonal red dotted line shows the dose response assuming linearity. See text for further explanation. Mouse, liver+ and Mouse, bladder+ show the results when neoplastic and preneoplastic lesions were combined for the analysis, respectively.

The patterns identified from the review in Objective 3.1 can also be seen when the data from the key mega-studies are plotted together in Figure 5. Figure 6 is an enlarged version of Figure 5, showing only the mega studies on a larger scale to facilitate closer examination.

Interpretation of Figures 5 and 6 is similar to that for ordinary dose-response graphs, except that dose is expressed as a fraction of $BMDL_{10}$ and response is expressed as frequency of excess cancers in the study species. A $BMDL_{10}$ was obtained for each study by benchmark dose modelling, carried out in accordance with current guidance (EFSA, 2009 - see report for Objective 4 (Annex 6) for details).

For each study, administered doses were expressed as a fraction of the $BMDL_{10}$ (Dose/POD) and the cancer incidences were expressed as excess over control level. These data were then plotted in Figures 5 and 6, using different symbols and colours to distinguish different studies

and cancer types. Symbols are plotted at the median estimate of excess cancers for each dose level, with a vertical line showing the upper 95% confidence limit[8].

Examining Figure 6 it can be seen that there is evidence of sub-linearity for all datasets except those for liver cancers in mice, with estimates or confidence intervals clearly below the diagonal dotted red line. For these datasets, assuming linearity below the $BMDL_{10}$ will lead to over-estimation of risk to the study species at lower doses from the range tested.

For liver cancers in mice there are indications of a degree of supra-linearity (estimates or confidence intervals above the diagonal dotted red line). Note that lower confidence limits are not shown in the graph, but that in fact the intervals for the three lowest doses for neoplastic and preneoplastic lesions combined (mouse, liver+) and the two lowest doses for neoplastic lesions only (mouse, liver only) do not extend to or below the diagonal red line. These data are potentially inconsistent with either linearity or sub-linearity although one should bear in mind that many intervals are shown in figure 6, each having a nominal coverage rate of only 95%.

It is important to bear in mind that excess cancer rates estimated from the mega studies relate to cancers in the animal study species, not humans. If humans were expected to be more sensitive than animals, e.g. for allometric or other reasons, then the excess cancer rates in humans would be higher at each dose. Similarly, if sensitivity in the human population is more variable (intra-species variation) than in the laboratory animals (e.g. due to inbreeding of the latter), then excess cancer rates at lower doses might be higher in humans. There are particular uncertainties in the trout study, as this is a non-mammalian species and the dosing regimen (via the aquatic environment) was very different from that used in mammalian studies. These are potential sources of uncertainty when using the results in Figures 5 and 6 to inform conclusions on the interpretation of MOEs in humans. In particular, they raise uncertainty as to whether assuming linearity would be protective within the dose ranges examined by the mega studies. However, no conclusions can be drawn from the mega studies about how these factors (deviations from linearity, inter-species differences and intra-species variation) might combine at lower doses, of more relevance to humans.

None of the mega-studies had sufficient power to explore dose-response relationships at incidence levels that would be considered to be of low concern in the human population, ≤ 1 in 10^5. This is clearly shown in Figure 5, where the mega study data are confined to the upper right part of the graph. A study large enough to examine incidence levels in the order of 10^{-5} would not be feasible in a vertebrate species, due to the number of animals that would be required and the background incidence of most tumour types.

Several groups have evaluated the published data sets for evidence of the existence of a threshold in the dose-response relationship for genotoxic carcinogens. In general, it is not

[8] Confidence intervals for the mega study results were obtained by a Bayesian approach with a Jeffreys prior and are more properly described as credibility intervals. Each point from a mega study represents a dose group, which is compared to a common control group. Each is a binomial sample with uncertain proportion. The two samples are independent. Therefore the Jeffreys prior has independent components for the two proportions and each of those is the standard Jeffreys prior for a binomial proportion: beta(0.5, 0.5). Writing x0 and n0 for number of cancers and number of animals in control and x1 and n1 correspondingly for dose, the posteriors for p0 and p1 are independent: p0 is beta(0.5+x0, 0.5+n0-x0) and p1 is beta(0.5+x1, 0.5+n1-x1). The excess risk is r=(p1-p0)/(1-p0) and the simplest way to find the posterior median and 2.5% and 97.5% credibility limits for r is to (a) take a large sample (N=1e6) of p0 values from the p0 posterior and a corresponding sample of p1 values from the p1 posterior (trivial in R), and (b) compute r for each pair of (p0,p1) values and then find the sample median and 2.5th and 97.5th percentiles. Simulations verified that N=1e6 is large enough that Monte Carlo sampling error is invisible in terms of the logarithmic scale in the plots.

possible to establish the presence or absence of a threshold on the basis of empirical observation, so conclusions about thresholding cannot be drawn from Figure 6. Mechanistic studies, in which the dose-response of necessary intermediate events in the mode of action for the carcinogenic response are characterised, are more promising for this purpose. Studies by Williams and colleagues reviewed in Objective 3.1 provide evidence of the advantages of such an approach. It was not possible within the current project to determine a suitable means of illustrating such dose-response data for intermediate events on the same graph as the tumour data. However, given the range of doses used in the studies of Williams et al, their inclusion in Figure 5 would not have changed the overall interpretation of the evidence shown. the potential contribution of mechanistic information was evaluated as part of the expert workshops in Objective 3.2 (below).

6.4 Findings from expert workshops (Objective 3.2)

At two workshops, experts were asked to make judgements about dose-responses in humans, taking account of the POD in experimental animals and what is known about modes of action of genotoxic carcinogens. The ranges of estimates they provided are shown as horizontal lines in the bottom half of Figure 5. The experts were asked to provide estimates for the human dose, expressed as a function of the POD from studies in experimental animal, at which particular levels of excess risk would occur in humans (1 in 10^4, 1 in 10^5 and 1 in 10^6). The black lines summarise the range of expert estimates for genotoxic carcinogens in general (workshop 1, Objective 3.2.1) and the coloured lines summarise estimates for specific examples of genotoxic carcinogens, taking account of specific information on factors that might influence the response (workshop 2, Objective 3.2.2). The range of each line on the horizontal axis reflects the range of estimates provided by the experts. Note that the coloured lines have been displaced vertically to improve visibility of the graph. The level of excess risk is the same as that shown by the respective black horizontal line. Information on the distribution of expert estimates within the range is provided in Figures 2 and 4 (see earlier) and the detailed judgements of each expert are provided in Annexes 4 and 5.

Examining the overall ranges in Figure 5, it can be seen that both for genotoxic carcinogens in general (workshop 1) and for the four case study chemicals (workshop 2), the ranges extend from supralinear to sublinear. Looking at the detailed responses for Workshop 1, central estimates were sublinear and the majority of experts considered linearity as a worst case (see Figure 2 and Annex 4). Looking at the detailed responses for Workshop 2, expert opinion ranged about equally above and below linearity for EMS and ENU, most favoured sublinearity for benzopyrene and almost entirely favoured sublinearity for aflatoxin, though still including linearity as a worst case (see Figure 4 and Annex 5)[9].

Experts at Workshop 1 agreed that extrapolating from the point of departure in experimental animals to relevant doses in humans requires consideration of mode of action, species differences and inter-individual variability. The results obtained in Workshop 2 showed that considering such information for specific chemicals was helpful, but that a high level of uncertainty remained and estimates still ranged from linearity (or above) to strongly sublinear. However, the experts commented that they would expect to have been able to give more confident responses if they had access to all the available evidence on these chemicals

[9] Note that these conclusions are subject to review when detailed analysis for workshop 2 is complete.

and more time to evaluate it. Overall, experts agreed that more detailed, case-by-case consideration of mode of action will be important in improving future assessments.

6.5 Findings from review of chemicals with evidence of carcinogenicity in humans (Objective 4.1)

The approach used for the carcinogens in 4.1 and 4.2 was to plot estimates of excess cancer risk, or the maximum upper bound estimates, in humans (from epidemiological studies) against the exposure associated with these estimates (obtained from the epidemiological studies and in some cases other studies) expressed as a ratio of the $BMDL_{10}$ (calculated from the critical animal studies).

In Objective 4.1, IARC Class 1 carcinogens, i.e. substances that are considered to be carcinogenic in humans, were investigated. Surprisingly, only 4 such substances were identified for which sufficiently adequate data could be retrieved to enable derivation of estimates of excess cancer risk in humans associated with a relevant exposure estimate, as well as a POD from animal data. In addition to those used here, aristolochic acid, PhIP and 1,3-butadine may be suitable examples for this purpose. Vertical uncertainty bars in Figure 5 reflect confidence intervals for relative risk. These do not include other sources of uncertainty affecting estimates of excess cancers, such as uncertainty in the baseline incidence of unexposed subjects.

Uncertainties affecting the rate of excess cancers (vertical axis of Figure 5)
Sources of uncertainty affecting the vertical axis are evaluated for the case carcinogens in Table 7. For each compound, the likely impact of these uncertainties has been considered under "**Evaluation of uncertainties**", leading to a qualitative judgement on the overall consequence in terms of the likelihood that assuming linearity would not be protective (**third column**).

A potential source of uncertainty affecting the vertical axis is site non-concordance: unless the mode of action is known it cannot be assumed that cancers caused in humans will be in the same tissue as those caused in experimental animals. For two of the 4 chemicals (chromium VI and benzidine), the excess risk estimate referred to cancers in a different tissue to that for the POD, although for benzidine, there is a good mechanistic explanation for the difference in site concordance. In the case of chromium VI, there is evidence that the respective tumour types observed in mice and humans were due to portal of entry effects (site-of-contact). This introduces uncertainty with respect to the appropriate dose metric and possible differences in site sensitivity between species. In general, given the potential for non-concordance of the systemic effects of genotoxic carcinogens, it seems possible that the total excess cancers (summed across all tissues) might be appropriate for assessing the level of protection provided by the MOE; rather than the excess for any one site[10]. If so, the points in the summary graph would move upwards for any chemicals where additional excess cancers occur in human tissues other than that to which the estimate refers.

[10] This would represent a novel approach to calculating the MOE, in general, and may merit a more general discussion by relevant experts.

Additional uncertainties affecting the excess cancer rate include the baseline cancer risk in the unexposed human population, and for 4.2 only, the influence of additional negative epidemiological studies. In the case of acrylamide, the tumour site with the highest background incidence (breast) was chosen, as this would result in the highest upper bound estimate for excess risk. However, relative risks (none of which was significant) for many other sites have been determined in subjects exposed to acrylamide. This information could be used to obtain a weighted estimate of upper bound risk, that would more accurately reflect the maximum possible excess of cancer in humans exposed to acrylamide. Such an estimate would be lower than that used in Figure 5, and hence would result in the point for acrylamide moving down, towards or even below the red dotted line. However, such an exercise was beyond the scope of the current project within the time available.

Table 7: Uncertainties Affecting Estimation of Excess Cancer Risk (Vertical Distance from Red Line)

Compound	**Ratio of excess cancer rates expected at the observed dose (assuming linearity) to the upper estimate for observed rate****	**Taking account of uncertainties affecting the estimate of observed excess cancer risk*, how likely is it that the estimate should lie above the diagonal red dotted line in Figure 5): very unlikely, unlikely or less unlikely).**	**Evaluation of uncertainties** **Key to symbols:** ↑ resolving uncertainty would increase estimate of observed excess cancers (move point upwards in Figure XX) ↓ resolving uncertainty would decrease excess cancers ↕ resolving uncertainty could increase or decrease estimate of excess cancers ● uncertainty is negligible and resolving it would have little effect
Objective 4.1			
Aflatoxin B1	8	Unlikely	The liver was the target for carcinogenicity in both rats and humans. This is supported by mechanistic data, so site concordance is not a major source of uncertainty (●)_.
Benzidine	2	Less unlikely	The target for carcinogenicity was the bladder in humans and the liver in mice. There are good mechanistic data to explain this difference in site, and the same primary hepatic metabolite is responsible for the tumour response in both species. Hence, site concordance is not a major source of uncertainty (●). Smoking was not adjusted for in the human epidemiology study so the risk maybe over-estimated (↓).
Chromium VI	0.35	Unlikely	The target for carcinogenicity was the lung in humans and the small intestine in mice. This may have reflected a portal of entry effect. There is no clear mechanistic basis for choice of site in humans. Hence, site concordance is a source of uncertainty (↕)
Vinyl chloride	270	Unlikely	ASL is very rare in humans hence numbers of deaths are small worldwide, leading to uncertainty in background incidence. Background liver cancer rate was used instead, which over-estimates potency in humans although it can probably be assumed that all ASLs are due to exposure to VCM (↓). The liver (angiosarcoma) was the target for carcinogenicity in both rats and humans. This is supported by mechanistic data, so site

			concordance is not a major source of uncertainty (●).
Objective 4.2			
Acrylamide	0.033	Less unlikely	No specific target was identifiable in humans, whilst mammary tumours were the most sensitive response in mice. The cancer type with the highest background incidence was used (coincidentally breast). This introduces uncertainty into the upper bound estimate of risk, so that potency may be over-estimated (↓). Furthermore, there are many additional studies in other tissues, which are mostly negative, suggesting that acrylamide is unlikely to be carcinogenic in humans. This reduces the likelihood that there is an excess risk of breast cancer due to acrylamide (↓). There was no mechanistic basis for the choice of cancer site in humans (↕).
Ethylene oxide	3,200	Very unlikely	No specific target was identifiable in humans but data available on leukaemia. Lung tumours (mice) and leukaemia (rats) were the most sensitive responses in experimental animals. No specific target in humans. There is no mechanistic basis for the choice of cancer site in humans. The background incidence of leukaemia in humans is 9.9 in 10^5, which is lower than for some forms of cancers and higher than for others. Hence, this introduces uncertainty into the upper bound estimate of risk, so that potency may be over or under-estimated (↕). Estimates of excess risk from other human studies were often small and statistically non-significant, with lower confidence limits well below 1 (↓).
Tamoxifen	180	Very unlikely	Upper confidence limit varied greatly between studies, partly due to small sample sizes, the highest of which was used here. (↓). The target for carcinogenicity in experimental animals was the liver, for which there is a good mechanistic understanding. The liver is the likeliest target site in humans for any carcinogenic effect by a genotoxic mode of action. Hence, the issue of site concordance is not a major source of uncertainty (●).

Uncertainties affecting the Dose/PoD (horizontal axis of Figure 5)

Sources of uncertainty affecting the horizontal axis are evaluated for the case carcinogens in Table 8. Note that the ratios in Table 8 differ from those in Table 7 because here they are calculated for the median estimate of the observed cancer risk, rather than the upper confidence limit.

The obvious source of uncertainty affecting the horizontal axis (Dose/POD) is the estimation of human exposure in the epidemiological studies. Information for this was generally poor and could either over- or underestimate the actual exposure of the population to which the relative risk estimate referred. Other sources of uncertainty included the number of animal studies available from which to select a data set for modelling, compared to the fewer studies likely to be available for more typical chemicals, the choice of the most sensitive tumour response in experimental animals from those suitable for modelling and the selection of the lowest $BMDL_{10}$ value for calculation of excess risk in humans.

The POD used was the $BMDL_{10}$ for the study most suitable for BMD analysis, for the most sensitive animal species, strain and tumour type. As such, it includes a measure of the uncertainty in the estimation of the BMD_{10} for that study. However, a less obvious source of uncertainty arises when the plotted points are used to assess whether assuming linearity is protective. This is because, in general, a large number of animal studies were available for the case study chemicals, whereas fewer studies will be available for the majority of chemicals to be assessed by the MOE approach in future. The fewer studies are available, the narrower the range of their results is likely to be. Therefore, if the case study chemical had a more typical number of animal studies, the range might be narrower and consequently the POD could be higher, decreasing the calculated value for Dose/POD and therefore moving the plotted point to the left in the summary graph.

Of the 4 human carcinogens investigated here, for two of them (aflatoxin B1, chromium (VI)), the cancer study used as the basis of BMD_{10} determinations was a typical NTP (type) study with the recommended strains. In the case of benzidine a non-typical, presumably more sensitive, mouse strain was used. In the case of vinyl chloride, although Sprague-Dawley rats were used, the response would also have been detected using Fischer rats. However, the number of dose groups was greater than normal.

Both these uncertainties affecting the horizontal axis (Dose/POD) are potentially substantial but difficult to quantify, so they were not plotted as intervals on the graph. Instead the relevant project team members assessed by expert judgement, for each substance, whether these uncertainties might be large enough to move the plotted point to the left of the red dotted line, since this is critical when evaluating whether the assumption of linearity is protective.

Exposure to some compounds was by the inhalation route in experimental animals, in humans or in both. Extrapolation of dose metrics by the inhalation route differs from that from the oral route. Whilst inhalation exposures were converted to equivalent systemic exposures, using default values for breathing, body weights, etc, the actual dose equivalence varies with each chemical. Hence, this introduces additional uncertainty into estimation of dose.

Table 8: Uncertainties Affecting Estimation of Dose/PoD (Horizontal Distance from Red Line)

Compound	**Ratio of dose required to cause observed level of excess cancers (assuming linearity) to the observed dose causing this excess***	**Taking account of uncertainties affecting Dose/POD*, how likely is it that the estimate should lie to the left of the diagonal red dotted line in Figure 5): very unlikely, unlikely or less unlikely.**	**Evaluation of uncertainties** **Key to symbols:** ← resolving uncertainty would decrease Dose/PoD (move point to left in Figure XX) → resolving uncertainty would increase Dose/PoD ↔ resolving uncertainty could increase or decrease Dose/PoD ● uncertainty is negligible and resolving it would have little effect
Objective 4.1			
Aflatoxin B1	22	Unlikely	Exposure for excess risk on inhalation in the Danish epidemiology study could be an over-estimate because of the assumption of equivalence by the inhalation and oral routes when extrapolating from the US study (←). Many of the studies were occupational and exposure is likely to have been over-estimated (←). NTP quality study in Fischer rats so number of available studies would have had little impact on data choice. This is therefore not an appreciable source of uncertainty (●). The liver was the target for carcinogenicity in both rats and humans. This is supported by mechanistic data, so site concordance is not a major source of uncertainty (●).
Benzidine	4	Less unlikely	Exposure in the key human data relies upon measurements in a very old paper and is likely to be of uncertain quality (↔). NTP quality study in F1 (not typical cross) mice so number of available studies could have had impact on data choice and introduced

			uncertainty (←). The target for carcinogenicity was the bladder in humans and the liver in mice. There are good mechanistic data to explain this difference in site, and the same primary hepatic metabolite is responsible for the tumour response in both species. Hence, site concordance is not a major source of uncertainty (●).
Chromium VI	1	Less unlikely	NTP quality study in typical F1 mice so number of available studies would have had little impact on data choice. This is therefore not an appreciable source of uncertainty (●). Potency overestimated in mice by summing sites and tumour types (←). The target for carcinogenicity was the lung in humans and the small intestine in mice. This may have reflected a portal of entry effect. There is no clear mechanistic basis for choice of site in humans. Hence, site concordance is a source of uncertainty (↔). The human epidemiology studies were occupational and exposure may have been over-estimated (←). Job exposure matrices were used, reducing the precision of the exposure classification (↔).
Vinyl chloride	3500	Very unlikely	The carcinogenicity data for experimental animals were for a large number of dose groups, in Sprague-Dawley rats. Hence, number of available studies could have had impact on data choice and introduced uncertainty (←). The potency is likely to be over-estimated in the rat because several tumour types in liver were combined (←). The liver (angiosarcoma) was the target for carcinogenicity in both rats and humans. This is supported by mechanistic data, so site concordance is not a major source of uncertainty (●).
Objective 4.2			

Acrylamide	0.033	Less unlikely	NTP quality study in Fischer rats so number of available studies would have had little impact on data choice. This is therefore not an appreciable source of uncertainty (●). Acrylamide intake in humans may be misclassified due to uncontrolled confounding and lack of adjustment for other relevant dietary components and/or lifestyle factors (●).
Ethylene oxide	3,200	Very unlikely	NTP quality study in Fischer rats. Study in mice was in A/J strain, not typical of NTP studies. Hence, number of available studies could have had impact on data choice for mouse tumours in mice and introduced some uncertainty (←).
Tamoxifen	180	Very unlikely	Experimental studies were performed in Wistar strain rats, which are more sensitive than Fischer rats. Hence, number of available studies could have had impact on data choice for and introduced some uncertainty (←). Human data were from clinical studies so exposure was well defined (●). The target for carcinogenicity in experimental animals was the liver, for which there is a good mechanistic understanding. The liver is the likeliest target site in humans for any carcinogenic effect by a genotoxic mode of action. Hence, the issue of site concordance is not a major source of uncertainty (●).

*The ratios in Table 8 differ from those in Table 7 because here they are calculated for the median estimate of the observed cancer risk, rather than the upper confidence limit.

Findings from review of known human carcinogens (Objective 4.1)

We now discuss each of the 4 human carcinogens in turn.

- **Aflatoxin B1**: The point for this substance lies 1-1.5 orders of magnitude below and to the right of the red dotted line in Figure 5. The most significant source of uncertainty associated with the evaluation of this substance is likely over-estimation of the potency of the substance in the epidemiology studies, because of route-to-route differences. There is much less uncertainty associated with study design and site concordance. Hence, it is judged that overall these uncertainties are such that it is unlikely their resolution would result in the substance moving on to or beyond the red dotted line.

- **Benzidine**: Estimates for this substance place the point for this substance either on the red dotted line in Figure 5 or less than 0.5 orders of magnitude below and to the right of the line. Significant sources of uncertainty associated with the evaluation of this substance are estimates of exposure in the epidemiological studies, possible confounding by smoking which might result in over-estimation of risk, potency in the critical experimental carcinogenicity study relative to that likely in a more typical study, and the influence of route of exposure in experimental animals and humans on potency. The overall impact of these uncertainties is such that their resolution might result in the substance moving on to or above the red dotted line.

- **Chromium VI**: The point for this substance lies on the red dotted line in Figure 5. Significant sources of uncertainty associated with the evaluation of this substance are the relevance of comparing different responses that are each potentially due to local effects at the portal of entry, in mice and humans, likely over-estimation of potency in the experimental studies by summing tumours from multiple sites in the small intestine, lack of a mechanistic basis for identifying likely sites of carcinogenesis in humans, relative to the target site in experimental animals, and possible over-estimation of exposure in the epidemiological studies. The overall impact of these uncertainties is such that it is unknown what the likely consequence of their resolution would be on the relative position of the substance with respect to the red dotted line.

- **Vinyl chloride**: The point for this substance lies almost 4 orders of magnitude below and to the right of the red dotted line in Figure 5. Significant sources of uncertainty associated with the evaluation of this substance are estimates of exposure in the epidemiological studies, the use of background incidence rates for liver (because angiosarcoma is very rare in unexposed individuals), the number of dose groups used in the critical experimental carcinogenicity studies relative to those in a more typical study, potency is likely to have been over-estimated in the experimental studies, because several tumour types in the liver were combined, and the influence of route of exposure (inhalation) on inter-species extrapolation. Nevertheless, the overall impact of these uncertainties is such that it is very unlikely that their resolution would result in the substance moving on to the red dotted line.

Taking into account the uncertainties associated with this assessment, it is concluded that assuming linearity of response from the $BMDL_{10}$ to human-relevant doses (e.g. assuming a

margin of exposure of 10,000 would cause 10^{-5} excess cancers), is considered protective for vinyl chloride and aflatoxin B1, and potentially unprotective for benzidine and chromium (VI).

Findings from review of potential human carcinogens (Objective 4.2)

Chemicals were sought for which there is good evidence that they cause carcinogenicity in experimental animals by a genotoxic mode of action, but for which available data show no association with any increase in the background incidence of cancer in humans.

Again, surprisingly few chemicals were found that had substantial epidemiological studies where the exposure could be quantified. Styrene is one possible additional example that could be investigated in this category.

Of the 3 experimental carcinogens investigated here, for two of them, acrylamide and ethylene oxide (rat), the cancer study used as the basis of BMD_{10} determinations was a typical NTP (type) study with the recommended strains. In the case of ethylene oxide (mouse) a non-typical, presumably more sensitive, mouse strain was used. For tamoxifen, Wistar-derived rats were used and there are data showing that Fischer rats are more resistant to the hepatocarcinogenic effects of this compound.

Data for these chemicals were evaluated and plotted in the same way as for those in Objective 4.1. The difference is that, for the chemicals in 4.2, the relative risk was not significantly different from 1 at the P=0.05 level, so the lower confidence interval and in some cases the median estimate of excess risk are negative (fewer cancers in the exposed population) and could not be plotted on a log scale. Therefore only the upper confidence bound and, where positive, the median estimate is plotted for these chemicals in Figure 5.

Where only a single epidemiological study was available, the plotted estimates can be interpreted in the same way as for the chemicals in Objective 4.1, and the same potential uncertainties need to be considered. However, for chemicals where there are several epidemiological studies showing no significant association, it is less likely that the chemical is carcinogenic , in which case the estimate of the study selected here might exaggerate the excess risk. This potential bias needs to be considered case by case when interpreting the plotted estimates.

We now discuss in turn the 3 chemicals assessed under Objective 4.2.

- **Acrylamide**: The point for this substance lies almost one order of magnitude above and to the left of the red dotted line in Figure 5. Significant sources of uncertainty associated with the evaluation of this substance are estimates of exposure in the epidemiological studies, choice of the commonest cancer type in humans as a basis for the risk estimates, and use of the highest confidence interval of those available for the chosen cancer site, which would tend to over-estimate potency. Hence, it is judged that overall these uncertainties are such that their resolution could result in the substance moving on to or even below the red dotted line.

- **Ethylene oxide**: The point for this substance lies just over four orders of magnitude below and to the right of the red dotted line in Figure 5. Significant sources of uncertainty associated with the evaluation of this substance are use of a strain of

mouse more sensitive than the typical strain used in cancer bioassays, leading to likely over-estimation of potency, choice of leukaemia as the cancer site in humans, for which the background rate is higher than some cancers and lower for others, and use of the highest confidence interval of those available for the chosen cancer site in humans, which would tend to over-estimate potency. The overall impact of these uncertainties is such that it is very unlikely that their resolution would result in the substance moving on to the red dotted line.

- **Tamoxifen**: The point for this substance lies approximately three orders of magnitude below and to the right of the red dotted line in Figure 5. Significant sources of uncertainty associated with the evaluation of this substance are use of a strain of rat more sensitive than the typical strain used in cancer bioassays, leading to likely over-estimation of potency and use of the highest confidence interval of those available for the chosen cancer site in humans, which would tend to over-estimate potency. The overall impact of these uncertainties is such that it is very unlikely that their resolution would result in the substance moving on to the red dotted line.

Taking into account the uncertainties associated with this assessment we conclude that assuming linearity of response, a margin of exposure of 10,000 is clearly protective for ethylene oxide and tamoxifen and is potentially unprotective for acrylamide.

7. Overall conclusions

There is more uncertainty than might have been expected about the relationship between the $BMDL_{10}$ in experimental animals and the levels of exposure causing lower incidences of excess cancer in either experimental animals or humans. The largest animal studies provide data for cancer rates at doses only 10 to 100 fold below the BMD_{10} (Objective 3.1); expert judgements about the relationship, taking account of available evidence on mode of action, range over several orders of magnitude (Objective 3.2); and there are fewer chemicals than expected where sufficient data exist to make a direct comparison (Objectives 4.1 and 4.2).

In summary,

1. Mega-studies on 2-acetylaminofluorene, N-nitrosodimethylamine, N-nitrosodiethylamine and dibenzo[a,l]pyrene were analysed. There is evidence of sub-linearity for all of the mega-study datasets except those for liver cancers in mice with 2-acetylaminofluorene, which show a degree of supra-linearity (Objective 3.1). However, no conclusions can be drawn from the mega studies about deviations from linearity at lower doses, of more relevance to humans, and it is uncertain how these would combine with inter-species differences and intra-species variation.

2. For genotoxic carcinogens in general, and for four specific examples, ethylmethane sulfonate, ethylnitrosourea, benzo[a]pyrene and aflatoxin B1, the overall range of expert judgements encompasses both superlinearity and sublinearity, though most experts considered sublinearity to be more probable (Objective 3.2).

3. The evaluation of four human carcinogens, aflatoxin B1, vinyl chloride monomer, benzidine and chromium VI, in Objective 4.1 gave no overall indication of whether an MOE of 10,000 is protective. For two of the compounds, aflatoxin B1 and vinyl chloride monomer, there was evidence that it was protective. For the other two, whilst the calculated MOE itself was protective, the uncertainties in the assessment were such that the MOE was potentially unprotective.

4. On the basis of the evaluation of three potential human carcinogens in Objective 4.2, acrylamide, ethylene oxide and tamoxifen it was not possible to reach an overall conclusion on whether an MOE of 10,000 is protective. For two of the compounds, ethylene oxide and tamoxifen, there was evidence that it was protective. For the other compound, acrylamide, the risk estimate based on the MOE from experimental data, together the uncertainties in the assessment were such that it was not possible to determine whether the MOE would be protective or not.

5. Taking into account all of the lines of evidence evaluated, it is apparent that it is not currently possible to establish with certainty the level of concern for an MOE of 10,000. A key limitation is the power of any study to explore the dose-response relationship at human relevant exposures. The most likely approach to reducing uncertainty in this area is a clearer understanding of mode or mechanism of action and the quantitative implications that this has for the dose-response relationship.

Whether the assembled evidence provides sufficient certainty that assuming an MOE of 10,000 is protective is a policy question, not a scientific one, because it depends on how much certainty society desires in managing risks of genotoxic carcinogens and on the economic costs and other consequences of requiring more or less certainty. If, after

considering the evidence above, it was decided that more certainty was required, then a degree of super-linearity could be assumed when interpreting MOEs: for example, an MOE of 10,000 would be interpreted as indicating a potential excess cancer risk of greater than 10^{-5}. If, on the other hand, it was considered that the evidence provide sufficient certainty that assuming a degree of sublinearity is still protective, then that could be adopted as a default assumption: for example, an MOE of 10,000 would be interpreted as indicating a potential excess cancer risk of less than 10^{-5}.

Whatever default assumptions are adopted for interpreting MOEs based on animal studies, consideration needs to be given to what options for refined assessment exist when the MOE raises concern. Where epidemiological data exist, they will generally make an important contribution to the assessment. The experts consulted in Objective 3.2 agreed that more use of information on mode of action would be beneficial. However, their estimates for 4 example substances remained very uncertain, suggesting more or better information on mode of action will be needed in practice, compared to what was considered in the exercise. However, the experts commented that because they did not have access to all relevant information at the workshop, their conclusions were preliminary and it may be possible to refine them with further information on mode of action once this was fully evaluated.

If a decision on default assumptions for interpreting the MOE can be reached on the basis of the evidence summarised above, those assumptions could then be applied in future assessments without the need for considering how multiple uncertainty or assessment factors should be combined. However, if an MOE raises concern, some options for refined assessment may involve the use of factors to represent different steps in extrapolation and/or different sources of uncertainty. There is some evidence for this from the approaches taken by the experts in the Second Workshop. If so, then the findings of Objective 2 will be relevant when considering how multiple factors should be combined to achieve a desired overall level of conservatism.

8. Recommendations for further work

Potential extensions to the current analysis:

- Identify and evaluate additional example substances for Objectives 4.1 and/or 4.2 – e.g. aristolochic acid, PhIP, 1,3-butadiene, styrene (see US EPA's Carcinogenic Potency Project, Swirsky Gold et al., 2008).

- More refined analysis of the examples already considered in 4.1 and/or 4.2, or more quantitative evaluation of the uncertainties affecting them, e.g. regarding a) site concordance, b) background risk, c) exposure estimation, d) number of animal studies compared to more typical chemicals and e) influence of additional negative studies on estimate of excess risk. Power calculations might also be helpful to determine how large studies would need to be to enable some of the uncertainties to be reduced in a meaningful way, i.e. such that they would materially impact on the conclusions above.

- Identification of an MOE for non-oral routes of exposure (inhalation, dermal) associated with the same level of concern as for an MOE of 10,000 following oral exposure.

Other avenues of research:

- Continue/expand research on approaches for the generation and use of mode and mechanism of action information in assessment of genotoxic carcinogens. Dose and time-response relationships for key events should be better defined, and inter-species differences characterised.

- Population variability in key events and their modifying factors would be of value in exploring interindividual differences in susceptibility.

- Ultimately, many of the questions would be best addressed by a quantitative, systems-based approach. The development of systems toxicology has implications well beyond the question of concern here, and investment in the development of such approaches is encouraged.

References

Akpambang VO, Purcaro G, Lajide L, Amoo IA, Conte LS and Moret S (2009). Determination of polycyclic aromatic hydrocarbons (PAHs) in commonly consumed Nigerian smoked/grilled fish and meat. *Food Addit Contam: Part A*, 26, 1096-1103

Andersen ME, Conolly RB and Gaylor DW (2003). Thresholds of carcinogenicity in the ED01 study. *Toxicol Sci*, 74, 486-487.

Andrade PD, de Mello MH, França JA and Caldas ED (2013). Aflatoxins in food products consumed in Brazil: a preliminary dietary risk assessment. *Food Addit Contam Part A*, 30, 127-136

Andres S, Appel KE and Lampen A (2013). Toxicology, occurrence and risk characterisation of the chloropropanols in food: 2-Monochloro-1,3-propanediol, 1,3-dichloro-2-propanol and 2,3-dichloro-1-propanol. *Food Chem Toxicol*, 58, 467-478.

Bailey GS, Reddy AP, Pereira CB, Harttig U, Baird W, Spitsbergen JM, Hendricks JD, Orner GA, Williams DE and Swenberg JA (2009). Nonlinear cancer response at ultralow dose: a 40800-animal ED(001) tumor and biomarker study. *Chem Res Toxicol*, 22, 1264-1276

Bakhiya N, Abraham K, Gürtler R, Appel KE and Lampen A (2011). Toxicological assessment of 3-chloropropane-1,2-diol and glycidol fatty acid esters in food. *Mol Nutr Food Res*, 55, 509-521

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 DJ, Page S, Rolland V, Schlatter J, Tritscher A, Tueting W and 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, Dinoni M, LeBlanc JC, Renwick AG, Setzer W, Schlatter J, Smith B, Slob W, Williams G and Wildeman 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, S2-24

Boobis A, Flari V, Gosling JP, Hart A, Craig P, Rushton L and 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, 106-118

COC (2007). CC/07/08: Further Consideration of the MOE Approach for Communicating the Risks of Exposure to Genotoxic Carcinogens [Online]. Available at http://www.iacoc.org.uk/papers/documents/cc0708.pdf [Last accessed on 01/08/2013]

Cooke R (2010). Conundrums with uncertainty factors (with discussion). *Risk Analysis*, 30, 330-338

COT (2007). Variability and Uncertainty in Toxicology of Chemicals in Food, Consumer Products and the Environment. Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment. Food Standards Agency, London. Available at http://cot.food.gov.uk/pdfs/vutreportmarch2007.pdf [Last accessed on 01/08/2013]

Crump KS and Clewell HJ (2003). Letter. *Toxicol Sci*, 74, 485

Doak SH, Jenkins GJ, Johnson GE, Quick E, Parry EM and Parry JM (2007). Mechanistic influences for mutation induction curves after exposure to DNA-reactive carcinogens. *Cancer Res*, 67, 3904-3911

EFSA (2005). Opinion of the Scientific Committee on a request from EFSA related to A Harmonised Approach for Risk Assessment of Substances Which are both Genotoxic and CarcinogenicLow dose extrapolation from animal carcinogenicity data using various models'. *EFSA Journal*, 282, 1-31 [Online]. Available at http://www.efsa.europa.eu/en/efsajournal/doc/282.pdf [Last accessed on 01/08/2013]

EFSA (2009). Guidance of the Scientific Committee on a request from EFSA on the use of the benchmark dose approach in risk assessment. *EFSA Journal*, 1150, 1-72 [Online]. Available at http://www.efsa.europa.eu/en/efsajournal/doc/1150.pdf [Last accessed on 01/08/2013]

EFSA (2010). Opinion of the Scientific Panel on food additives, flavourings, processing aids and materials in contact with food (AFC) on the safety of polyvinylpyrrolidone-vinyl acetate copolymer for the proposed uses as a food additive. *EFSA Journal*, 8, 1948 (28 pp) [Online]. Available at http://www.efsa.europa.eu/en/efsajournal/doc/1948.pdf [Last accessed on 02/08/2013]

EFSA (2011). Opinion of the Scientific Panel on Contaminants in the Food Chain (CONTAM) on pyrrolizidine alkaloids in food and feed. *EFSA Journal*, 9, 2406 (134 pp) [Online]. Available at http://www.efsa.europa.eu/en/efsajournal/doc/2406.pdf [Last accessed 02/08/2013]

EFSA (2012a). Opinion of the Scientific Panel on food additives, flavourings, processing aids and materials in contact with food (AFC) on the re-evaluation of vegetable carbon (E 153) as a food additive. *EFSA Journal*, 10, 2592 (34 pp) [Online]. Available at http://www.efsa.europa.eu/en/efsajournal/doc/2592.pdf [Last accessed 02/08/2013]

EFSA (2012b). Statement from the Scientific Committee on a request from EFSA on the applicability of the Margin of Exposure approach for the safety assessment of impurities which are both genotoxic and carcinogenic in substances added to food/feed . *EFSA Journal*, 10, 2578 (5 pp). Available at http://www.efsa.europa.eu/en/efsajournal/doc/2578.pdf [Last accessed on 02/08/2013)

EFSA (2013). Opinion of the Scientific Panel on food additives, flavourings, processing aids and materials in contact with food (AFC) on the safety of polyvinyl alcohol-polyethylene

glycol-*graft*-co-polymer as a food additive. *EFSA Journal*, 11, 3303 (30 pp). Available at http://www.efsa.europa.eu/en/efsajournal/doc/3303.pdf [Last accessed on 10/11/2013]
EMA (European Medicines Agency) (2008). Press Release of July 24, 2008. Studies assessed by the EMEA indicate no increased risk of developing cancer for patients who have taken Viracept contaminated with ethyl mesilate [Online]. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2009/11/WC500015040.pdf.

EPA (2005). Guidelines for Carcinogen Risk Assessment [Online]. Available at http://www.epa.gov/raf/publications/pdfs/CANCER_GUIDELINES_FINAL_3-25-05.PDF [Last accessed on 01/08/2013]

European Commission (2009). Risk assessment methodologies and approaches for genotoxic and carcinogenic substances [Online]. Available at http://ec.europa.eu/health/ph_risk/committees/04_scher/docs/scher_o_113.pdf [Last accessed on 05/08/2013]).

FSA (2010). Draft Minutes of Open Board Meeting of 12 October 2010, FSA 10/12/01 [Online]. Available at http://www.food.gov.uk/multimedia/pdfs/board/fsa101201.pdf [Last accessed on 02/08/2013]

FSA (2011). Minutes of Open Board meeting of 15 November, 2011. Report from Director of Food Safety, FSA 11/11/06 [Online]. Available at http://www.food.gov.uk/multimedia/pdfs/board/fsa111106.pdf [Last accessed on 02/08/2013]

Gaylor DW (1980). The ED01 study: summary and conclusions. J Environ Pathol Toxicol, 3, 179-183

Gaylor D, Greenman DL and Frith CH (1985). Urinary bladder neoplasms induced in BALB/c female mice with low doses of 2-acetylaminofluorene. *J Environ Pathol Toxicol Oncol*, 6, 127-136

Gaylor DW and Kodell RL (2000). Percentiles of the product of uncertainty factors for establishing probabilistic reference doses. *Risk Analysis*, 20, 245-250

Gocke E and Müller L (2009). In vivo studies in the mouse to define a threshold for the genotoxicity of EMS and ENU. *Mutat Res*, 678, 101-107

Haseman JK (2003). An alternative perspective: A critical evaluation of the Waddell threshold extrapolation model in chemical carcinogenesis. *Toxicol Pathol*, 31, 468-470

Hernández-Moreno D, Soffers AE, Wiratno, Falke HE, Rietjens IM and Murk AJ (2013). Consumer and farmer safety evaluation of application of botanical pesticides in black pepper crop protection. *Food Chem Toxicol*, 56, 483-490

JECFA (2005). WHO Technical Report Series 930. Evaluation of certain food contaminants. Sixty-fourth report of the Joint FAO/WHO Expert Committee on Food Additives, 8-17 February 2005. WHO, Geneva. Available at http://whqlibdoc.who.int/trs/WHO_TRS_930_eng.pdf [Last accessed on 01/08/2013]

Lachenmeier DW, Gumbel-Mako S, Sohnius EM, Keck-Wilhelm A, Kratz E and Mildau G (2009a). Salivary acetaldehyde increase due to alcohol-containing mouthwash use: a risk factor for oral cancer. *Int J Cancer*, 125,730-735

Lachenmeier DW, Kanteres F, Kuballa T, López MG, Rehm J (2009b). Ethyl carbamate in alcoholic beverages from Mexico (tequila, mezcal, bacanora, sotol) and Guatemala (cuxa): market survey and risk assessment. *Int J Environ Res Public Health*, 6, 349-360

Lachenmeier DW, Kanteres F and Rehm J (2009c). Carcinogenicity of acetaldehyde in alcoholic beverages: risk assessment outside ethanol metabolism. *Addiction*, 104, 533-550

Lachenmeier DW, Reusch H and Kuballa T (2009d). Risk assessment of furan in commercially jarred baby foods, including insights into its occurrence and formation in freshly home-cooked foods for infants and young children. *Food Addit Contam Part A*, 26, 776-785

Lachenmeier DW, Kuballa T, Reusch H, Sproll C, Kersting M and Alexy U (2010). Benzene in infant carrot juice: further insight into formation mechanism and risk assessment including consumption data from the DONALD study. *Food Chem Toxicol*, 48, 291-297

Lachenmeier DW, Kanteres F and Rehm J (2011). Epidemiology-based risk assessment using the benchmark dose/margin of exposure approach: the example of ethanol and liver cirrhosis. *Int J Epidemiol*, 40, 210-218

Lachenmeier DW, Maser E, Kuballa T, Reusch H, Kersting M and Alexy U (2012a). Detailed exposure assessment of dietary furan for infants consuming commercially jarred complementary food based on data from the DONALD study. *Matern Child Nutr*, 8, 390-403

Lachenmeier DW, Przybylski MC and Rehm J (2012b). Comparative risk assessment of carcinogens in alcoholic beverages using the margin of exposure approach. *Int J Cancer*, 131, E995-1003

Lachenmeier DW, Monakhova YB, Markova M, Kuballa T and Rehm J (2013). What happens if people start drinking mouthwash as surrogate alcohol? A quantitative risk assessment. *Food Chem Toxicol*, 51:173-178

Lutz WK (2003). Letter. *Toxicol Sci*, 75, 223

Monakhova YB and Lachenmeier DW (2012a). The margin of exposure of 5-hydroxymethylfurfural (HMF) in alcoholic beverages. *Environ Health Toxicol*, 27. e2012016. doi: 10.5620/eht.2012.27.e2012016

Monakhova YB, Jendral JA and Lachenmeier DW (2012b). The margin of exposure to formaldehyde in alcoholic beverages. *Arh Hig Rada Toksikol*, 63, 227-237

Moro S, Chipman JK, Wegener JW, Hamberger C, Dekant W and Mally A (2012). Furan in heat-treated foods: formation, exposure, toxicity, and aspects of risk assessment. *Mol Nutr Food Res*, 56, 1197-1211

O'Brien J, Renwick AG, Constable A, Dybing E, Muller DJG, Schlatter J, Slob W, Tueting W, Van Benthem J, Williams GM and Wolfreys A (2006). Approaches to the risk assessment of genotoxic carcinogens in food: A critical appraisal. *Food Chem Toxicol*, 44, 1613-1635

Oplatowska M, Donnelly RF, Majithiya RJ, Glenn Kennedy D and Elliott CT (2011). The potential for human exposure, direct and indirect, to the suspected carcinogenic triphenylmethane dye Brilliant Green from green paper towels. *Food Chem Toxicol*, 49, 1870-1876

Peto R, Gray R, Brantom P and Grasso P (1991a). Effects on 4080 rats of chronic ingestion of N-nitrosodiethylamine or N-nitrosodimethylamine: a detailed dose-response study. *Cancer Res*, 51, 6415-6451

Peto R, Gray R, Brantom P and Grasso P (1991b). Dose and time relationships for tumor induction in the liver and esophagus of 4080 inbred rats by chronic ingestion of N-nitrosodiethylamine or N-nitrosodimethylamine. *Cancer Res*, 51, 6452-6469

Poirier MC, Fullerton NF, Kinouchi T, Smith BA and Beland FA (1991). Comparison between DNA adduct formation and tumorigenesis in livers and bladders of mice chronically fed 2-acetylaminofluorene. *Carcinogenesis*, 12, 895-900

Purchase IF and Auton TR (1995). Thresholds in chemical carcinogenesis. *Regul Toxicol Pharmacol*, 22, 199-205

Schuetze A, Heberer T and Juergensen S (2008). Occurrence of residues of the veterinary drug malachite green in eels caught downstream from municipal sewage treatment plants. *Chemosphere*, 72, 1664-1670

Slob W and Pieters MN (1998). A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: general framework. *Risk Analysis*, 18, 787-798

SOT (1981). Re-examination of the ED01 study: Review of statistics: the need for realistic statistical models for risk assessment. *Fundam Appl Toxicol*, 1, 124-126

Tardiff RG, Gargas ML, Kirman CR, Carson ML and Sweeney LM (2010). Estimation of safe dietary intake levels of acrylamide for humans. *Food Chem Toxicol*, 48, 658-667

van den Berg SJ, Serra-Majem L, Coppens P and Rietjens IM (2011a). Safety assessment of plant food supplements (PFS). *Food Funct*, 2, 760-768

van den Berg SJ, Restani P, Boersma MG, Delmulle L and Rietjens IM (2011b). Levels of genotoxic and carcinogenic ingredients in plant food supplements and associated risk assessment. *Food Nutr Sci*, 2, 989-1010

Waddell WJ (2003a). Thresholds of carcinogenicity in the ED01 study. *Toxicol Sci*, 72, 158-163.

Waddell WJ (2003b), Letters to the Editor. *Toxicol Sci*, 74, 485-488

Waizenegger J, Winkler G, Kuballa T, Ruge W, Kersting M, Alexy U and Lachenmeier DW (2011). Analysis and risk assessment of furan in coffee products targeted to adolescents. *Food Addit Contam Part A*, 29, 19-28

Wang J, Liu XM and Zhang ZQ (2009). Exposure assessment of liver cancer attributed to dietary aflatoxins exposure in Chinese residents [Article in Chinese]. *Zhonghua Yu Fang Yi Xue Za Zhi*, 43, 478-481.

Williams GM, Gebhardt R, Sirma H and Stenbäck F (1993). Non-linearity of neoplastic conversion induced in rat liver by low exposures to diethylnitrosamine. *Carcinogenesis*, 14, 2149-2156

Williams GM, Iatropoulos MJ and Jeffrey AM (2005). Thresholds for DNA-reactive (genotoxic) organic carcinogens. *J Toxicol Pathol*, 18, 69-77

Williams GM, Iatropoulos MJ and Jeffrey AM (2004). Thresholds for the effects of 2-acetylaminofluorene in rat liver. *Toxicol Pathol*, 32 (Suppl 2), 85-91.

Williams G, Iatropoulos MJ, Jeffrey AM, Luo F, Wang C and Pittman B (1999). Diethylnitrosamine exposure-responses for DNA ethylation, hepatocellular proliferation, and initiation of carcinogenesis in rat liver display non-linearities and thresholds. *Arch Toxicol*, 73, 394-402

Williams GM, Iatropoulos MJ, Wang CX, Jeffrey AM, Thompson S, Pittman B, Palasch M and Gebhardt R (1998). Nonlinearities in 2-acetylaminofluorene exposure responses for genotoxic and epigenetic effects leading to initiation of carcinogenesis in rat liver. *Toxicol Sci*, 45, 152-161

Zeilmaker MJ, Bakker MI, Schothorst R and Slob W (2010). Risk assessment of N-nitrosodimethylamine formed endogenously after fish-with-vegetable meals. *Toxicol Sci*, 116, 323-335