This is a draft paper for discussion.
It does not represent the views of the Committee and must not be quoted, cited or reproduced.

CC/2022/06

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

Draft document on how the Committees evaluate the relevance and reliability of data when assessing a chemical of concern? – draft 0.e

Introduction

1. The topic of 'biological relevance and statistical significance' has been raised as an area of interest during Committee horizon scanning activities for a number of years. A scoping paper was presented at the Joint COC/COM meeting in November 2020 (CC/MUT/2020/03) also attended by some COT members, which outlined some of the more relevant and significant work that has been published on this issue in recent years. As a result, it was agreed that the general public would benefit from guidance that provided clarity on how the expert Committees evaluate data with respect to consideration of biological relevance and statistical significance.

2. A draft document was prepared which provides a brief outline of the Committee evaluation process focussing on the relevance and reliability of data, written specifically to inform the lay person. It has been reviewed by lay members of the three Committees and was discussed by COC in March 2021 (CC/2021/06), COT in March 2021 (TOX/2021/18) and COM in March 2022 (MUT/2022/03) and June 2022 (MUT/2022/04).

3. The document attached in Annex A is an updated version that has been amended following consideration by COM at its June 2022 meeting (MUT/2022/04).

This is a draft paper for discussion.
It does not represent the views of the Committee and must not be quoted, cited or reproduced.

Questions for the Committee

4. Members are asked to consider the updated draft document and whether they are content that this can now be published following review by COC at the July 2022 meeting.

IEH-C under contract supporting the PHE Secretariat
July 2022

DRAFT

This is a draft paper for discussion.
It does not represent the views of the Committee and must not be quoted, cited or reproduced.

TOX/2022/XX Annex A

Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment

Draft document on how the Committees evaluate the relevance and reliability of data when assessing a chemical of concern? – draft 0.e

Draft document

IEH-C under contract supporting the PHE COC Secretariat

June 2022

DRAFT

This is a draft opinion for discussion. It does not represent the views of the Committee and must not be quoted, cited or reproduced.

COMMITTEES ON CARCINOGENICITY, MUTAGENICITY AND TOXICITY OF CHEMICALS IN FOOD, CONSUMER PRODUCTS AND THE ENVIRONMENT (COC/COM/COT)

How do the Committees evaluate the relevance and reliability of data when assessing a chemical of concern?

Background

1. This document provides information on how the Committees on Carcinogenicity, Mutagenicity, and Toxicity (COC, COM, and COT) objectively evaluate study data and how they consider whether the information is both relevant and reliable. Readers are also referred to the Synthesis and Integration of Epidemiological and Toxicological Evidence subgroup (SETE) report and the preparatory discussion document, 'Biological Relevance and Statistical Significance' (CC/MUT/2020/03), which discuss in greater detail many of the concepts introduced here.

The Committee process

2. One of the key roles of the COC, COM, and COT is to evaluate whether chemicals that people may be exposed to in their daily lives can damage their health. The Committees are made up of Members who have expertise spanning a wide range of relevant fields, including biologists and toxicologists, pathologists, clinicians, epidemiologists and medical statisticians, as well as one or more Lay Members representing the interests of the public. The Committees adhere to the Nolan Principles of Public Life[1] and in doing so make their best efforts to consider all of the available evidence and provide advice that is both independent and transparent.

3. Questions on whether a chemical or exposure has the potential to cause adverse health effects are referred to the Committee from Government Departments and Agencies, and may come from professional scientific or health bodies and authorities, or from individuals, or be generated by the Committee itself. Where a specific evaluation is considered to be warranted, the Committee will endeavour to establish the likely adverse health effects associated with the chemical(s) or

[1] The Seven Principles of Public Life. https://www.gov.uk/government/publications/the-7-principles-of-public-life/the-7-principles-of-public-life--2

exposure(s) in question and to determine how these relate to the way in which people are exposed. For each new issue or topic, the Committee begins by defining the question to be addressed - 'problem formulation'. The types of question that might be tackled by the Committee include:

- 'Is the presence of chemical A in the environment likely to cause harm to the health of people in the general population?'
- 'Does the use of additive B in food products pose a risk to development of the fetus during pregnancy?'
- 'Is skin contact with product C linked with an increased risk of developing cancer?'

4. Addressing the question to hand begins with the identification of all available relevant data and information, with consideration of the objectivity and or/reliability of the data sources themselves. A substantial number of individual pieces of information may be gathered, and the totality of the information amassed (the 'evidence base') is then assessed and evaluated using a 'weight-of-evidence' approach. It is essential throughout, that the evaluation, interpretation and reporting of data is carried out in a fully objective way, i.e., without bias, judgment, or prejudice.

5. The European Food Safety Authority (EFSA) has published helpful guidance on use of the weight-of-evidence approach in scientific assessments (EFSA, 2017). In this document, EFSA notes three key steps in the weight-of-evidence assessment process: 'assembling the evidence', 'weighing the evidence', and 'integrating the evidence'. These three elements are described in detail in the following paragraphs.

Assembling the evidence

6. In general, information gathered will comprise peer-reviewed publications or other types of study reports describing findings from scientific and/or clinical studies, including dossiers provided by chemical and/or product manufacturers. This information is generally identified and assessed using a systematic process, with the aim to ensure that only relevant information is selected for evaluation and none is missed in so far as is possible.

7. Committees can use data from different types of studies to form the evidence base for addressing a particular question. This includes information taken from studies in individual humans or human populations ('clinical' or 'epidemiological' studies), laboratory animals ('*in vivo*' studies), or living or inert biological materials, for example cells maintained in culture or DNA extracted from biological samples ('*in vitro*' studies). Additional information may come from theoretical and/or computer-based evaluations of how a chemical might cause effects based on existing knowledge on similar types of chemicals ('*in silico*' studies). The Committees will also take into account information included in systematic reviews, meta-analyses, and opinion pieces published by authoritative bodies.

8. Using data from clinical or epidemiological studies can provide useful information about the potential human health impact of specific exposures and this avoids the uncertainty that may derive from experimental studies conducted in animals, where the biological make-up will differ to a greater or lesser extent to that of humans.

9. However, with human data it can be difficult to separate out the effect of the chemical under investigation from others that the individual is also exposed to. Studies using animals can generally be much more strictly designed and controlled than is possible with human studies, and this allows the possibility of obtaining clearer results. Animal studies can also allow for more extensive and detailed investigation of aspects such as how effects vary with the amount of exposure (the 'dose-response'), and the mechanisms by which an exposure causes biological damage or disease.

10. There are now, for ethical reasons, increasing efforts to reduce the use of animals in experimental studies (through a concept known as 'the 3Rs', namely replacement, reduction, and refinement) and sophisticated *in vitro* and *in silico* methods are being developed and validated for use wherever possible.

11. In assembling the evidence, the Committees will independently evaluate all data to reach their objective assessment of the evidence base. The sources of the data typically include study reports and peer-reviewed journal articles, as well as opinions from other authoritative bodies. The evaluation will consider study design, whether it has been conducted to Good Laboratory Practice (GLP), where the study is published, and any funding sources acknowledging the independence on contract research organisations. The Committees will be aware that in some instances positive results are more likely to get into the literature as investigators may either not wish to publish negative data or find that journals do not accept negative data.

Weighing the evidence

12. EFSA (2017) defines 'relevance' and 'reliability' as two major aspects to be considered when weighing evidence. These can be explained briefly as the contribution a piece of evidence would make to answering the question (relevance), and the extent to which the information being considered is valid and correct (reliability).

Relevance

13. Exposure to a chemical may result in changes ('biological effects') that affect the body at one or more different 'levels'; organs, tissues, cells, or individual molecules. The body has a substantial capacity to reverse or adapt to many such changes (through a process known as homeostasis), meaning that the majority of exposures that people experience during their lives will not lead to any adverse effects on health. However, in some cases changes may occur that cannot be

reversed or kept within the margins of normal body functioning, and which may eventually lead to negative impacts on health. Such adverse health effects might be caused directly by the exposure ('primary' effects) or may occur as a consequence of the initial changes induced by the exposure ('secondary' effects).

14. The ability of an exposure to cause biological effects depends not only on the type of exposure (i.e. the particular substance), but also on a number of other factors, including the amount of the substance to which the person is exposed, how they are exposed (for example, if the substance is swallowed, inhaled, or comes into contact with the skin) and for what duration and/or frequency that exposure occurs. Effects may also differ between different people, and for each person at different times during their life (e.g. during childhood, adolescence, pregnancy, in older age). In many cases, exposures below a certain level ('threshold') will be considered to be too low to be of human health concern.

15. When conducting scientific studies to evaluate whether exposure to a substance may produce harmful effects, the aim is to try to identify and discriminate between biological changes that signal a potential problem and those that would be considered to be normal or non-problematic. A critical part of this question is determination of the 'biological relevance' or 'biological importance' of an observed change; that is to say, to what extent does the effect observed represent an adverse change in terms of biological function. This concept can be extended to 'clinical relevance', that is, if an effect is considered to be of biological relevance, could it then lead subsequently to adverse effects on human health?

16. These are questions that need to be judged by people with expertise in the relevant fields, for example specialists in toxicology, pathology and immunology. The process of assessing and establishing biological and clinical relevance/importance is a key step in the evaluation of evidence. Consideration of study and historical control data will support identifying normal ranges for variation that may occur, in both animal and human studies.

Reliability

17. Although establishing the biological relevance of findings is very important, this is not the only aspect that needs to be taken into account when assessing study data; it is also important to look at how probable it is that the study findings are valid and dependable – i.e. to assess of the quality of the data.

18. In assessing study outcomes, it is necessary to determine whether observed changes are truly likely to have been caused by the exposure being investigated (or, conversely, whether a lack of changes genuinely indicates that the test exposure does *not* cause adverse effects). In addition, the study data, especially from the study controls, should also be considered in light of historical control data, i.e. do study control data sit within the historical range of control data for the species and strain used.

19. Of central importance here is whether the research is replicable; can the findings as reported in the study under consideration be reproduced? If the work has been replicated in other studies, this can provide a greater level of confidence that the findings are reliable.

20. The reliability or 'soundness' of research results may also be indicated by aspects such as whether a particular piece of research has been subjected to 'peer review' and has been conducted according to validated methodological and/or quality guidelines established by authoritative bodies, such as the Organisation for Economic Co-operation and Development (OECD), and whether it has been conducted to GLP. Consideration should also be made of whether the study design is appropriate to address the question asked of the Committee. Such considerations and assessments should aim to evaluate the impact of any identified study design limitations on the results and conclusions drawn.

21. Statistical analysis plays a crucial role in the wider evaluation of the reliability of a data set. Specific outcomes that Committees will want to be able to determine from a reported statistical analysis include a 'best' estimate of the size of the observed effect ('effect size') and the uncertainty (shown, for example, by confidence intervals) associated with the observation. The planning and design of a study is key to this, and needs to include determination of what size of effect would be considered to indicate a biologically relevant change and how large the study sample needs to be in order to be able to detect such effects reliably.

22. Over the years, it has become common practice for the results of statistical analyses to be reported in terms of '*P*-values'. A *P*-value may be defined as the probability of obtaining results at least as extreme as those actually observed, under the assumption that the null hypothesis[2] is correct.

23. It is commonplace for researchers to consider a *P*-value less than or equal to 0.05 ($P \leq 0.05$) to indicate a result that is 'statistically significant', and furthermore that this can be taken to support the finding of a genuine effect, or 'true' result. Some researchers use a *P*-value of 0.01 rather than 0.05 to signify statistical significance, while a *P*-value of 0.001 or less is sometimes considered to represent a result that is 'highly significant'. These 'cut-off' values for *P* are a matter of judgement or convention and are entirely arbitrary. P values may even differ for the same data depending on the assumptions made by the researcher, for example about how the distribution of data compares with a 'normal distribution'. It should also be noted that *P*-values can be used for a number of different types of statistical tests, e.g. trend tests and between group comparisons, all of which will differ in their outcomes.

24. Although many researchers continue to report study results solely in terms of *P*-values, this practice – which places a large emphasis on an arbitrary threshold value – has recognised limitations and should be avoided where possible. The

[2] The *null hypothesis* is the assumption that the treatment or factor of interest has no effect.

Committees support the reporting of a more complete spectrum of data obtained from a statistical analysis, including effect sizes and measures of uncertainty such as confidence intervals, as described above.

25. To ensure appropriate planning and statistical analysis of studies, scientists who conduct research should be well educated in statistical methods, their uses and their limitations. Also, when reporting study findings, the experimental results should be made available as 'raw data' so that they are available for analysis by other investigators. The Committees usually assess findings from statistical analyses as reported by the study investigators, but may also decide to conduct their own analyses if they consider that this will be useful and the raw study data are available.

26. The apparent significance of the results of a study as determined by statistical analysis should not be interpreted as representing the biological relevance or importance of the findings to the human population, whether in general or to a specific subpopulation; statistical 'significance' refers only to a result of the statistical analysis of the study in question and not the biological effect. Biological relevance and statistical significance are thus separate aspects that are both of key importance when making judgements about the results of a study. Such evaluations and judgments form an essential part of Committee deliberations and it is important that they are always clearly explained. Readers are referred to the EFSA document, 'Statistical Significance and Biological Relevance' for more detailed information.

Integrating the evidence

27. Following identification and weighing of the evidence, a full overall evaluation is carried out to objectively integrate the evidence, that is to combine all the information into a single overview. This helps the Committee use its expert judgement to reach an overall conclusion on the question being addressed, based on all the evidence available at the time the evaluation is carried out. This process is described in some detail in the SETE work mentioned in paragraph 1 of this document. The aim is to use the Committee's expertise to identify chemical exposures that are genuinely likely to present a human health hazard and to evaluate the nature and magnitude of the potential risk, to inform subsequent decisions taken by risk managers.

28. It is likely that new information will continue to become available beyond the date of the Committee's evaluation; for example, the results of new studies may be published. For this reason, Committees often keep a 'watching brief' on topics that have been evaluated, and as new information becomes available this can be integrated into the evidence base. However, a study or piece of evidence should not be taken to be of greater importance simply because it is new; as new information becomes available it must be weighed and considered in the same way as the earlier evidence, to become a contributing part of the full, available evidence base.

29. In their evaluation the Committee may also highlight data gaps, noting areas where information was not available and making suggestions for future studies.

This is a draft opinion for discussion. It does not represent the views of the Committee and must not be quoted, cited or reproduced.

Summary

30. The role of the COC, COM, and COT is to objectively evaluate whether chemicals to which people may be exposed in their daily lives can damage their health. The purpose of this document is to provide an overview of how the Committees carry this out and, in particular, how they evaluate the relevance and reliability of the data that are assessed.

31. The Committees will assemble the evidence objectively from an appropriate range of sources according to the question being considered. In this process they will note the funding sources, peer-review status and any issues of balance between positive and negative findings.

32. The subsequent assessment of the individual pieces of evidence incorporates the evaluation of two major aspects: relevance and reliability. This requires expert judgment; such evaluations and judgments form an essential part of Committee deliberations and should always be clearly explained.

33. Determining the 'biological relevance' or 'biological importance' of changes that are associated with exposure to a chemical involves establishing the extent to which observed effects represent meaningful and relevant changes in terms of biological function. Following from this, the concept of 'clinical relevance' relates to whether a biologically relevant effect could lead to adverse effects on human health.

34. It is of equal importance to establish whether the data evaluated are true and dependable, using statistical analysis to provide an objective measure on which to base conclusions. Committees will want to establish the size of an identified biologically and/or clinically relevant effect, and also the uncertainty associated with the observation.

35. Once all available pieces of evidence have been assessed, a full evaluation is carried out to objectively integrate the evidence. The aim is to reach a conclusion in response to the question posed and to note any areas where potentially useful data were lacking.

36. Committees often keep a watching brief on topics that have been evaluated previously. As new information becomes available, this can as required be assessed and integrated into the full evidence base using the same robust process as before.