Committee on TOXICITY

Committee on CARCINOGENICITY

Report of the Synthesising Epidemiological Evidence Subgroup (SEES) of the Committee on Toxicity and Committee on Carcinogenicity

Report of the Synthesising Epidemiological Evidence Subgroup (SEES) of the Committee on Toxicity and Committee on Carcinogenicity

Table of Contents

Executive summary

The Synthesising Epidemiology Evidence Subgroup (SEES) of the Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment (COT) and Committee on Carcinogenicity (COC) was set up in 2015. Its aim was to review and document current practice, given recent international and national development of methods by which evidence is synthesised, and to make recommendations for COT/COC guidance.

Human studies can provide direct evidence of health impacts of particular exposures. However, much of the evidence comes from observational epidemiological studies, where control of chance, bias (including exposure misclassification) and confounding may be problematic. Systematic review and meta-analysis are gold standard methods for combining epidemiological studies, but may not be available, or practical or possible to conduct for many of the questions considered by COT/COC.

Epidemiological reviews leading to statements or opinions in the last 10 years by COT/COC were identified and reviewed. A wide range of topics were identified relating to infant feeding, alcohol consumption, asbestos exposure, organophosphate exposure and vitamin E intake. The review methods used by the Committees varied by topic and requirement.

Evidence synthesis in the World Health Organization (WHO), the International Agency for Research on Cancer (IARC) and European Food Safety Authority (EFSA) was discussed and a number of well documented major systems for evidence synthesis were reviewed. These were:

- Systems initially designed for clinical medicine but now applied more widely, the Cochrane collaboration, GRADE (Grading of Recommendations Assessment Development and Evaluation) and SIGN (the Scottish Intercollegiate Guidelines Network). GRADE, with modifications, is being increasingly used in systematic reviews of environmental exposures.
- US Federal programmes, the National Toxicology Program (NTP)-OHAT, National Toxicology Program (NTP)-Report on Carcinogens and EPA-IRIS – these programmes were considered too time-consuming and resource intensive to be replicated in their entirety for COT/COC
- The Navigation Guide, first published in 2014, designed to speed up implementation of health protection measures for hazardous chemicals in the environment.

SEES considered evidence synthesis methodologies and tools available in order to draw up guidance points for scoping, conducting and reporting. For systematic reviews and meta-analysis, SEES recommended use of the Meta-analysis of Observational Studies in Epidemiology (MOOSE) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidance. Quality assessment of studies was considered an integral part of review. A large number of numerical scoring tools are available; the subgroup did not recommend any one tool and considered that if employed, these should be used (i) to aid narrative assessment rather than in place of it and (ii) can help direct sensitivity analyses of the meta-analysis e.g. by exclusion of low-scoring studies. Specific issues related to quantitative risk assessment and meta-analysis were identified, particularly around consideration of study heterogeneity. Documentation of uncertainty and of (potential conflict of) interests was considered important.

SEES also considered methods for combining epidemiological and toxicological evidence. These are less well developed than those for systematic review, particularly in a quantitative framework. There

are currently international initiatives in this area e.g. the Systematic Review and Integrated Assessment (SYRINA) and COT/COC will need to keep this methodological area under regular review.

1. Introduction

This document is an output from a joint subgroup of the Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment (COT) and the Committee on Carcinogenicity of Chemicals in Food, Consumer Products and the Environment (COC), that also included a member of the Committee on the Medical Effects of Air Pollutants (COMEAP). Members had extensive experience of UK and international scientific advisory committees. Details of the subgroup membership are given at the end of this document.

COT/COC generally review epidemiological evidence to (i) assess evidence for causality (hazard identification) and (ii) to determine appropriate dose-response estimates (hazard characterisation). It is hoped this initiative will prove of use to groups beyond COT, COC and COMEAP who are reviewing human non-experimental (epidemiological) studies. For example, the Committee on Mutagenicity of Chemicals in Food, Consumer Products and the Environment (COM), which from time to time reviews human data on genetic endpoints in its assessments, such as levels of chromosomal damage in studies comparing genetic damage in groups of humans with different levels of exposure.

Synthesis of epidemiological, toxicological and other evidence for risk assessment purposes is an integral part of the work conducted by scientific advisory groups. Toxicological studies provide mechanistic and experimental evidence of potential for causal associations and can form a basis of dose-response estimation if appropriate information is not available from human studies. However, toxicological studies are not always good predictors of impact of an exposure on the whole system in humans, including where biologic response in humans may be affected by concurrent other exposures (e.g. lifestyle factors, diet) or influenced by variability in toxicokinetics or the microbiome etc.

Risk assessment requires health based guidance values (HBGVs), to which exposure data can be compared. HBGVs are derived from a point of departure (POD) which can be a no observed adverse effect level (NOAEL), or a benchmark dose (BMD). The World Health Organization (WHO) have defined the NOAEL as the "greatest concentration or amount of a substance, found by experiment or observation, which causes no detectable adverse alteration of morphology, functional capacity, growth, development, or lifespan of the target organism under defined conditions of exposure" (FAO/WHO, 2009b). The BMD approach was developed as an alternative to the NOAEL. The full dose-response data is used in the statistical analysis and in this instance a POD is defined by the exposure level which produces a defined (non-zero) response level. This approach has the advantage of the "possibility to extrapolate outside of the experimental dose range and respond appropriately to sample size and the associated uncertainty" (FAO/WHO, 2009a). The BMD approach is increasingly used in preference to NOAEL and is the method recommended by EFSA. Appropriate uncertainty factors may then be applied to the POD to derive an HBGV.

The adequacy of uncertainty factors was considered in the COT's 2007 report on *Variability and Uncertainty in Toxicology of Chemicals in Food, Consumer Products and the Environment* (COT, 2007), which concluded that "Data from the available research in which compounds have been studied in both animals and man suggest that the default uncertainty factor of 10 allows adequately for interspecies differences." However, further considerations of developmental toxicity by COT at its December 2013 (COT, 2013) and May 2014 (COT, 2014) meetings led to conclusions that that the

10-fold uncertainty factor for interspecies variation in developmental toxicity was not adequate in all cases.

Epidemiological studies can provide direct evidence of human health impacts of particular exposures so interspecies uncertainty factors are not needed, but additional factors may be needed to account for other sources of variability. For risk assessment, human studies are used and preferred if available and of suitable quality. Experimental designs (for example, randomised controlled trials, intervention studies, natural experiments, chamber studies, food challenge) can be particularly powerful in evaluating dose-response. However, often epidemiological studies rely on observational designs (cohort, case-control, cross-sectional, descriptive). A cohort study, or case-control study nested within a cohort usually provides the most robust evidence. For any epidemiological study, it is important to carefully consider whether chance, bias and confounding might affect observed associations. A common problem encountered is uncertainty about the exposure characterisation. For risk assessment, it is also important to consider whether results are generalizable from the study population to the population for whom the risk assessment is being carried out (e.g. UK general population, UK babies and infants).

Systematic review is the formal optimal process to ensure all available evidence has been identified and rigorously assessed to provide the best estimate of the exposure-response relationship, but is resource-intensive and time-consuming. It is frequently used for clinical and epidemiological studies but can also be applied to toxicology (https://www.ncbi.nlm.nih.gov/pubmed/28501917) (Hoffman *et al.*, 2017). It is recognised that it is not always feasible for a scientific advisory committee to conduct a systematic review, for example if timeframes are short, resources are limited, and/or a systematic review has been recently published and only a short update is needed. However, scientific advisory committees need to be able to appraise quality of published systematic reviews and the methods used. Also, some methods in systematic review will be applicable in other forms of literature review e.g. documenting search terms and databases used.

Assessing causality

The majority of epidemiological studies relating to exposures from the environment and lifestyle including from food are observational rather than experimental in design – in most cases experimental studies would be unethical and natural experiments are rare. Where experimental studies are possible these are generally with low doses designed not to produce toxicity. Observational studies are usually regarded as showing associations rather than demonstrating cause and effect.

The conclusion as to whether an epidemiological association may be causal is therefore based on scientific judgement, considering epidemiological and other sources of information in a weight of evidence approach. Assessments are usually based on the Bradford Hill considerations, originally published in 1965 (Bradford Hill, 1965). These comprise strength, consistency, specificity, temporality, biological gradient, biological plausibility, coherence, experiment and analogy – not all need to be satisfied for causality to be met and absence of one or more is not proof of lack of causality. They should not be used as a checklist but to inform a weight of evidence approach. The Bradford Hill considerations are very extensively used, e.g. by the International Agency for Research on Cancer (IARC) (http://www.iarc.fr/) (Pearce *et al*., 2015), the US Office of Health Assessment and Translation (OHAT) of the US National Institute of Environmental Health Sciences

(http://ntp.niehs.nih.gov/pubhealth/hat/index.html) (Rooney *et al*., 2014), and the Grading of Recommendations Assessment, Development and Evaluation (GRADE) group (http://www.gradeworkinggroup.org/) (Schünemann *et al*., 2011).

IARC is the leading internationally recognised body assessing evidence of causality for carcinogenicity of chemical and other exposures. Opinions on carcinogenic and non-carcinogenic effects may be given as part of risk assessments for specific exposures by, for example, expert groups of the WHO e.g. Joint FAO/WHO Expert Committee on Food Additives (JECFA); OHAT; the Integrated Risk Information System (IRIS) of the US Environmental Protection Agency; the European Chemicals Agency (ECHA) Committee for Risk Assessment, the European Food Safety Authority (EFSA); and national committees such as COC, COT, COMEAP for the UK;. COT and COC may be asked to assess opinions and conclusions given by other bodies when conducting risk assessments for the UK population.

Quantifying exposure-response – single study or meta-analysis?

As in toxicology it is possible to set a dose-response function based on a single high quality epidemiological study. This is usually the one with the lowest point of departure if several studies are available and cannot be combined in a meta-analysis and/or a high-quality study that is well powered statistically to detect the elevation in risk under investigation (this may be hundreds or hundreds of thousands of subjects depending on the exposure). A single robust study has been used by COT for a risk assessment of arsenic in the infant diet (COT, 2016a), following a similar approach to that taken by JECFA (FAO/WHO, 2011). However, in their assessment of arsenic published prior to the study used by JECFA and the COT in their evaluations of inorganic arsenic, EFSA (EFSA, 2009) used a range of BMDL values calculated from different studies, rather than a single value or study, because none were considered particularly robust.

However, if several epidemiological studies are available and similar enough to be combined in a meta-analysis, this is usually preferred as (i) it helps increase power and precision (excess risks are often small, particularly where they relate to environmental exposures) and (ii) it provides better allowance for inherent bias and incomplete control for confounding than use of a single study, (iii) allows exploration of heterogeneity and quantifies variability.

For example, a single very large air pollution study (the American Cancer Society study following 500,000 individuals in the US) was used by COMEAP to provide dose-response estimates for chronic health effects of air pollution up to its 2009 report 'Long-term Exposure to Air Pollution: Effect on Mortality' (COMEAP, 2009), but more recent reports have used dose-response estimates based on systematic review and meta-analysis. (e.g. COMEAP, 2015; WHO, 2013). This provides a more robust evidence base and provides information on variability and therefore calculation of uncertainty intervals.

Aims & Objectives

The aim of this report was to review the approaches to synthesising epidemiological evidence that are used by COT and COC in chemical risk assessments and to make recommendations for COT/COC guidance. The objectives were:

- To review recent use of epidemiological evidence in committee statements and reports
- To provide an overview of initiatives and guidance of other groups of relevance to this topic

- To develop a systematic approach to reporting be used by COT and COC to improve transparency in committee conduct, taking into account the complexity and diversity of risk assessments conducted by COT and COC and the urgency of the work.

2. Recent epidemiological reviews considered and/or conducted by COT and COC

The systematic reviews and meta-analyses considered by the committees come from a variety of sources, including being conducted by the secretariat, commissioned by FSA or PHE and conducted by external contractors for review by the committee, or published in scientific literature.

Eight major reviews leading to statements or opinions carried out by COT and COC in 2008-2015 were identified and discussed by the subgroup, using a proforma to identify the type of review and methods (Table 1). The review illustrated the range of evidence assessments conducted by COT and COC committees – from evaluation of a recently published cancer prevention study (vitamin E and prostate cancer) or combination of case-reports (asbestos risks in children) to a series of extensive reviews (on risks arising from the infant diet) conducted by the secretariat and/or by consultants employing standardised methodology agreed in advance with the committee. Methods of review were not standardised across topics, but this would have been difficult given the heterogeneity of both topics and literature identified. Further, there was a mix of assessing reviews conducted by others, combining information from several reviews and reviews conducted in-house (i.e. secretariat and/or committee members). Quantitative assessment (i.e. stating the exposure-response relationship) was used where possible, but in several reviews this was not possible due to heterogeneity and/or study designs used. Where methods or uncertainty had not been described fully in statements, members of the SEES subgroup indicated that this had been part of the assessment but was not documented in the final statement.

The member from the UK Committee on the Medical Effects of Air Pollutants (COMEAP) noted that COMEAP also makes extensive use of systematic reviews in coming to conclusions. For many years, the Committee's work was supported by a Department of Health funded Air Pollution Epidemiology Database (APED) at St. George's, University of London which extracted key details from epidemiological time-series studies on air pollution on an ongoing basis (Anderson *et al*, 2007). This meant that any subsequent COMEAP need for a systematic review and meta-analysis, could be responded to more quickly (e.g. COMEAP, 2006). Population of the database involved systematic literature searching and screening for minimum quality criteria such as sufficient quantitative information to enable the calculation of standardized effect estimates; minimum time period of 1 year; some method of seasonal adjustment; some adjustment for temperature and analyses of effects in the general population rather than specific sub-groups. Information was extracted from the journal articles into many different fields to allow appropriate grouping for meta-analysis (e.g. ICD codes, type of health outcome, pollutant and averaging times) and for analysis of heterogeneity (e.g. WHO Region) (Anderson *et al*, 2007). Several further publications have arisen from this work on publication bias (Anderson et al, 2005) and on reviews of effects of specific pollutants (Atkinson et al, 2014a, 2014b, 2015; Mills et al 2015,2016; Walton et al, 2015).

Accessibility of past reviews of the committees was raised as being important. The subgroup noted that not all previous literature reviews included in the consideration by the subgroup were currently readily accessible on the COC website (potentially related to migration of websites), which was important if these were to be used as the basis of updating evidence. Publication of reviews undertaken on behalf of committees in peer-reviewed journals would be ideal and also useful for mid-career scientists, but this can be a lot of work and is unlikely to be feasible as routine practice.

However, there would potentially be scope to discuss with a journal regarding publishing overviews of committee work (e.g. with Occupational and Environmental Medicine).

Table 1. COT and COC reviews of epidemiological evidence in recent years

Year	COT/COC	Topic	Literature identification	Evidence synthesis	Uncertainty expressed?
2008	COT	Statement on the review of the 1998 COT recommendations on peanut avoidance. https://cot.food.gov.uk/cotstatements/cotstatementsyrs/cotstatements2008/cot200807peanut (Considered again in 2015-2016 COT reviews of infant feeding)	Described. British Nutrition Foundation review supplemented by BMA review and additional expert reviews carried out by individual experts for COT as COT found the systematic review alone insufficient.	Narrative review. SIGN scoring system	Mentioned
2013	COC	Relative Vulnerability of Children to Asbestos compared to Adults – Epidemiology and Case Reports on asbestos exposure in childhood and the risk of mesothelioma in later life. https://www.gov.uk/government/publications/relative-vulnerability-of-children-to-asbestos-compared-to-adults	Not described	Narrative summary – appropriate as evidence mainly relates to case-reports	Implied but not described
2014	COT	Statement on long-term neurological, neuropsychological and psychiatric effects of low- level exposure to organophosphates in adults https://cot.food.gov.uk/cotstatements/cotstatementsyrs/cotstatements2014/cotstatorg	Described. COT secretariat & Toxicology Unit at Imperial conducted the literature search, a working group reviewed the papers	Narrative summary with a description of each study – studies considered too heterogeneous for meta-analysis	Mentioned
2015 (a)	COC	First draft of statement on consumption of alcoholic beverages and risk of cancer – consideration of significance to public health.[1]	Described. Pubmed used from 2008 to identify studies published since last IARC review	Narrative including review of published meta-analysis. Study quality reviewed using modified Newcastle-Ottawa scale.	Mentioned

[1] The subgroup considered the first draft, not the final report.

		https://www.gov.uk/government/publications/consumption-of-alcoholic-beverages-and-risk-of-cancer			
2015 (b)	COC	Statement on vitamin E and the risk of prostate cancer https://www.gov.uk/government/publications/vitamin-e-and-the-risk-of-prostate-cancer	Not described. This was not a systematic review but explored the literature on vitamin E from human, animal and mechanistic data in order to determine whether the Selenium and Vitamin E Cancer Prevention Trial (SELECT) trial was plausible.	Narrative	Implied
2016 (b & c)	COT	(Three reviews) Review of risks arising from the infant diet and the development of atopic and autoimmune disease: Systematic review C Part I; review C Part II; review A (reserved business). Statement for Systematic review C Part1: https://cot.food.gov.uk/cotstatements/cotstatementsyrs/cot-statement-on-hydrolysed-cows-milk-formulae Combined statement for systematic reviews A and C Part II: https://cot.food.gov.uk/cotstatements/cotstatementsyrs/cot-statements-2017/statementonrevsaandc	Detailed description. Carried out by Imperial College consultants on behalf of FSA. Registered on PROSPEROs. CRD42013003802 – REVIEW A; CRD42013004239 – REVIEW B; CRD42013004252 – REVIEW C;	Meta-analysis where possible with methods set out in advance. PRISMA guidelines for interventions, MOOSE for observational studies, AMSTAR for systematic reviews	Detailed discussion of bias and strength of evidence expressed using GRADE

3. Systems for synthesising evidence

There are a number of major international and national established systems in use for synthesising evidence of relevance to the committees that were discussed by the subgroup, some of whom had participated in evidence synthesis using these systems (e.g. at IARC). These were:

- Cochrane collaboration
- GRADE Grading of Recommendations Assessment, Development and Evaluation
- SIGN
- National Toxicology Program (NTP)- Office of Health Assessment and Translation (OHAT)
- Navigation Guide
- US Environmental Protection Agency (EPA) IRIS
- National Toxicology Program (NTP)-Report on Carcinogens

It was noted that national and international bodies such as EFSA, IARC and WHO also produce guidance and some members of the subgroup had also participated in evidence synthesis at these bodies. A brief overview of evidence synthesis methods of these 10 systems and bodies is given in sections 3.1 and 3.2. The subgroup acknowledged that this was not exhaustive and that there were a number of other potentially useful guidance information and documents available e.g. from the University of York Centre for Reviews and Dissemination (https://www.york.ac.uk/crd/), which provides practical guidance for undertaking systematic reviews evaluating the effects of health interventions.

3.1 Evidence synthesis systems developed for clinical interventions

The longest established and best known evidence synthesis systems are those developed to evaluate and make recommendations on clinical interventions.

Cochrane collaboration

http://www.cochrane.org/

The Cochrane collaboration is the 'gold standard' system set up to synthesise evidence and produce recommendations to improve human health. It was originally set up to evaluate evidence in the field of clinical interventions and healthcare, but has wide relevance to evidence synthesis of experimental and observational epidemiological studies for non-healthcare related objectives. Awareness of methods and resources available is important to be covered within scientific advisory committee membership and secretariat.

The collaboration describes itself as "*a global independent network of researchers, professionals, patients, carers and people interested in health. We are a not-for-profit organization with contributors from more than 120 countries working together to produce credible, accessible health information that is free from commercial sponsorship and other conflicts of interest. We do this by producing reviews that summarize the best available evidence generated through research to inform decisions about health*." (http://uk.cochrane.org/about-us).

There is a detailed website with online training available and an online high quality handbook (Cochrane, 2011) that can be considered the most authoritative textbook for conducting systematic reviews currently available (see part 2 'General Methods for Cochrane reviews' in http://handbook.cochrane.org/).

The COT/COC subgroup noted that the Cochrane review group are (starting 2016-) developing guidelines for systematic review methods for public health, including nonhuman toxicology, relevant to improving evidence based regulation and guidance for environmental and occupational health policy, as well as drug and food safety (lead Dr Ellen Silbergeld, Johns Hopkins University).

GRADE (Grading of Recommendations Assessment, Development and Evaluation)

http://www.gradeworkinggroup.org/
GRADE is very widely used in evidence synthesis and has been used in some commissioned reviews by the committees. GRADE is formulated for a clinical setting and downgrades evidence based on observational (epidemiological) studies. This needs particular consideration if it is used in risk assessment of environmental exposures, which may need to rely on such studies.

The GRADE Working Group began in 2000 as an informal collaboration of people interested in addressing the shortcomings of present grading systems in health care – i.e. it was set up to evaluate evidence for and make recommendations on healthcare interventions, not to consider the type of environmental and lifestyle exposures usually considered by the scientific advisory committees covered by this report.

There is a GRADE handbook (GRADE, 2013) and detailed website that offers to "provide a guide for systematic review and health technology assessment authors, guideline panellists and methodologists on how to apply the GRADE methodology framework in more detail: GRADE evidence profiles, framing the question and deciding on important outcomes, rating the quality of evidence, risk of bias, publication bias, imprecision, inconsistency, indirectness, rating up, resource use, overall rating, Summary of Findings tables (binary) and (continuous), presentation of recommendations, and recommendation's direction and strength" (http://www.gradeworkinggroup.org/#pub).

Evidence is assigned one of four categories: HIGH, MODERATE, LOW or VERY LOW depending on the strength of evidence (Table A2) relying on a careful assessment of factors such as bias, inconsistency, precision and treatment of confounders. The interpretation of GRADE evidence assessments is that for HIGH level assessments, further research is very unlikely to change confidence in the estimate of effect; for MODERATE evidence further research is likely to have an important impact on confidence in the estimate of effect and may change the estimate; for LOW level evidence, further research is likely to have a very important impact on confidence in the estimate of effect and is likely to change the estimate. VERY LOW level evidence means that, although evidence is available, any estimate of effect is very uncertain (www.gradeworkinggroup.org/publications/Grading_evidence_and_recommendations_BMJ.pdf).
The GRADE assessment of evidence is widely used, including in some COT reviews e.g. of infant feeding, but of most usefulness when evaluating clinical interventions. Modifications for its use in environmental epidemiology are in development by the GRADE Environmental Health Project Group (papers in submission in 2018).

Very recently, adaptations of the GRADE system have been advocated to evaluate and integrate evidence from human, animal, *in vitro*, and *in silico* (computer modelling) studies when determining whether an environmental factor represents a potential health hazard or risk (Morgan *et al*., 2016).

Assessment of the hazard of exposures can produce analyses for use in the GRADE evidence-to-decision (EtD) framework to inform risk-management decisions about removing harmful exposures or mitigating risks, and this EtD framework allows for grading the strength of the recommendations.

The Guidelines Review Committee of the WHO (http://www.who.int/publications/guidelines/guidelines_review_committee/en/) has published a Handbook for Guideline Development (WHO, 2014a) that has adopted GRADE. Adaptations can be made to GRADE to serve a specific purpose, and a good example is the approach taken for the indoor air quality guidelines for household fuel combustion (WHO, 2014b) (http://www.who.int/indoorair/guidelines/hhfc/Evidence_review_methods.pdf?ua=1).

The Scottish Intercollegiate Guidelines Network (SIGN)

http://www.sign.ac.uk/ and http://www.sign.ac.uk/pdf/sign50.pdf (SIGN, 2015)
This system may be used in published reviews encountered by the committees, but is chiefly used for clinical guideline development.

SIGN was established in 1993 by the Academy of Royal Colleges and their Faculties in Scotland, to develop evidence based clinical guidelines for the National Health Service in Scotland. The methodological assessment of the literature evaluation is based on a number of criteria that focus on those aspects of the study design that have significant impact on risk of bias in the results reported and conclusions drawn.

The SIGN checklist for:

- systematic reviews is based on AMSTAR (considered below).
- Randomised controlled trials (RCTs) is based on an internal project (1997)
- Observational studies is based on MERGE (Method for Evaluating Research and Guideline Evidence) checklists
- Diagnostic accuracy studies is based on QUADAS (Quality Assessment of Diagnostic Accuracy Studies) programmes

SIGN uses a grading system for the studies used in systematic review based on study design, called levels of evidence. Scores range from 1++ for high quality meta-analyses to 4 for expert opinion (Table A3). The levels of evidence are used in grading the quality of evidence underpinning the recommendations in the clinical guidance (Table A4), with grading ranging from A (highest quality evidence e.g. based on RCTs) to D (e.g. based on expert opinion or case reports).

3.2 Evidence synthesis systems developed for environmental and lifestyle exposures

Three US Federal programmes were considered: National Toxicology Program (NTP)-OHAT, National Toxicology Program (NTP)-Report on Carcinogens and EPA-IRIS. These are described below, but were generally felt to be too time-consuming and resource intensive to be replicated within the scientific advisory committee setting, but that awareness within the committees of some of the methods used would be useful.

National Toxicology Program (NTP) - Office of Health Assessment and Translation (OHAT)

OHAT was established by the US National Toxicology Program (NTP) and the National Institute of Environmental Health Sciences to serve as an environmental health resource to the public and to regulatory and health agencies (Bucher, Thayer and Birnbaum, 2011). It conducts evaluations to

assess the evidence that environmental chemicals, physical substances, or mixtures (collectively referred to as "substances") may cause adverse health effects with an explicit focus on environmental health questions. It then provides opinions on whether these substances may be of concern given what is known about current human exposure levels. There are a number of papers available in Environmental Health Perspectives (e.g. “Evidence Integration for Literature-Based Environmental Health Science Assessments describing the approach”, (Rooney, 2014))

OHAT uses a seven-step framework for systematic review and evidence integration for reaching hazard identification conclusions: 1) problem formulation and protocol development, 2) search for and select studies for inclusion, 3) extract data from studies, 4) assess the quality or risk of bias of individual studies, 5) rate the confidence in the body of evidence, 6) translate the confidence ratings into levels of evidence, and 7) integrate the information from different evidence streams (human, animal, and “other relevant data” including mechanistic or in vitro studies) to develop hazard identification conclusions.

US Environmental Protection Agency (EPA) Integrated Risk Information System (IRIS)

IRIS is a human health assessment program that evaluates information on health effects that may result from exposure to environmental contaminants that is funded (at least up to 2016) by the US Environmental Protection Agency (EPA) http://www2.epa.gov/iris. It is used by the EPA and others to support decisions to protect human health.

It has the following steps:

1. Scoping and Problem Formation
    a. Identify needs of EPA program and regional offices
    b. Problem formulation – frame scientific questions specific to the assessment
    c. Draft Development
        Apply principles of systematic review to identify pertinent studies, evaluate study methods and quality, integrate evidence each health outcome, select studies for deriving toxicity values and finally to derive toxicity values
2. Review by scientists in EPA’s program and regional offices
3. Interagency Science Consultation – review by other federal agencies and Executive Office of the President
4. Public Comment and External Peer review
5. Revision of assessment
6. Final agency review and interagency discussion as in 3.

National Toxicology Program (NTP)-Report on Carcinogens

The biannual Report on Carcinogens (RoC) is a science-based public health report to the United States Secretary of Health and Human Services listing substances in the environment that pose a hazard to those living in the USA. It is mandated by the US Congress and prepared by the US National Toxicology Program (NTP). Substances for inclusion can be nominated and are then assessed. The NTP scopes the available evidence for a concept document and a RoC Monograph is prepared for those substances selected for evaluation. Substances are listed in the Report on Carcinogens as either “known to be a human carcinogen” or as “reasonably anticipated to be a human carcinogen”. (NTP, 2017)

A “Handbook for Preparing Report on Carcinogens Monographs” (NTP, 2015) provides detailed information on the methods including systematic reviews used to develop the Monographs. The approach is described as a “transparent process using systematic review methods guides the development of this report. Once candidate substances are selected, an extensive scientific review process begins with multiple opportunities for public comments. The review process also includes input from external scientific experts and government scientists from federal health and regulatory agencies.”

Detailed information on the review process is at http://ntp.niehs.nih.gov/go/727393. The Handbook notes that “It is anticipated that this handbook will be refined as new tools for conducting literature-based systematic reviews are developed.” At the time of writing, the most recent report, the 14th Report on Carcinogens released on November 3, 2016 included 248 listings of agents, substances, mixtures, and exposure circumstances that are known or reasonably anticipated to cause cancer in humans – this includes viruses as well as chemicals and metals and other exposures.

Navigation Guide

The Navigation Guide was developed by an interdisciplinary team from governmental and nongovernmental organisations and academia to address and shorten the time between scientific discovery of toxicity from chemicals in the environment and implementation of health protection measures. The Guide built on methods of research synthesis developed in clinical sciences, Cochrane and GRADE, and used by IARC and the US EPA. It aimed to provide a systematic and rigorous approach to research synthesis that would reduce bias and maximize transparency in the evaluation of environmental health information (Woodruff and Sutton, 2014).

There are four steps outlined in the Navigation Guide and steps 1-3 are applied to the different types of evidence (in vitro, in vivo and in silico and human observational studies), which are then combined.

1. Specify the study question
2. Select the evidence
3. Rate the quality and strength of the evidence
4. Grade the strength of the recommendations (modelled after GRADE)

Unlike clinical evidence synthesis, human observation studies are *a priori* assigned a ‘moderate’ quality rating, which is then upgraded or downgraded depending on *a priori* criteria. This is in contrast to systematic reviews in clinical sciences using, for example, Cochrane and GRADE, which generally assign an *a priori* rating to the body of human observational studies of “low quality”. Additional features identified are

- A protocol is developed prior to the review following a PECO – participants, exposure, comparator, outcomes – approach
- Standardised and transparent documentation including expert judgement
- Assessment of risk of bias
- A comprehensive and efficient search strategy
- Separation of science from values and preferences

The Guide's authors comment that conflicts of interests are not currently addressed in the system's assessments of risk of bias.

3.3 Use of evidence synthesis systems in other bodies

IARC

Shortly after IARC's establishment, its parent entity, the World Health Organization (WHO), asked IARC to prepare a list of agents known to cause cancer in humans (Pearce et al, 2015).

IARC assessments of carcinogenicity are based on evidence from epidemiologic studies, animal bioassays, pharmacokinetic/mechanistic experiments, and surveys of human exposure. The aim is to include all relevant papers on cancer in humans and experimental animals that have been published, or accepted for publication, in peer-reviewed scientific journals, and also any publicly available government or agency documents that provide data on the circumstances and extent of human exposure.

Evaluations involve consideration of all of the known relevant evidence from epidemiologic, animal, pharmacokinetic/mechanistic, and exposure studies to assess cancer hazard in humans (Tomatis, 2002).

The IARC classification categories are summarized below and given in more detail in Table A5 in the Appendix.

Group 1: In general, this category is used when there is *sufficient evidence of carcinogenicity* in humans.

Group 2: This category includes agents for which, at one extreme, the degree of evidence of carcinogenicity in humans is almost *sufficient*, as well as those for which, at the other extreme, there are no human data but for which there is evidence of carcinogenicity in experimental animals.

> *Group 2A:* The agent is probably carcinogenic to humans. For these agents there is limited evidence of carcinogenicity in humans and sufficient evidence of carcinogenicity in experimental animals.
>
> *Group 2B:* The agent is possibly carcinogenic to humans. For these agents there is limited evidence of carcinogenicity in humans and less than sufficient evidence of carcinogenicity in experimental animals.

Group 3: This category is used most commonly for agents for which the evidence of carcinogenicity is *inadequate* in humans and *inadequate* or *limited* in experimental animals.

Group 4: The agent is probably not carcinogenic to humans. This category is used for agents for which there is evidence suggesting lack of carcinogenicity in humans and in experimental animals.

EFSA

Outputs from a project in 2014-16, PROMETHEUS (Promoting methods for evidence use in scientific assessments), will be used by EFSA to improve their methods for handling data and evidence. POMETHEUS delivered 2 reports (EFSA, 2015; EFSA, 2016). The first report identified the principles and the processes for dealing with data and evidence in scientific assessments. The second reported on the analysis of methodological needs of EFSA.

EFSA defined the principles for dealing with data and evidence as: impartiality; excellence in scientific assessments; transparency and openness; and responsiveness. Based on these principles, EFSA defined the process for handling data and evidence in a scientific assessment, in four fundamental phases.

PROMETHEUS recognises that methods and process will need to be flexible to fit each assessment , but puts forward a stepwise "plan-conduct-verify-document-report" structure, that emphasises planning the strategy for the assessment as a key initial step.

1. Planning a strategy for the assessment upfront, before starting the assessment

The strategy for the assessment is defined upfront and includes the clarification of the assessment scope, conceptual framework definition and evidence needs and the way in which data and evidence should be dealt with. This includes collecting or extracting relevant data, validating or appraising evidence and analysing and integrating evidence. However, modifications to the strategy are acceptable. (EFSA, 2015)

2. Conducting the assessment in line with the strategy

The assessment should be carried out through completion of all the phases of the process and according to the agreed strategy as much as possible. This is to encourage the principles of impartiality and excellence in scientific assessments. Strategy modifications may be used if thoroughly documented and justified. (EFSA, 2015)

3. Verifying the process

The assessment should be continually verified as to whether the process was compliant with the planned strategy. Any deviations should be assessed. (EFSA, 2015)

4. Documenting and reporting the process, modifications to the strategy, results and conclusions, and ensuring accessibility of methods and data.

EFSA recommends that the process and its results be thoroughly and systematically documented and reported. This should be done for all steps of the assessment including assumptions made relating to uncertainty and also any strategy revisions. There are efforts in the scientific community to improve the quality of reporting and EFSA also has a number of projects with respect to this. Processes should be documented and reported along with results and conclusions, and accessibility of methods and data (without confidentiality violation). This is essential for transparency and verification of the scientific assessment process. (EFSA, 2015)

Expert judgement is fundamental in all phases of the process.

The second report (EFSA, 2016) identified the need for cross-cutting methodological development, training for staff and experts, instructions for applicants to integrate the existing regulatory frameworks and specialised data repositories. The analysis of the EFSA methodological needs would be updated in 4 years' time.

WHO

The World Health Organisation does not have a single approach to evidence synthesis in different topic areas but does have a common approach to guideline development including a chapter on systematic review (WHO, 2014a, pages 83-108).

4. Methods used in systematic literature review and quantitative synthesis

Initial problem formulation is important to determine resources needed to address the research question and scoping guidance was identified and is presented in section 6. A new extensive systematic review would not be necessary in many situations encountered by committees. However, published systematic reviews are commonly used by committees and an understanding of key elements of these is important. Also, some of the principles used for systematic review can also help inform reporting of more limited reviews.

4.1 Conducting and/or evaluating systematic reviews and meta-analyses

The two most widely accepted over-arching guidance systems for both conducting and evaluating systematic reviews and meta-analyses come from Meta-analysis of Observational Studies in Epidemiology (MOOSE) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).

PRISMA (http://www.prisma-statement.org/) aims to help authors improve the reporting of systematic reviews and meta-analyses. PRISMA may also be useful for critical appraisal of published systematic reviews, although it is not a quality assessment instrument to gauge the quality of a systematic review. A checklist of items is provided http://www.prisma-statement.org/documents/PRISMA%202009%20checklist.pdf (PRISMA Group, 2009)

MOOSE was developed following a workshop in 1997, led from the US Centers for Disease Control and Prevention with 27 US academic and government agency partners (Stroup *et al*., 2000). It provides a checklist with specifications for reporting of meta-analyses of observational studies in epidemiology, including background, search strategy, methods, results, discussion, and conclusion (https://www.editorialmanager.com/jognn/account/MOOSE.pdf). Key elements of MOOSE are:

- Reporting of background including a clear problem definition
- Reporting of search strategy –ensuring a comprehensive literature search for which search terms are included (if time is limited this may be limited to years following a published systematic review) and list of included and excluded papers.
- Reporting of methods including selection, classification, assessment of confounding, study quality and statistical methods
- Reporting of results including graphics, tables, sensitivity analyses, statistical uncertainty
- Report of discussion including assessment of bias, justification of exclusions, assessment of quality of included studies
- Reporting of conclusions including alternative explanations for results, generalisation of conclusions
- Disclosure of funding

A caveat to the use of these by COT/COC is that they are generic i.e. not specific to environmental and personal exposures that might be considered in COT and COC. This is of particular concern around exposure assessment e.g. if considering health effects of low-level pesticide exposure, was this: inferred as in an agricultural occupation; taken from self-reported working with undefined pesticides; self-report of working with specific pesticides; details of working with pesticides verified with occupational or farm records; contemporaneous biomarkers taken at relevant time windows.

However, MOOSE or a similar modified checklist would be helpful for COT/COC when conducting and assessing reviews and is recommended. A modified MOOSE checklist incorporating subgroup comments can be found in Appendix Table A1.

4.2 Quality assessment and use of numerical scoring tools

Quality assessment is an integral part of systematic review. However, there is no agreed 'gold standard' appraisal tool. The subgroup discussed a systematic review of tools for assessing quality and susceptibility to bias in observational epidemiological studies published in 2007 (Sanderson, Tatt and Higgins, 2007) which identified 86 tools, comprising 53 checklists and 33 scales. Most of these identified selection methods, measurements of study variables, sources of bias, confounder control and use of statistics, but the authors noted that distribution and weighting of domains across tools was variable and inconsistent and that half the tools did not describe their development, validity and reliability. Other scoring systems discussed by the subgroup were AMSTAR and the Newcastle-Ottawa Scale.

A systematic review should always assess quality and make transparent the quality criteria. However, the subgroup considered that a quantitative score of quality is a guide and may not always distinguish well between good and poor quality studies. It is not a replacement for expert opinion and, if used, needs to be used in conjunction with a narrative assessment of study strengths.

The subgroup therefore did not recommend regular use of a numerical scoring system for systematic reviews conducted by the committee, but acknowledged these could sometimes be useful in identifying good quality key studies and meta-analyses (e.g. a sensitivity analysis confined to higher scoring studies). If a scoring system has been used in a published systematic review, the method, its advantages and disadvantages, likely influence on the review and whether its use was appropriate should be discussed and documented.

The Strengthening Reporting of Observational Studies in Epidemiology (STROBE) statement (Vandenbroucke *et al.,* 2007) has helped improve the quality of reporting and transparency of epidemiological studies – some journals ask prospective authors to indicate how they have followed this reporting system as a condition of submission. The statement sets out a minimum set of recommendations for reporting, consisting of a set of 22 items covering cohort studies, case-control studies, and cross-sectional design studies. Adaptations are available for study areas with specific requirements, e.g. the STROBE Extension for Nutritional Epidemiology (STROBE-nut) (Lachat *et al.*, 2016), whilst useful to determine what items to cover, it was not designed as an instrument to evaluate the quality of observational research.

4.3 Specific issues for quantitative synthesis

Systematic review methods are standard and covered in statistical and epidemiological textbooks and online resources (e.g. http://handbook.cochrane.org/). However, for committee use, the following points were felt to be important for risk assessment:

- Consider if there is sufficient homogeneity of study design and outcomes to be able to combine studies in a meta-analysis (Der Simonian and Laird, 1986).
- Fully describe methods, especially whether using fixed or random effects models

- Take a decision on the level of unacceptable heterogeneity – and this may vary if heterogeneity can be readily explained
- Assess degree of (Huedo-Medina et al, 2006) and explore reasons for heterogeneity by stratification and by meta-regression where appropriate
- Include a graphical display for results e.g. a forest plot (Lewis and Clarke, 2001).
- Explore publication bias e.g. with a funnel plot (Light and Pillemer, 1984), noting that heterogeneity as well as publication bias can be responsible for an asymmetric funnel plot (Egger *et al.*, 1997). There are also statistical tests for publication bias (Egger *et al.*, 1997; Begg and Berlin, 1989).
- Trim and fill (Duval and Tweedie, 2000) can be useful to identify and correct for funnel plot asymmetry arising from publication bias. However, results need to be interpreted with caution as asymmetry in the funnel plot may represent true heterogeneity.
- Undertake sensitivity analyses to check how robust the findings are to differing assumptions.
- It may be important to distinguish between primary analysis conducted to address original problem formulation and secondary analysis.

Systematic reviews often include epidemiological studies of fundamentally different designs, such as cohort studies, case-control studies and randomised controlled trials. Cross-design methods for combining results from human studies of different designs have been developed for example, for matched and unmatched case-control studies (Moreno et al, 1996), case-control and cohort studies (Bhatia *et al.*, 1998), randomised clinical trials and observational studies (Prevost *et al.*, 2000). Bayesian methods of synthesis have also been developed that are sufficiently flexible to allow, if appropriate, for prior evidence and/or expert judgement (for example, on the relative appropriateness of certain types of evidence) to be incorporated into the analysis of the observed data (Sutton and Abrams, 2001).

4.4 Mixed approaches – quantitative synthesis of epidemiological and toxicological evidence

In assessing risks to human health from exposure to chemical substances in the environment, relevant evidence comes from both animal and human research. This is an evolving field that Committees need to keep up to date with. Toxicological data can be used to provide mechanistic information to support epidemiological findings, while human data can be used to validate evidence and extrapolations made from toxicological studies; combining both toxicological and human data helps in establishing causality (EFSA, 2017a). There has been relatively little exploration of methods to date for quantitative synthesis of evidence from human *and* animal studies, or even of toxicological studies alone (Roberts *et al.*, 2002b; Sandercock and Roberts, 2002). However, DuMouchel and Harris (1983) and DuMouchel and Groër (1989) have investigated alternative Bayesian models for combining dose-response slopes from animal and human studies. Current approaches usually consider epidemiological evidence separately from toxicological evidence, and then combine information at the end, but a common dose response is often difficult to establish.

Methodological issues in combining epidemiological and toxicological evidence

Two papers have investigated the potential usefulness of methods for combining human and animal data in human health risk assessment of exposure to environmental chemicals using the examples of (i) low birth weight and exposure to trihalomethanes (Peters at al 2005) (ii) assessment of the neurobehavioural effects associated with exposure to manganese (Peters et al 2008):

- The trihalomethane example identified 13 relevant studies (five epidemiological and eight toxicological, the latter including different species and animal strains). Issues that the authors had to resolve included: the use of odds ratios in the epidemiological studies the odds ratios for low birth weight which were adjusted for different covariates in each study compare with means (and standard deviations) of weight at each dose level in the toxicological studies; exposures were reported as parts per billion (ppb) in the epidemiological studies, but in the toxicological studies as mg/kg body weight/day. Study-specific dose-response slope estimates were obtained for each of the studies and synthesised using Bayesian meta-analysis models. Sensitivity analyses demonstrated that results were sensitive to the various assumptions made such as the choice of priors, defining the percentage of control group animals that were of low birth weight in the toxicological studies, the choice of dose-response model etc.

- The second example by Peters et al (2008) identified many more studies, 92 (55 human and 37 animal), potentially relevant to an assessment of the neurobehavioural effects associated with exposure to manganese. These studies were quite diverse covering a range of exposure routes (e.g. oral, inhalation, injection), species (e.g. humans, rats, rabbits, monkeys) and study design (e.g. occupational and environmental epidemiological studies). The types of neurobehavioural outcomes assessed and tests to measure them were also quite diverse in both the human and animal studies. Challenges in combining information from such diverse studies include different types of data available from (i) the epidemiological studies (mean scores and standard errors from a questionnaire for controls and exposed subjects; proportion of exposed and control subjects reporting negative activity symptoms; proportion of exposed subjects reporting negative activity symptoms; assessment of exposure in human population studies) (ii) the toxicological studies (activity scores; proportion of animals observed to have an adverse activity effect). For the activity data the authors explored the use of animal data as a prior for synthesis of the human data i.e. the relevance of the animal to the human data.

These two examples demonstrate that systematic review methods can offer improved transparency and structure in a risk assessment process, and that meta-analysis methods, particularly the more flexible models incorporating judgements on relevance of certain types of evidence, have potential for use in this context. They also show how effects in different species can be compared.

This approach needs investigating in other examples to identify assumptions and issues that are of general importance and those only relevant to specific examples and to investigate the influence of incorporating additional relevant information such as different routes of exposure, different types of exposure (e.g. individual chemicals vs. mixtures of chemicals) and available data on biological effects and mechanisms.

Weight of Evidence approach

A weight of evidence (WoE) and/or systematic review approach is used for chemical risk assessment in the European Union. Nine regulatory frameworks were reviewed in a paper published in 2016 (Agerstrand and Beronius, 2016), of which four (the REACH regulation, the Biocides directive, the Cosmetics regulation and the regulation for Classification, Labelling and Packing (CLP)) explicitly mention WoE, while other frameworks include this in guidance (including that used for food contaminants by EFSA). However, the 2016 review concluded that there was limited guidance in the frameworks on how to perform WoE syntheses and that this could be improved, using guidance from the European Commissions' Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR), IARC and the Navigation Guide.

EFSA published guidance on the weight of evidence approach in 2017 (EFSA, 2017b), giving a number of examples. The approach should be used where more than one piece of evidence is used to answer a scientific question. In brief, the assessment consists of the following three steps, (1) assembling the data into lines of evidence of similar type, (2) weighing the evidence and (3) integrating the evidence. Weighing includes assessment of how applicable the evidence is (includes relevance and reliability), the quality of the evidence and how consistent it is. The guidance considers four weighing methods, two qualitative and two quantitative. Qualitative methods are best professional judgement (e.g. narrative systematic review) and causal criteria (e.g. Bradford Hill considerations). Quantitative methods are rating (e.g. GRADE, IARC and OHAT mentioned in this document (section 3.1)) and quantification (e.g. statistical models including regression, meta-analysis, meta-regression, Bayesian models, machine learning, *in silico* tools including QSAR). One or more weighing method may be used in the WoE. Integrating the evidence in the EFSA guidance has the following steps: considering the conceptual model for integrating the evidence; assessing the consistency across different lines of evidence for which a hierarchical approach is suggested, starting with evidence lines that are closely related; if an expert judgment step is required, using appropriate procedures (e.g. formal expert knowledge elicitation). There is no explicit guidance on how to integrate toxicological and epidemiological lines of evidence.

International work on methods to combine human and toxicological data

Work is in progress on this at an international level e.g. under discussion in the GRADE Environmental Health Project Group (as of 2018) as well as in specific research areas. For example, a proposed framework for the systematic review and integrated assessment (SYRINA) of endocrine disrupting chemicals was published in 2016, which included authors involved in the Navigation Guide and from the US EPA, IARC and university departments in a number of countries (Vandenberg *et al.*, 2016). The World Cancer Research Fund is funding research into methods for reviewing mechanistic evidence (http://www.wcrf.org/int/research-we-fund/continuous-update-project-cup/mechanisms-research). This has developed a systematic review protocol integrating evidence from human, animal and other mechanistic studies to aid in situations where for example, systematic review of human observational studies is suggestive but not conclusive of an effect and insight may be obtained from systematic review of mechanistic studies. The initial work was on studies linking diet, nutrition and physical activity to cancer (e.g. milk intake and prostate cancer). There do not appear to be any publications to date (March 2018) but presentations are available online e.g.

http://www.slideshare.net/wcrf/the-continuous-update-project-novel-approach-to-reviewing-mechanistic-evidence-on-diet-nutrition-physical-activity-and-cancer

4.5 Expressing uncertainty in the findings

This was considered important by subgroup, but there is no 'gold standard' method of doing so.

Expression of uncertainty have been addressed in different ways in previous committee systematic reviews (see Section 2) e.g. the systematic reviews provided to COT on infant feeding used the GRADE rankings of HIGH, MODERATE, LOW or VERY LOW and also considered *post hoc* trial sequential analysis (TSA) to quantify statistical reliability of findings graded as MODERATE or HIGH (TSA quantifies statistical reliability of data in a cumulative meta-analysis in a similar way to an interim analysis in a single randomized clinical trial).

The COT narrative review on low-level organophosphates, where studies were too heterogeneous for meta-analysis, expressed uncertainty in a narrative format e.g. 'There is uncertainty as to whether long-term exposure to organophosphates causes detectable impairment of sensory thresholds, but if there is an effect then it is likely to be small'.

COT considered uncertainty in 2010 through a workshop and report, *Assessment of the COT uncertainty framework from a social science perspective: A theoretical evaluation* (Rowe, 2010). The conclusions of the report were:

- People are not good at understanding and using uncertainty estimates of verbal or numeric form
- Context in which an uncertainty is expressed (for example what the uncertainty is about, the situation in which it is being given, who is expressing the uncertainty) plays an important role in how people understand terms
- Findings relate to all people, including experts, although some might favour presenting uncertainty information in, for example, numerical terms rather than verbal in certain cases
- The COT framework should probably not endorse a verbal means for communicating uncertainty because of differential and inconsistent interpretation within and between people.

EFSA also considered expression of uncertainty in draft guidance in 2016. The COT considered they should wait to see how this is implemented during its trial phase before deciding on whether to change current practice.

5. Assessment and reporting of potential conflict of interest

Conflicts of interest may affect conduct and interpretation of systematic reviews and synthesis of evidence (Jørgensen, Hilden and Gøtzsche, 2006), while perceived conflicts of interest may affect confidence in the findings. To address this, scientific advisory committees have a published approach on annual declarations of interest, which are available on committee websites (e.g. the Food Standards Agency, which provides the secretariat for COT provides the following guidance https://www.food.gov.uk/news-alerts/consultations/our-approach-to-managing-the-interests-of-its-external-scientific-advisers(Annex B)). Additionally, declarations are requested before each agenda item at meetings and recorded in minutes; if the interest potentially constitutes a conflict, members are not allowed to participate in discussions. It should be noted that committees such as the COC, COM and COT expect that their members will have interests to declare because these will reflect their breadth of experience and is partly why they were appointed as experts to these committees.

There are multiple types of conflict of interest, which include direct and indirect financial support, acting as an expert witness on a topic and entrenched beliefs (which may also form interests rather than conflict of interests). Attention also needs to be given to assess potential conflicts of interest if relying on external reviews, but these may be more difficult to identify and evaluate.

Methods used by committees to synthesise evidence can come under intense scrutiny that result in real or perceived conflicts of interest being publicly highlighted. For example, the US Dietary Guidelines Advisory Committee (Advisory Committee) review of US dietary guidelines published in 2015 (Dietary Guidelines Advisory Committee, 2015) was criticised for reliance on external systematic reviews such as from the American Heart Association and the American College of Cardiology, who report 20% and 38% of revenue from industry – although no evidence was presented to suggest that there was any attempt to directly affect the reviews (Teicholz, 2015).

6 Guidance on epidemiological evidence synthesis

As a result of the previous considerations, the subgroup identified the following over-arching guidance on epidemiological evidence synthesis. It was recognised that questions considered by committees are varied and that it was therefore not possible to recommend a single evidence synthesis method. For example, in some situations (e.g. establishing a TUL of a nutrient) case-reports in humans may provide the most valuable information, whereas established epidemiological evidence synthesis systems (usually set up with respect to clinical interventions) regard case-reports as the lowest quality of evidence.

6.1 Scoping and problem formulation

The first step in the process of evidence synthesis is scoping and problem formulation. This helps make efficient use of resources and to identify the best method in a given situation. The following points should be considered.

- Why is a review of epidemiological evidence needed?
- Is a systematic review required?
    - How quickly is the review needed? Quick advice will require limited literature search and/or use of an existing review. Long-term important issues may merit investment in a new or updated systematic review.
    - What is the importance of the issue and consequences of Committee advice? The greater the importance, the more likely a systematic review will be needed.
    - Is qualitative information about hazard enough, or does risk need to be quantified? The latter is more likely to require systematic review to ensure all relevant papers are identified.
    - Is there another recent review available in the literature or by a reputable body e.g. IARC, WHO, EFSA?
        - If yes, can this be used? Is it systematic and good quality?
        - Does the review need updating only, or does it need to be redone?
        - Is the review missing older literature that could be valuable?
        - Was the risk estimate identified justified?
        - Does an existing meta-analysis need updating?
        - Does a meta-analysis need to be conducted or will forest plots be enough? [Extraction of data is time-consuming and although statistical analysis itself is relatively straightforward, meta-regression may also be needed to account for study differences]

6.2 Overarching principles

- An established system or guideline should be followed where appropriate (e.g. for a systematic review) and this should be stated in publications or reports.
- The evidence synthesis would usually include an expression of uncertainty in the findings.
- Potential conflicts of interest should be considered, including for published reviews.

6.3 Limited literature search

This might be needed if updating an existing meta-analysis or a quick review of literature is required. As a minimum this should include:

- Purpose of search (e.g. to identify papers on iodine toxicity in children published since an EFSA review in *date*)
- Database searched (e.g. PubMed)
- Time period covered by search (e.g. 2015 to March 2017)
- Search terms (e.g. iodine excess children, iodine toxicity)
- Numbers of papers identified, and numbers included in the review
- Reasons for exclusion of papers
- Extraction of key information from papers in narrative, graphical and/or tabular format
- Discussion and conclusion

6.4 Evaluating an existing systematic review

- As a minimum, an adapted checklist from MOOSE (Appendix A1) should be consulted when evaluating systematic reviews and meta-analyses and this should be referred to in reports.
- Committees should explicitly discuss and document the evidence synthesis methodology and any scoring system used in reviews to be aware of how these affect inferences e.g. systems developed to synthesise evidence from clinical trials such as GRADE give lower weight to evidence from non-experimental (observational) studies, even if studies are very large, high quality and consistent; systematic review methods that exclude studies with zero cases may introduce bias.
- If a scoring system has been used in a published systematic review, the method, its advantages and disadvantages, likely influence on the review and whether its use was appropriate should be discussed and documented.
- The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist[2] (Moher *et al.*, 2009) may also be useful to assist in critical appraisal of published systematic reviews, bearing in mind it was not specifically designed to be a quality assessment instrument to gauge the quality of a systematic review.

6.5 Conducting systematic review

- An adapted checklist from Meta-analysis of Observational Studies in Epidemiology MOOSE guidelines (Appendix Table A1) should be used when conducting systematic reviews and meta-analyses and this should be referred to in reports. To better cover the type of evidence synthesis conducted by the committee, the following elements have been added to the published MOOSE checklist:
    - Include a flow chart for identification of papers in systematic review
    - Adequate presentation of study data – descriptive paragraphs and/or in tables
    - Description of data extraction
    - Use of a forest plot to illustrate findings from the studies reviewed

[2] http://www.prisma-statement.org/documents/PRISMA%202009%20checklist.pdf

    - Consideration of patterns of association and confidence intervals are preferred to the use of conventional statistical significance ($p<0.05$) to determine evidence or absence of proof of an association.
- Quality assessment and use of scoring systems
    - The quality of papers and reviews should always be assessed but this does not necessarily need to involve a numerical scoring system.
    - Use of a numerical scoring system does not replace the need for narrative assessment of quality.
    - Scoring systems may be helpful to identify good quality key studies, especially for use in meta-analyses (e.g. a sensitivity analysis confined to higher scoring studies).
    - There are a lot of numerical scoring systems in use and the committee did not make any specific recommendations.
    - The Strengthening Reporting of Observational Studies in Epidemiology (STROBE) statement (Vandenbroucke *et al.*, 2007) provides a useful number of areas to consider when evaluating a study, but it was developed as a tool to improve reporting and transparency and not as an instrument to assess quality.

6.6 Conducting quantitative synthesis

For quantitative synthesis, the following elements are important to include (these have been added to the adapted MOOSE checklist, Appendix A1):

- Consider if there is sufficient homogeneity of study design and outcomes to be able to combine studies
- Fully describe methods, especially whether using fixed or random effects
- Take a decision on the level of unacceptable heterogeneity – and this may vary if heterogeneity can be readily explained
- Explore reasons for heterogeneity by stratification and by meta-regression where appropriate
- Include a graphical display for results
- Explore publication bias e.g. with a funnel plot

6.7 Reporting

- Methods used, even for limited literature review, should always be documented, in particular the databases searched, the detail of search terms used and the papers identified.
- The PRISMA checklist[3] (Moher *et al.*, 2009) should be followed for reporting of systematic reviews and meta-analyses.

[3] http://www.prisma-statement.org/documents/PRISMA%202009%20checklist.pdf

7. Conclusions

A review of COT and COC epidemiological evidence synthesis in recent years confirmed the opinion that the UK scientific advisory committees consider a number of very different topics and scientific questions. There are already a large number of existing systems and methodologies to synthesise epidemiological evidence, with methodologies that can be adapted for Committee use. Members therefore did not consider a need to develop a UK-specific new system to synthesise epidemiological evidence. Keeping informed about methodology development on evidence synthesis is important as this area is currently undergoing rapid development, especially with respect to consideration of environmental exposures and of synthesis of evidence from epidemiological and toxicological data.

Guidance points for Committees and their secretariats when conducting epidemiological review have been formulated using existing guidance and a checklist for meta-analyses and systematic reviews of observation studies, modified from MOOSE guidelines. Guidance is deliberately short to maximise uptake and use. These need to be considered by the Committees and, if adopted, their use should be evaluated.

A separate report on SEES methods of working and wider recommendations for COT (e.g. on secretariat training) has also been prepared.

Appendix

Table A1. COT/COC checklist for meta-analyses and systematic reviews of observation studies, modified from MOOSE guidelines (items in italics not relevant in all searches)

Section		Present✓/×
Introduction	**Present?**	
	The study question	
	The hypothesis under test	
	Statement of objectives: study population, exposure, outcomes	
Sources	**Described?**	
	Qualifications of searchers (librarians/researchers)	
	Search strategy – time period, keywords	
	Databases and registries searched	
	Search software (name, version, special features e.g. explode term)	
	Use of hand searching (of references of papers identified)	
	Other efforts to include all studies e.g. contact with authors	
	List of citations included and excluded (with justification)	
	Method of addressing articles not in English	
	Method of handling abstracts and unpublished studies	
	A flow chart describing identification of papers	
Methods	**Described?**	
	Types of study designs considered and included/excluded	
	Relevance & appropriateness of studies to answer study question	
	Rationale for selection and coding of data	
	Documentation of how data were extracted, classified and coded (*including if more than one person extracted data, blinding, inter-rater reliability if assessed*)	
	Explicit description of exposure assessment methods	
	Confounding dealt with, appropriate confounder adjustments in analyses	
	Assessment of bias (e.g. comparability of cases and controls)	
	Assessment of study quality (*blinding of quality assessors?, stratification or regression by study quality parameters*)	
	Documented if studies are not homogeneous enough (in design, exposures or outcomes), or there are not enough studies (usually <3) to be able to proceed to meta-analysis	
	Statistical methods fully described including whether using fixed on random effects	
	Decision on level of unacceptable heterogeneity in meta-analyses	
Results	**Present?**	
	A graph (usually Forest plot) summarizing individual study estimates and overall estimate	
	Assessment of heterogeneity in meta-analyses (how much? can it be explained? Explore reasons for heterogeneity by stratification and meta-regression where appropriate)	
	Adequate presentation of each included study in descriptive paragraphs and preferably also a more detailed table	
	Results of sensitivity testing e.g. subgroup analysis	
	Consideration of patterns of association and confidence intervals rather than solely using statistical significance (e.g. p<0.05) to determine proof/absence of an association	
	Indication of statistical uncertainty of findings	
	Explore publication bias e.g. with a funnel plot	
Discussion	**Described?**	
	Strengths and weaknesses of studies	
	Potential biases in the review e.g. publication bias that may affect conclusions	
	Justification for exclusion (e.g. citations not in English)	
	Assessment of quality of included studies	
	Consideration of alternative explanations for observed results	
	Discussion of generalisability of the conclusions	
	Guidelines for future research	
	Disclosure of funding source and potential conflicts of interest	

Adapted from: Stroup, DF.; Berlin, JA.; Morton, SC.; Olkin, I.; Williamson, GD.; Rennie, D.; Moher, D.; Becker, BJ.; Sipe, TA. and Thacker, SB. (2000) 'Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group' *Journal of the American Medical Association* 283(15) pp.2008-2012 Available at: http://www.ncbi.nlm.nih.gov/pubmed/10789670

Table A2. GRADE Levels of Evidence

Quality of evidence	Definition	Examples of when this is the case
High	Further research is very unlikely to change our confidence in the estimate of effect	Several high-quality studies with consistent results
Moderate	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate	One high-quality study Several studies with some limitations
Low	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate	One or more studies with severe limitations
Very low	Any estimate of effect is very uncertain	No direct research evidence One or more studies with very severe limitations

Table A3. SIGN Levels of Evidence

1++	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
1+	Well-conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
1-	Meta-analyses, systematic reviews, or RCTs with a high risk of bias
2++	High quality systematic reviews of case control or cohort or studies High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
2+	Well-conducted case control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal
2-	Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, e.g. case reports, case series
4	Expert opinion

Table A4. SIGN grading of recommendations

Grading	
A	At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; or A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results

B	A body of evidence including studies rated as 2++, directly applicable to the target population, and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+

Table A5. IARC classification categories

Group 1	The agent is *carcinogenic* to *humans*	This category is used when there is *sufficient evidence of carcinogenicity* in humans. Exceptionally, an agent may be placed in this category when evidence of carcinogenicity in humans is less than *sufficient* but there is *sufficient evidence of carcinogenicity* in experimental animals and strong evidence in exposed humans that the agent acts through a relevant mechanism of carcinogenicity.
Group 2		This category includes agents for which, at one extreme, the degree of evidence of carcinogenicity in humans is almost *sufficient*, as well as those for which, at the other extreme, there are no human data but for which there is evidence of carcinogenicity in experimental animals. Agents are assigned to either Group 2A (*probably carcinogenic to humans*) or Group 2B (*possibly carcinogenic to humans*) on the basis of epidemiological and experimental evidence of carcinogenicity and mechanistic and other relevant data. The terms *probably carcinogenic* and *possibly carcinogenic* have no quantitative significance and are used simply as descriptors of different levels of evidence of human carcinogenicity, with *probably carcinogenic* signifying a higher level of evidence than *possibly carcinogenic*.

Group 2A	The agent is *probably carcinogenic* to *humans*	This category is used when there is *limited evidence of carcinogenicity* in humans and *sufficient evidence of carcinogenicity* in experimental animals. In some cases, an agent may be classified in this category when there is *inadequate evidence of carcinogenicity* in humans and *sufficient evidence of carcinogenicity* in experimental animals and strong evidence that the carcinogenesis is mediated by a mechanism that also operates in humans. Exceptionally, an agent may be classified in this category solely on the basis of *limited evidence of carcinogenicity* in humans. An agent may be assigned to this category if it clearly belongs, based on mechanistic considerations, to a class of agents for which one or more members have been classified in Group 1 or Group 2A.
Group 2B	The agent is *possibly carcinogenic* to *humans*	This category is used for agents for which there is *limited evidence of carcinogenicity* in humans and less than *sufficient evidence of carcinogenicity* in experimental animals. It may also be used when there is *inadequate evidence of carcinogenicity* in humans but there is *sufficient evidence of carcinogenicity* in experimental animals. In some instances, an agent for which there is *inadequate evidence of carcinogenicity* in humans and less than *sufficient evidence of carcinogenicity* in experimental animals together with supporting evidence from mechanistic and other relevant data may be placed in this group. An agent may be classified in this category solely on the basis of strong evidence from mechanistic and other relevant data.
Group 3	The agent is *not classifiable as to its carcinogenicity to humans*	This category is used most commonly for agents for which the evidence of carcinogenicity is *inadequate* in humans and *inadequate* or *limited* in experimental animals. Exceptionally, agents for which the evidence of carcinogenicity is *inadequate* in humans but *sufficient* in experimental animals may be placed in this category when there is strong evidence that the mechanism of carcinogenicity in experimental animals does not operate in humans. Agents that do not fall into any other group are also placed in this category.

Group 4	The agent is probably not carcinogenic to humans.	This category is used for agents for which there is evidence suggesting lack of carcinogenicity in humans and in experimental animals. In some instances, agents for which there is inadequate evidence of carcinogenicity in humans but evidence suggesting lack of carcinogenicity in experimental animals, consistently and strongly supported by a broad range of mechanistic and other relevant data, may be classified in this group.

Adapted from: Pearce, N.; Blair, A.; Vineis, P. *et al*. (2015) 'IARC Monographs: 40 Years of Evaluating Carcinogenic Hazards to Humans' *Environmental Health Perspectives* 123(6) pp.507-514 Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4455595/

Reference List

Agerstrand, M. and Beronius, A. (2016) 'Weight of evidence evaluation and systematic review in EU chemical risk assessment: Foundation is laid but guidance is needed' *Environment International* 92-93 pp 590-596. Available at: https://www.ncbi.nlm.nih.gov/pubmed/26682868

Anderson, HR.; Atkinson, RW.; Peacock, JL.; Sweeting, MJ. and Marston, L. (2005) 'Ambient particulate matter and health effects—publication bias in studies of short-term associations' *Epidemiology* 16 pp 155–163. Available at: https://www.ncbi.nlm.nih.gov/pubmed/15703529

Anderson, HR.; Atkinson, RW.; Bremner, SA.; Carrington, J. and Peacock, J. (2007) 'Quantitative Systematic Review of Short Term Associations between Ambient Air Pollution (Particulate Matter, Ozone, Nitrogen Dioxide, Sulphur Dioxide and Carbon Monoxide), and Mortality and Morbidity' *Corporate report to the United Kingdom Department of Health.* Available at: www.gov.uk/government/publications/quantitative-systematic-review-of-short-termassociations-between-ambient-air-pollution-particulate-matter-ozone-nitrogen-dioxide-sulphur-dioxide-andcarbon-monoxide-and-mortality-and-morbidity

Atkinson, RW.; Kang, S.; Anderson, HR.; Mills, IC. and Walton, HA. (2014a) 'Epidemiological time series studies of PM2.5 and daily mortality and hospital admissions: a systematic review and meta-analysis' *Thorax* 69(7) pp 660-5. Epub 2014/04/08. Available at: https://www.ncbi.nlm.nih.gov/pubmed/24706041

Atkinson, RW.; Mills, IC.; Walton, HA.; Kang, S. and Anderson, HR. (2014b), editors 'Systematic Review and Quantitative Meta-analysis of the Evidence for Associations between Chronic and Short-term Exposure to Outdoor Air Pollutants and Health' *Department of Health Policy Research Programme Project: 002/0037* Available at: http://www.sgul.ac.uk/images/Research/Pop_Health/DHreport_chronic_shortterm_exposure_to_outdoor_pollutants092015.pdf

Atkinson, RW.; Mills, IC.; Walton, HA. and Anderson, HR. (2015) 'Fine particle components and health--a systematic review and meta-analysis of epidemiological time series studies of daily mortality and hospital admissions' *Journal of Exposure Science and Environmental Epidemiology.* 25(2) pp 208-14. Epub 2014/09/18. Available at: https://www.ncbi.nlm.nih.gov/pubmed/25227730

Bhatia, R.; Lopipero, P. and Smith, AH. (1998) 'Diesel exhaust exposure and lung cancer' *Epidemiology* 9 pp 84-91 Available at: https://www.jstor.org/stable/3702619?seq=1#page_scan_tab_contents

Begg, CB. and Berlin, JA. (1989) 'Publication bias and dissemination of clinical research' *Journal of the National Cancer Institute* 81 pp 107–115. Available at: https://www.ncbi.nlm.nih.gov/pubmed/2642556

Bucher, JR.; Thayer, K. and Birnbaum, L. (2011) 'The Office of Health Assessment and Translation: A Problem-Solving Resource for the National Toxicology Program' *Environmental Health Perspectives* 119 ppa196-a197 Available at: https://ehp.niehs.nih.gov/1103645/

COC (2013) 'Statement on the relative vulnerability of children to asbestos compared to adults' Available at: https://www.gov.uk/government/publications/relative-vulnerability-of-children-to-asbestos-compared-to-adults

COC (2015a) 'Statement on consumption of alcoholic beverages and risk of cancer' Available at: https://www.gov.uk/government/publications/consumption-of-alcoholic-beverages-and-risk-of-cancer

COC (2015b) 'Statement on vitamin E and the risk of prostate cancer' Available at: https://www.gov.uk/government/publications/vitamin-e-and-the-risk-of-prostate-cancer

Cochrane (updated March 2011) Cochrane Handbook for Systematic Reviews of Interventions - Version 5.1.0 Higgins, JPT. and Green, S. (eds) Available at: http://handbook.cochrane.org/

Cochrane (Copyright 2018. The Cochrane Collaboration) Cochrane Community Glossary. Available at: https://community.cochrane.org/glossary

COMEAP (2006) 'Cardiovascular disease and air pollution' Available at: https://www.gov.uk/government/publications/comeap-cardiovascular-disease-and-air-pollution

COMEAP (2009) 'Long-term Exposure to Air Pollution: Effect on Mortality' Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/304667/COMEAP_long_term_exposure_to_air_pollution.pdf

COMEAP (2015) 'Quantification of mortality and hospital admissions associated with ground-level ozone' Available at: https://www.gov.uk/government/publications/comeap-quantification-of-mortality-and-hospital-admissions-associated-with-ground-level-ozone

COT (2007) 'Variability and Uncertainty in Toxicology of Chemicals in Food, Consumer Products and the Environment' Available at: http://cot.food.gov.uk/cotreports/cotwgreports/cotwgvut

COT (2008) 'Statement on the review of the 1998 COT recommendations on peanut avoidance' Available at: http://cot.food.gov.uk/cotstatements/cotstatementsyrs/cotstatements2008/cot200807peanut

COT (2014) 'Statement on organophosphate' Available at: http://cot.food.gov.uk/cotstatements/cotstatementsyrs/cotstatements2014/cotstatorg

COT (2016a) 'Statement on potential risks from arsenic in the diet of infants aged 0 to 12 months and young children aged 1 to 5 years' Available at: https://cot.food.gov.uk/cotstatements/cotstatementsyrs/cot-statements-2016/statement-on-potential-risks-from-arsenic-in-the-diet-of-infants-aged-0-to-12-months-and-children-aged-1-to-5-years

COT (2016b) 'Statement on the timing of introduction of allergenic foods' Available at: https://cot.food.gov.uk/cotstatements/cotstatementsyrs/cot-statements-2016/statement-on-the-timing-of-introduction-of-allergenic-foods-to-the-infant-diet-and-influence-on-the-risk-of-development-of-atopic-outcomes-and-autoimmune-disease

COT (2016c) 'Statement on hydrolysed cows' milk formulae' Available at: https://cot.food.gov.uk/cotstatements/cotstatementsyrs/cot-statement-on-hydrolysed-cows-milk-formulae

DerSimonian, R. and Laird, NM. (1986) 'Meta-analysis in clinical trials' *Controlled Clinical Trials* 7 pp 177–88. Available at: https://www.ncbi.nlm.nih.gov/pubmed/3802833

Dietary Guidelines Advisory Committee (2015) 'Scientific Report of the 2015 Dietary Guidelines Advisory Committee (Advisory Report)' Available at: https://health.gov/dietaryguidelines/2015-scientific-report/

DuMouchel, W. and Groër, PG. (1989) 'A Bayesian methodology for scaling radiation studies from animals to man' *Health Physics* 57 *(suppl 1)* pp 411-418 Available at: https://www.ncbi.nlm.nih.gov/pubmed/2606701

DuMouchel, WH. and Harris, JE. (1983) 'Bayes methods for combining the results of cancer studies in humans and other species' *Journal of the American Statistics Association* 78 *pp* 293-308 Available at: https://www.jstor.org/stable/2288631?seq=1#page_scan_tab_contents

Duval, S. and Tweedie, R. (2000) 'A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis' *Journal of the American Statistical Association*; 95 pp 89–98. Available at: https://www.jstor.org/stable/2669529?seq=1#page_scan_tab_contents

Egger, M.; Davey, SG.; Schneider, M. *and Minder, C.* (1997) 'Bias in meta-analysis detected by a simple, graphical test' The *BMJ* 315 pp 629–634. Available at: https://www.ncbi.nlm.nih.gov/pubmed/9310563

EFSA (2009) 'Scientific opinion on arsenic in food' *EFSA Journal 7(10):1351*. Available at: https://www.efsa.europa.eu/en/efsajournal/pub/1351

EFSA (2015) 'Principles and process for dealing with data and evidence in scientific assessments' *EFSA Journal* 13(6) pp 4121-4156 Available at: https://www.efsa.europa.eu/en/efsajournal/pub/4121

EFSA (2016) ' Analysis of EFSA methodological needs for evidence use in scientific assessments' EFSA Supporting publication 13(10) 1092E DOI: 10.2903/sp.efsa.2016.EN-1092 Available at: http://onlinelibrary.wiley.com/doi/10.2903/sp.efsa.2016.EN-1092/abstract

EFSA (2017a) 'Scientific Opinion of the PPR Panel on the follow-up of the findings of the External Scientific Report 'Literature review of epidemiological studies linking exposure to pesticides and health effects'. *EFSA Journal* 15(10):5007 [101 pp.]. DOI: 10.2903/j.efsa.2017.5007. Available at: https://www.efsa.europa.eu/en/efsajournal/pub/5007

EFSA (2017b) 'Guidance on the use of the weight of evidence approach in scientific assessments'. *EFSA Journal* 15(8):4971 [69 pp.]. DOI: 10.2903/j.efsa.2017.4971. Available at: https://www.efsa.europa.eu/en/efsajournal/pub/4971

FAO/WHO (2009a) 'Environmental Health Criteria 239: Principles for modelling dose-response for the risk assessment of chemicals' *Geneva: World Health Organization* Available at: http://apps.who.int/iris/bitstream/10665/43940/1/9789241572392_eng.pdf

FAO/WHO (2009b) 'Environmental Health Criteria 240: Principles and methods for the risk assessment of chemicals in food' *Geneva: World Health Organization* Available at: http://www.who.int/foodsafety/publications/chemical-food/en/

FAO/WHO (2011) 'Safety evaluation of certain contaminants in food. Arsenic (addendum)' WHO Food Additives series: 63: 153-316. Available at: http://apps.who.int/food-additives-contaminants-jecfa-database/chemical.aspx?chemID=1863

Government Office for Science (2011) 'Code of Practice for Scientific Advisory Committees' London, UK Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/278498/11-1382-code-of-practice-scientific-advisory-committees.pdf

GRADE (updated October 2013) GRADE Handbook Schünemann, H.; Brożek, J.; Guyatt, G. and Oxman, A. (eds) Available at: http://gdt.guidelinedevelopment.org/app/handbook/handbook.html

Hill, AB. (1965) 'The Environment and Disease: Association or Causation?' *Proceedings of the Royal Society of Medicine* 1965 58 pp.295-300 Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1898525/

Hoffmann S, de Vries RBM, Stephens ML, Beck NB, Dirven HAAM, Fowle JR 3rd, Goodman JE, Hartung T, Kimber I, Lalu MM, Thayer K, Whaley P, Wikoff D, Tsaioun K. (2017). 'A primer on systematic reviews in toxicology'. *Arch Toxicol*. Jul;91(7):2551-2575. doi: 10.1007/s00204-017-1980-3. Epub 2017 May 13. Available at: https://www.ncbi.nlm.nih.gov/pubmed/28501917

Huedo-Medina, T B.; Sánchez-Meca, J.; Marín-Martínez, F. and Botella, J. (2006) 'Assessing heterogeneity in meta-analysis: *Q* statistic or I^2 index?' *Psychological Methods* 11(2) pp 193-206 Available at: https://www.ncbi.nlm.nih.gov/pubmed/16784338

International Epidemiological Association. (2008) 'A Dictionary of Epidemiology. 5th edition. Porta, M. (ed). Oxford University Press.

Jørgensen, AW.; Hilden, J. and Gøtzsche, PC. (2006) ' Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review' The BMJ 333 doi: https://doi.org/10.1136/bmj.38973.444699.0B

Lachat, C.; Hawwash, D.; Ocké, MC.; Berg, C.; Forsum, E.; Hörnell, A.; Larsson, C.; Sonestedt, E.; Wirfält, E.; Åkesson, A.; Kolsteren, P.; Byrnes, G.; De Keyzer, W.; Van Camp, J.; Cade, JE.; Slimani, N.; Cevallos, M.; Egger, M. and Huybrechts, I. (2016) 'Strengthening the Reporting of Observational Studies in Epidemiology—Nutritional Epidemiology (STROBE-nut): An Extension of the STROBE

Statement' *PLoS Medicine* 13(6) e1002036. doi:10.1371/journal.pmed.1002036 Available at: http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002036

Lewis, S. and Clarke, M. (2001) 'Forest plots: trying to see the wood and the trees' *The BMJ* 322 (7300) pp 1479–1480 Available at: https://www.ncbi.nlm.nih.gov/pubmed/11408310

Light, RJ. and Pillemer, DB. (1984) 'Summing up. The science of reviewing research' Cambridge, MA: Harvard University Press

Mills, IC.; Atkinson, RW.; Kang, S.; Walton, H. and Anderson, HR. (2015) 'Quantitative systematic review of the associations between short-term exposure to nitrogen dioxide and mortality and hospital admissions' *BMJ Open* 5(5):e006946. Epub 2015/05/15. Available at: http://bmjopen.bmj.com/content/5/5/e006946

Mills, IC.; Atkinson, RW.; Anderson, HR.; Maynard, RL and Strachan, DP. (2016) 'Distinguishing the associations between daily mortality and hospital admissions and nitrogen dioxide from those of particulate matter: a systematic review and meta-analysis' *BMJ Open* 6(7):e010751. Available at: https://www.ncbi.nlm.nih.gov/pubmed/27443553

Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, DG. and The PRISMA Group. (2009) 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement' *PLoS Medicine* 6(7) e1000097. doi:10.1371/journal.pmed.1000097 Available at: http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000097

Moreno, V.; Martin, ML.; Bosch, FX.; de Sanjosé, S.; Torres, F. and Muñoz, N. (1996) 'Combined analysis of matched and unmatched case-control studies: comparison of risk estimates for different studies' *American Journal of Epidemiology* 143(3), pp 293-300 Available at: https://academic.oup.com/aje/article/143/3/293/78083/Combined-Analysis-of-Matched-and-Unmatched-Case

Morgan, RL.; Thayer, KA.; Bero, L.; Bruce, N.; Falck-Ytter, Y.; Ghersi, D.; Guyatt, G.; Hoojimans, C.; Langendam, M.; Mandrioli, D.; Mustafa, RA.; Rehfuess, EA.; Rooney, AA.; Shea, B.; Silbergeld, EK.; Sutton, P.; Wolfe, MS.; Woodruff, TJ.; Verbeek, JH.; Holloway, AC.; Santesso, N. and Schünemann, HJ. (2016) 'GRADE: Assessing the quality of evidence in environmental and occupational health' *Environment International* 92-93 pp.611-616 Available at: http://www.sciencedirect.com/science/article/pii/S0160412016300034

NTP (2015) 'Handbook for preparing report on carcinogens monographs'. Available at: https://ntp.niehs.nih.gov/ntp/roc/handbook/roc_handbook_508.pdf

NTP (2017) 'About the report on carcinogens'. Available at: https://ntp.niehs.nih.gov/pubhealth/roc/

Pearce, N.; Blair, A.; Vineis, P. *et al*. (2015) 'IARC Monographs: 40 Years of Evaluating Carcinogenic Hazards to Humans' *Environmental Health Perspectives* 123(6) pp.507-514 Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4455595/

Peters, JL.; Rushton, L.; Sutton, AJ.; Jones, DR.; Abrams, KR. and Mugglestone, MA. (2005) 'Bayesian methods for the cross design synthesis of epidemiological and toxicological evidence' *Journal of the*

Royal Statistical Society: Series C (Applied statistics) 54(1) pp 159-172 Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9876.2005.00476.x/abstract

Peters, JL.; Sutton, AJ.; Jones, DR.; Abrams, KR. and Rushton, L**.** (2008) 'The contribution of systematic review and meta-analysis methods to human health risk assessment: Neurobehavioural effects of manganese' *Human and Ecological Risk Assessment: An International Journal* 14(6) pp 1250-1272 Available at: http://www.tandfonline.com/doi/abs/10.1080/10807030802494592?scroll=top&needAccess=true&journalCode=bher20

PRISMA Group: Moher, D.; Liberati, A.; Tetzlaff, J. and Altman DG. (2009) PRISMA 2009 Checklist Available at: http://www.prisma-statement.org/documents/PRISMA%202009%20checklist.pdf from 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement' *PLoS Medicine* 6(7) Available at: http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000097

Prevost, TC.; Abrams, KR. and Jones, DR. (2000) 'Hierarchical models in generalized synthesis of evidence: An example based on studies of breast cancer screening' *Statistics in Medicine* 19 pp 3359-3376 Available at: https://www.ncbi.nlm.nih.gov/pubmed/11122501

Roberts, I.; Kwan, I.; Evans, P. and Haig, S. (2002) 'Does animal experimentation inform human healthcare? Observations from a systematic review of international animal experiments on fluid resuscitation' The *BMJ* 324 pp 474-476 Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1122396/

Rooney, AA.; Boyles, AL.; Wolfe, MS.; Bucher, JR. and Thayer, KA. (2014) 'Systematic review and evidence integration for literature-based environmental health science assessments' *Environmental Health Perspectives* 122(7) pp.711-718 Available at: http://ehp.niehs.nih.gov/1307972/

Rowe, G. (2010) 'Assessment of the COT uncertainty framework from a social science perspective: A theoretical evaluation (A report for the FSA)' Available at: https://www.food.gov.uk/science/research/ssres/crosscutss/evaluncertframework

Sandercock, P. and Roberts, I. (2002) 'Systematic reviews of animal experiments' *The Lancet* 360 pp 586 Available at: http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(02)09812-4/abstract

Sanderson, S.; Tatt, ID. and Higgins, JPT. (2007) 'Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography' *International Epidemiology Association* 36 pp.666-676 Available at: http://ije.oxfordjournals.org/content/36/3/666.full.pdf+html

Schünemann, H.; Hill, S.; Guyatt, G.; Akl, EA. and Ahmed, F. (2011) 'The GRADE approach and Bradford Hill's criteria for causation' *Journal of Epidemiology & Community Health* 65(5) pp.392-395 Available at: http://www.ncbi.nlm.nih.gov/pubmed/20947872

SIGN (2015) 'SIGN 50: A guideline developer's handbook' Edinburgh: SIGN Available at: http://www.sign.ac.uk/pdf/sign50.pdf

Stroup, DF.; Berlin, JA.; Morton, SC.; Olkin, I.; Williamson, GD.; Rennie, D.; Moher, D.; Becker, BJ.; Sipe, TA. and Thacker, SB. (2000) 'Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group' *Journal of the American Medical Association* 283(15) pp.2008-2012 Available at: http://www.ncbi.nlm.nih.gov/pubmed/10789670

Sutton, AJ. and Abrams, KR. (2001) 'Bayesian methods in meta-analysis and evidence synthesis' *Statistical Methods in Medical Research* 10 pp, 277-303 Available at: https://www.ncbi.nlm.nih.gov/pubmed/11491414

Teicholz, N. (2015) 'The scientific report guiding the US dietary guidelines: is it scientific?' *BMJ* 351 Available at: http://www.bmj.com/content/351/bmj.h4962

Tomatis, L. (2002). 'The IARC monographs program: Changing attitudes towards public health' *International Journal of Occupational and Environmental Health* 8 pp 144-152 Available at: https://www.ncbi.nlm.nih.gov/pubmed/12019681

Vandenberg, LN.; Agerstrand, M.; Beronius, A.; Beausoleil, C.; Bergman, Å.; Bero, LA.; Bornehag, C-G.; Boyer, CS.; Cooper, GS.; Cotgreave, I.; Gee, D.; Grandjean, P.; Guyton, KZ.; Hass, U.; Heindel, JJ.; Jobling, S.; Kidd, KA.; Kortenkamp, A.; Macleod, MR.; Martin, OV.; Norinder, U.; Scheringer, M.; Thayer, KA.; Toppari, J.; Whaley, P.; Woodruff, TJ. and Rudén C. (2016) 'A proposed framework for the systematic review and integrated assessment (SYRINA) of endocrine disrupting chemicals' *Environmental Health.* 15 pp 74 Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4944316/

Vandenbroucke, JP.; von Elm, E.; Altman, DG.; Gøtzsche, PC.; Mulrow, CD.; Pocock, SJ.; Poole, C.; Schlesselman, JJ. and Egger, M. for the STROBE initiative. (2007) ' Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration' *Epidemiology* 18(6) pp 805-835 Available at: http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0040297

Walton, HA.; Kang, S.; Atkinson, RW.; Mills, IC. and Anderson, HR. 'Quantitative systematic review of the associations between short-term exposure to ambient ozone and mortality and hospital admissions. In Atkinson RW, Mills IC, Walton HA, Kang S, Anderson HR (2014b), editors. 'Systematic Review and Quantitative Meta-analysis of the Evidence for Associations between Chronic and Short-term Exposure to Outdoor Air Pollutants and Health' *Department of Health Policy Research Programme Project: 002/0037* http://www.sgul.ac.uk/images/Research/Pop_Health/DHreport_chronic_shortterm_exposure_to_outdoor_pollutants092015.pdf

WHO (2013). 'Health risks of air pollution in Europe – HRAPIE project. Recommendations for concentration–response functions for cost–benefit analysis of particulate matter, ozone and nitrogen dioxide' *Copenhagen: World Health Organization* Available at: http://www.euro.who.int/en/health-topics/environment-and-health/air-quality/publications/2013/health-risks-of-air-pollution-in-europe-hrapie-project.-recommendations-for-concentrationresponse-functions-for-costbenefit-analysis-of-particulate-matter,-ozone-and-nitrogen-dioxide

WHO (2014a) 'WHO handbook for guideline development' – 2nd ed. WHO, Geneva. Available at: http://apps.who.int/iris/bitstream/10665/145714/1/9789241548960_eng.pdf?ua=1&ua=1

WHO (2014b) 'WHO Indoor Air Quality Guidelines: household fuel combustion. Methods used for evidence assessment' WHO, Geneva. Available at: http://www.who.int/indoorair/guidelines/hhfc/Evidence_review_methods.pdf?ua=1

Woodruff, TJ. and Sutton, P. (2014) 'The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes' *Environmental Health Perspectives 122 pp.1007-1014 Available at:* https://ehp.niehs.nih.gov/wp-content/uploads/122/10/ehp.1307175.alt.pdf

Glossary

Bias	This is a specific term in epidemiology relating to problems in the study design that may affect the observed measure of association in the statistical analysis. Bias cannot be removed by including larger numbers and it cannot be adjusted for in the statistical analysis. The two main types of bias in epidemiological studies are selection bias and information bias (i.e. measurement error). For example, a study relying on occupational health records to investigate a specific exposure, will not have information on those who developed disease after they left their job (selection bias).
Case-control studies	Case control studies compare individuals with a specific disease or outcome of interest (cases) to individuals from the same population that don't have that disease or outcome (controls). Studies aim to find associations between the disease or outcome and prior exposure to a particular risk factor, but are prone to various biases. (Cochrane glossary, 2018)
Cohort studies	A cohort study is an observational study in which a defined group of individuals (the cohort) is followed over time. The outcomes of individuals in subgroups of the cohort are compared, to examine individuals who were exposed or not exposed (or exposed at different levels) to a particular intervention or other factor of interest. These can be prospective or retrospective in nature. (Cochrane glossary, 2018)
Confounder	A confounder is a factor that is independently associated with both an intervention (or exposure) and the outcome of interest. Failure to account for this will distort the observed measure of association in the statistical analysis. For example, if people in the experimental group of a controlled trial (or the exposed group) are younger than those in the control group, it will be difficult to decide whether a lower risk of death in one group is due to the intervention (or exposure) or the difference in ages. (Cochrane glossary, 2018).
Cross-sectional studies	For example, a survey. Information on outcome and exposures is taken at the same point in time. These are relatively easy to conduct, but it is more difficult to ascribe causality than in a cohort study.
Descriptive studies	A descriptive study describes the characteristics or health status of a sample of individuals. In this type of study the investigators do not actively intervene to test a hypothesis, but just describe the health status or characteristics of a sample from a defined population. (Cochrane glossary, 2018).
Epidemiology	The study of the health status of populations and communities, not just particular individuals. (Cochrane glossary, 2018)

Evidence synthesis	Evidence synthesis involves the development of techniques to combine multiple sources of quantitative evidence. Synthesis techniques such as systematic reviews and meta-analysis, are increasingly being adapted and applied.
Experimental study	In this type of study, the investigators actively intervene to test a hypothesis. In a controlled trial, one type of experimental study, the subjects receiving the treatment being tested are said to be in the experimental group (or arm) of the trial. (Cochrane glossary, 2018).
Intervention studies	This type of study involves an intervention of people, groups, entities or objects in an experimental study. An intervention is sometimes used to describe the regimens in all comparison groups, including placebo and no-treatment arms in a controlled trial. (Cochrane glossary, 2018).
Meta-analysis	A meta-analysis is the use of statistical techniques in a systematic review to integrate and quantify the results of included studies. This term is sometimes misused as a synonym for systematic reviews, where the review includes a meta-analysis. (Cochrane glossary, 2018).
Natural experiments	These are naturally occurring circumstances in which subsets of the population are exposed to different levels of a supposed causal factor, in a situation resembling an actual experiment where human subjects would be randomly allocated to groups, for example, accidental contamination of food or water with a substance. (International Epidemiological Association., 2008)
Observational studies	A non-experimental study - the investigators do not seek to intervene, and simply observe the course of events. Most epidemiological studies are observational. Changes or differences in one characteristic (e.g. whether or not people received the intervention of interest) are studied in relation to changes or differences in other characteristic(s), without action by the investigator. There is a greater risk of selection bias than in experimental studies. (Cochrane glossary, 2018).
Point of departure	In toxicology, the point of departure (POD) is defined as the point on a toxicological dose-response curve established from experimental data or observational data generally corresponding to an estimated low effect level or no effect level. The POD can then be used to calculate a toxicological reference dose (RfD). Points of departure include the BMD, BMDL, LOAEL, and carcinogenic potency estimates, such as the T25. (FAO/WHO, 2009a).
Randomised controlled trials	These are experiments in which two or more interventions, possibly including a control or no intervention, are compared through random allocation to study participants. Most trials,

	assign one intervention to each individual but sometimes assignment is to defined groups of individuals (for example, in a household). Interventions may also be assigned within an individual (for example, in different orders or to different body parts). (Cochrane glossary, 2018).
Sensitivity analyses	An analysis used to determine how sensitive the results of a study or systematic review are to changes in parameters e.g. excluding earlier years, excluding studies with low quality scores from a meta-analysis, only including cohort studies.
Systematic review	"A review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyse and summarise the results of the included studies." (Cochrane glossary, 2018). A systematic review that does not include meta-analysis is sometimes referred to as a narrative review.
Uncertainty intervals/estimates	The term uncertainty intervals is used to refer to confidence intervals. This is the measure of uncertainty around a statistical analysis result. There will be an upper and lower confidence limit. Most estimates use a 95% confidence interval which means that if a study were continually repeated the true value would be contained in 95% of the confidence intervals from those studies. (Cochrane glossary, 2018).
Uncertainty factors	Uncertainty factors or safety factors are used in toxicology to extrapolate from experimental animal or human data to the average human situation. Uncertainty factors may be used for a number of reasons including: to account for inter- and intra-species differences; differences in duration of exposure; issues related to dose-response and the quality of the whole database. (FAO/WHO, 2009b)

SEES membership list

Anna Hansell (Chair)

Professor Anna Hansell is a Fellow of the UK Faculty of Public Health and a Member of the Royal College of Physicians. From July 2018, she took up a position as Professor of Environmental Epidemiology and Director of the Centre for Environmental Health and Sustainability, University of Leicester, also holding an Honorary Consultant position with the University of Leicester Hospitals Trust. Prior to moving to Leicester, she was a Clinical Reader in Environmental Epidemiology and Assistant Director of the Small Area Health Statistics Unit (SAHSU) at the MRC-PHE Centre for Environment and Health, based at Imperial College London and Honorary Consultant with Imperial College Healthcare NHS Trust. She is an environmental epidemiologist and public health doctor, with research interests in the health effects of environmental exposures and areas of study include long term effects of air pollution exposures and the health effects of environmental noise. She has extensive experience of working with large health datasets such as mortality and hospital admissions. Professor Hansell is a past president of the Section of Epidemiology & Public Health at the Royal Society of Medicine and an expert member of the International Volcanic Health Hazard Network.
Professor Hansell was an appointed member of the COT from September 2008 to March 2016.

Declaration of interests available at: https://cot.food.gov.uk/membership/members/ahansell

Derek Bodey

Mr Bodey was appointed to the Committee on 1 April 2010 as the Public Interest Representative. Previously a Principal of a Sixth Form College, Mr Bodey has been a Board Member of the Association of Colleges and is currently a Council Member of the Association of School and College Leaders. His career within education has given him wide experience of committee work, risk assessment and decision making processes. Mr Bodey has for many years had an interest in and commitment to issues of Fairtrade and Trade Justice. Among his charitable interests, he is treasurer of his local Christian Aid committee.

Mr Bodey was an appointed Member of the COT from April 2010 to March 2016. He is an appointed Member of COC from 2013.

Declaration of interests available at: https://cot.food.gov.uk/membership/members/dbodey

Alan Boobis

Professor Alan Boobis is currently Professor of Biochemical Pharmacology and Director of Toxicology Unit (funded by PHE & DH) in the Faculty of Medicine at Imperial College London. He has been a member of Imperial College London (initially at the Royal Postgraduate Medical School, which merged with the College in 1997) for almost 40 years. His main research interests lie in mechanistic toxicology, drug metabolism, toxicity pathway analysis and in the application of knowledge in these areas to risk assessment. He has published around 230 original research papers (H-factor 63) and for several years served as an Editor-in-Chief of Food and Chemical Toxicology.

He is a member of a number of national and international advisory committees, including the Committee on the Medical Effects of Air Pollutants, the WHO Study Group on Tobacco Product Regulation (TobReg) and the WHO Chemical Risk Assessment Network Coordinating Group. He is co-chair of the WHO Mode of Action Steering group, JECFA (veterinary residues – co-chair) and JMPR (alternating co-chair). He has been a member and deputy chair of the UK Advisory Committee on Pesticides, a member and deputy chair of the UK Committee on Toxicity (2003-2012), a member of the UK Committee on Carcinogenicity, the EFSA Panel on Contaminants in the Food Chain and a member and deputy chair of the EFSA Panel on Plant Protection Products. He has also served as a member of the HPA Board Sub-Committee for Radiation, Chemical and Environmental Hazards and the Veterinary Residues Committee.

He is chair of the Board of Trustees of the International Life Sciences Institute (ILSI) a member and a past chair of the Board of Trustees of ILSI HESI and vice-president of the Board of Directors of ILSI Europe. He is involved in several HESI, ILSI Research Foundation and ILSI Europe projects. He is a fellow of the Society of Biology and of the British Toxicology Society. He has served as president and is an honorary member of Eurotox and received the Merit Award in 2009. He is a past chair of the British Toxicology Society and received the John Barnes Prize Lectureship in 2013. He was recipient of the Royal Society of Chemistry Toxicology Award in 2013. He received an OBE in 2003 for his work on the risk assessment of pesticides.

Professor Boobis was appointed chair of the Committee on Toxicity with effect from 1 April, 2015 for 3 years, and reappointed for a further 3 years from April 2018.

Declaration of interests available at: https://cot.food.gov.uk/cot-membership/cot-chair

Janet Cade

Professor Janet Cade was appointed to the Committee on 1 September 2010. She leads the Nutritional Epidemiology Group in the School of Food Science and Nutrition at the University of Leeds. She is a nutritional epidemiologist with particular interests in dietary assessment methodology. She runs the large UK Women's Cohort Study which is characterising dietary exposures in relation to chronic disease outcomes. Other recent work has explored approaches to improve the quality of diets in children; the impact of foods, nutrients and dietary patterns associated with adverse health outcomes including obesity and cancer risk; development and validation of a mobile phone application to support weight loss. Professor Cade chairs the registration committee of the Association for Nutrition, which is involved in the professionalization of nutritionists.

Professor Cade was appointed as a member of the COT from September 2010.

Declaration of interests available at: https://cot.food.gov.uk/membership/members/janetcade

David Lovell

Dr Lovell is a Reader in Medical Statistics in the Division of Biomedical Sciences at St. Georges, University of London. He was previously Associate Director and Head of Biostatistics Support to

Clinical Pharmacogenomics at Pfizer Global Research and Development in Kent, where he provided data management and statistical support to pharmacogenetics and genomics.

David has conducted and managed research programmes on genetics, statistics and quantitative risk assessment. Dr Lovell has been a member of COM since 2006 and the Chair of COM since 2012. He was a Member of COC from 2009 until 2012 and is now an ex officio member of COC. He has been a member of the Scientific Committee of EFSA and a member of the Independent Scientific Advisory Committee, an expert committee of the Medicines and Healthcare Products Regulatory Agency.

Declaration of interests available at: https://gacs.food.gov.uk/sites/default/files/gacs-members-interests.pdf

Neil Pearce

Professor Neil Pearce joined the London School of Hygiene and Tropical Medicine (LSHTM) at the beginning of 2011, after working in New Zealand for the last 30 years. He originally trained in biostatistics, before moving over to do a PhD in epidemiological methods. Since the completion of his PhD in epidemiology in 1985 he has been engaged in a wide range of public health research activities. In 1988 he co-founded the Wellington Asthma Research Group (WARG) at the Wellington School of Medicine. In 2000 he established the Massey University Centre for Public Health Research. He is a Fellow of the Royal Society of New Zealand (FRSNZ) and the Academy of Medical Sciences (FMedSCi) and is currently Past-President of the International Epidemiological Association (IEA). He is also currently a Member of the COC.

Professor Neil Pearce currently teaches epidemiology, biostatistics and public health courses at the LSHTM. He also teaches at the annual European Educational Programme in Epidemiology (EEPE) summer course, and on various IEA courses in developing countries.

He has a broad range of research interests with a common theme of applied epidemiological and biostatistical methods, particularly methods of study design and data analysis for non-communicable diseases (NCDs). In terms of substantive research, during 1980-1988 his main research interest was in occupational epidemiology, and during this time he co-authored the leading textbook of occupational epidemiology, published by Oxford University Press in 1989. During the 1990s, at the Wellington Asthma Research group, he conducted a wide range of research projects including the identification of the role of the asthma drug fenoterol in the New Zealand asthma mortality epidemic, studies of the management of asthma in the community, and more recently studies of the causes of the increases in asthma prevalence in New Zealand and worldwide. He co-authored a textbook of asthma epidemiology which was published by Oxford University Press in 1998. During his ten years at the Massey University Centre for Public Health Research, they conducted a wide range of public health research including respiratory disease, cancer, diabetes, Maori health, Pacific health and occupational and environmental health research. His current research interests focus on epidemiological and biostatistical methods, and their application to studies of neurological disease, occupational and environmental health, asthma, cancer, and health inequalities.

Declaration of interests available at:
https://cot.food.gov.uk/sites/default/files/annualreport2015.pdf

Julian Peto

Professor Julian Peto is the Cancer Research Campaign Professor of Epidemiology and Chairman of the section of Epidemiology at the Institute of Cancer Research, and is Professor of Epidemiology at the LSHTM. He has made major contributions in cancer epidemiology, genetics and clinical trials. He is a leading expert on occupational carcinogens, especially asbestos. His early work showed that industrial asbestos control levels were dangerously high and he developed dose response models that led to more stringent regulations worldwide. His studies of the international epidemic of mesothelioma have led to a Europe-wide ban on asbestos use. Professor Peto showed in 1980 that large genetic effects must underlie the moderate risks observed in relatives of cancer patients and this led to the systematic studies that have identified a series of important genes in breast cancer. This work led ultimately to the cloning of the breast cancer 2 gene. Professor Peto has also been involved in the conduct and analysis of many cancer treatment trials and he has made substantial contributions to statistical methodology. He was appointed as a Member of the COC from 2012 until 2017.

Declaration of interests available at:
https://cot.food.gov.uk/sites/default/files/annualreport2015.pdf

Lesley Rushton

Dr Lesley Rushton is an epidemiologist/statistician with extensive research experience into occupational and environmental causes of ill health. She has worked in several UK academic institutions and is currently a Reader in Occupational Epidemiology at Imperial College London.

Dr Rushton has specialised in health studies in various industries, including leukaemia and related diseases from benzene in the petrochemical industry, studies of lung cancer and silicosis in the silica sand industry, and dermatitis in the printing industry. She has also carried out several studies in the area of indoor and outdoor air pollution, particularly in relation to children's health. She led the study to estimate the current and future burden of cancer due to occupation in Britain. A major new project involves the design and application of an occupational module for UK Biobank. Dr Rushton's methodological research includes systematic review and meta-analysis in the areas of risk assessment and cross-design synthesis. She has been a member of several UK government committees including the COT, the Industrial Injuries Advisory Council, and the EU Scientific Committee on Emerging and Newly Identified Health Risks. She is currently a member of COC and the HSE Scientific and Engineering Assurance Committee.

Declaration of interests available at:
https://cot.food.gov.uk/sites/default/files/annualreport2015.pdf

Heather Walton

Dr Heather Walton is Senior Lecturer in Environmental Health at King's College London. She has a strong interest in the application of science to the optimisation of policy on air pollution and health. A biochemist by training, she subsequently trained in toxicology and is a member of the UK Register of Toxicologists.

She worked as a scientific civil servant for many years at the Department of Health and then the Health Protection Agency, writing review papers and reports for Expert Committees such as the Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment (COT) and the Committee on the Medical Effects of Air Pollutants (COMEAP); developing calls for research proposals for the Department of Health Air Pollution and Health Research Programme and providing scientific advice on air pollution and health to Defra.

The latter included quantification of the benefits to health of reductions in air pollution for the Interdepartmental Group on Costs and Benefits' economic analysis of the air quality strategy. Her interest in epidemiology has developed as a result of these various areas of work. International work has included acting as a temporary special adviser to World Health Organisation meetings on air pollution and health, several presentations in Brussels and presentations to members of the US Health Effects Institute.

Particular interests at King's include the health effects of ozone and whether there is a threshold, distinguishing the effects of nitrogen dioxide from those of particles, health impact assessment, systematic reviews of the evidence on air pollution and health and the use of lifetables to assess the effects of long-term exposure to particles. I am a member of the Committee on the Medical Effects of Air Pollutants, its' quantification sub-group and of a drafting group for a forthcoming statement on recommending methods for quantifying the health impacts of ozone.

Declaration of interests available at:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/376590/COMEAP_Register_of_Members_Interests_2014.pdf

COT secretariat

Diane Benford (Until January 2017)

Dr Diane Benford is a toxicologist with particular expertise in risk assessment. Until May 2017, she was head of the Risk Assessment Unit at the UK Food Standards Agency. The Unit had overall responsibility for advice associated with all types of chemicals in food and of microbial contamination, but much of Diane's work focussed on chemical contaminants, food additives and natural toxicants. She was also scientific secretary to the Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment (COT) and part of the joint secretariat to its sister committees on Mutagenicity (COM) and Carcinogenicity (COC). In a personal capacity Diane was a member of the Scientific Panel on Contaminants in the Food Chain (CONTAM) of the European Food Safety Authority (EFSA) from 2005 to 2015, acting as chair of the panel for the final 3 year term of office. Since 2015 she has been a member of the EFSA Scientific Committee. Also in a personal capacity she has participated in meetings of the Joint FAO/WHO Committee on Food Additives (JECFA) since 2001, firstly as a WHO Temporary Advisor and since 2013 as a member and on two occasions as vice-chair.

Lily Williams (October 2015 to October 2016)

Lily Williams (née Buckley) was a Toxicologist at the Food Standards Agency from 2014 to 2016, carrying out risk assessments of chemicals in food. Alongside this role she was a member of the COT secretariat and wrote papers for the Committee to discuss. She left the Food Standards Agency in 2016 to take up a role as a Technical Expert in Consumer Safety at Syngenta, Jealott's Hill, Berkshire. Prior to working at the Food Standards Agency, she completed a Bachelor's degree in Pharmacology, and a Master's degree in Toxicology.

Claire Potter

Claire Potter is a Senior Toxicologist at the Food Standards Agency carrying out risk assessments of chemicals in food. Alongside this role she is a member of the COT secretariat and writes papers for the Committee to discuss. Prior to working at the Food Standards Agency she worked at Pfizer, Sandwich as a Cellular Toxicologist. During this time Ms Potter was awarded an MSc in Applied Toxicology. Her Bachelors degree was in Biochemistry.

Claire has been a member of the COT secretariat since November 2011.

COC secretariat

Britta Gadeberg

Britta Gadeberg is Principal Toxicologist at Public Health England (PHE) and was appointed PHE Scientific Secretary to COT and COC in November 2016.

Britta works in the PHE Toxicology Department advising on the risks to human health of chemicals in land, water and from waste processes and has been part of the Secretariat for the COC since 2011. At PHE's predecessor organisation, Britta worked in the Air Pollution Unit and supported the work of COMEAP on quantification of the effects of PM2.5¬. Prior to working for PHE, Britta was a Toxicologist at the Food Standards Agency, advising on the health risks from contaminants in food and was part of the COT Secretariat.

Frances Pollitt (Until April 2016)

Frances Pollitt is a toxicologist with nearly 40 years' experience. She began her career in the pharmaceutical industry before moving to Government, where she provided advice on food additives and contaminants and on environmental chemicals. She has also been scientific secretary of two of the Government's expert advisory committees on chemicals in food, consumer products and the environment and, for many years, managed a research contract for the Small Area Health Statistics Unit (SAHSU).