Guidance for recommending classifications and health-based occupational exposure limits

By
the Dutch Expert Committee on Occupational Safety (DECOS)
the Subcommittee on the Classification of Carcinogenic Substances
the Subcommittee on the Classification of Substances Toxic to Reproduction

The Hague, December 21, 2021

Health Council of the Netherlands

contents

01 introduction

The Health Council of the Netherlands is an independent scientific advisory body whose legal task is to advise the Dutch Government and Parliament in the field of public health, and health and health care research. One of the areas is healthy working conditions. Advisory reports are drawn up by committees, which are composed of scientific experts from various expert fields. More information about the advisory reports, the Health Council and its tasks and procedures, can be found on the website, www.healthcouncil.nl.

1.1 Advice on chemical substances

At the request of the Minister of Social Affairs and Employment, the Health Council evaluates the toxicity of substances to which workers can be exposed in the workplace. The evaluations form the basis for recommending a classification as a carcinogenic, mutagenic or reproduction toxic substance, or a health-based occupational exposure limit (health-based OEL). The Minister then uses the recommendations to decide on the inclusion of the substance in the official List of Carcinogens, Mutagens and substances toxic to Reproduction (CMR; Dutch Working Conditions Act), and to set legally binding occupational exposure limits.

The evaluations and recommendations are performed by the Dutch Expert Committee on Occupational Safety (DECOS) and its two subcommittees: the Subcommittee on the Classification of Carcinogenic Substances, and the Subcommittee on the Classification of Substances Toxic to Reproduction. In this guidance, the DECOS and its subcommittees are indicated as 'the committees'.

1.2 Goal of the guidance

Over the past years, the Health Council has published several guidelines and advisory reports on the evaluation process of chemical substances and the methodologies used, including:

- Guideline for the calculation of occupational cancer risk values (2012);[1]
- Guideline to the classification of carcinogenic compounds (2010);[2]
- Prevention of work-related airway allergies: recommended occupational exposure limits and periodic screening (2008);[3]
- Standard for dermal exposure in the workplace (2001; in Dutch, summary also in English);[4]
- Adjustment of occupational exposure limits in case of unusual work schedules (2001; in Dutch, summary also in English);[5]
- Toxicology-based recommended exposure limits (1996);[6]
- *Gezondheidskundige aspecten van het begrip blootstelling en van het meten/schatten ervan* (1990, only in Dutch).[7]

The present guidance is a supplement to the technical guidelines mentioned above. It outlines the advisory process and the general principles used by the committees to advise. It has been compiled under supervision of the committees.

The present guidance incorporates current methodologies and principles also used by other scientific organisations, such as the Risk Assessment Committee (RAC, European Chemicals Agency [ECHA], EU, 2017 and 2019), the Scientific Committee on Occupational Exposure Limits (SCOEL, EU, 2017), the International Agency for Cancer Research (IARC), and the collective expert appraisal report published by the French National Agency for Food, Environmental and Occupational Health & Safety (Anses, 2014).[8-13] In addition, the committees reviewed and discussed the latest developments and views in the public scientific literature on evaluating scientific papers for hazard and risk assessment purposes.

The committees may deviate from the guidance for each recommendation. In that case, the reason will be explained in the advisory report. Furthermore, the committees will continue to monitor developments in methodologies, and adjust the guidance in the future if necessary.

1.3 Public consultation

In addition to standard Health Council procedures, draft advisory reports that are approved by the committees and the Board of the Health Council are published on the website to give interested parties or persons a one-time opportunity for a written comment on the scientific studies used in the advisory report and on the judgement of these studies by the committees, and to point out possible inconsistencies in the argumentation in the advisory report. The committees consider the received comments for the final decision of the recommendation. Comments related to feasibility, practicality, commercial or political views are not considered, because the mandate of the Health Council is restricted to the state of the art of health science. The duration of this public comment round is three months.

1.4 International activities

The Health Council strives for a uniform European approach in determining health-based recommended exposure limits. Therefore, the DECOS works together, where possible, with Europe's Nordic Expert Group (NEG), the US National Institute for Occupational Safety and Health (NIOSH) and Anses. The DECOS also liaises with the German Ausschuss für Gefahrstoffe (AGS) of the Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, and the Senatskommission zur Prüfung gesundheitsschädlicher Arbeitsstoffe of the Deutsche Forschungsgemeinschaft (DFG). Furthermore, the DECOS comments on draft advisory reports submitted by the RAC and ECHA to third parties under the European Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH[a]) legislation concerning substances for which the Council previously issued advisory reports.

[a] The REACH Regulation aims to improve the protection of human health and the environment from the risks that can be posed by chemicals, while enhancing the competitiveness of the EU chemicals industry.

1.5 Substances

The Health Council publishes its work programme for the following year in September. Substances for which the Council received a request by the Minister of Social Affairs and Employment are listed in the chapter on Working conditions. The request concerns substances of concern that are used in, or produced by, Dutch industry, and which fall outside the private responsibility of the manufacturers or importers according to the Dutch Working Conditions Law and Regulations and the REACH Regulation (legislation regulating the authorisation of chemicals within the European Union).

The substances can be mono-constituent chemicals, mixtures, complex substances, and emissions, and can be grouped. They can occur in a natural state or as a produced, used or intentionally or unintentionally released agent.

- *Mixtures* are produced for certain purposes and contain several substances, whereby each substance retains its own properties. An example is ink.
- *Complex substances* are of natural origin and contain many constituents, such as substances of unknown or variable composition, complex reaction products, or biological materials (UVCBs). Examples are crude oil and wood dust.
- *Emissions*. Workers may be unintentionally exposed to emissions during their work. Examples are welding fumes, diesel engine exhaust, and emissions from iron and steel founding. The composition of the emissions is often complex and variable, depending on the work processes and actions.
- *Groups of substances*. The committees consider grouping of closely related substances with comparable physicochemical properties and structural similarities that are likely to have a similar toxicity profile, and for which no complete toxic profile is available for part of the individual substances. Examples are inorganic lead and lead compounds, benzo(a)pyrene and polycyclic aromatic hydrocarbons, diisocyanates, and wheat and other cereal dusts.

Medicines or plant protection products and biocides are excluded, because evaluations of the safe use of these substances for humans, animals, and the environment are performed by the Medicines Evaluation Board and the Dutch Board for the Authorization of Plant Protection Products and Biocides.

The committees assess the mixture, the complex substance, or the emission as a whole, basing their advice on data and evidence of tested material that can be considered 'typical' in composition. If insufficient data are available, the committees check whether data from the individual constituents can be used. The committees also make a choice about the best exposure parameter. This can be a chemical constituent, for example a substance that is associated with adverse effects of interest and is

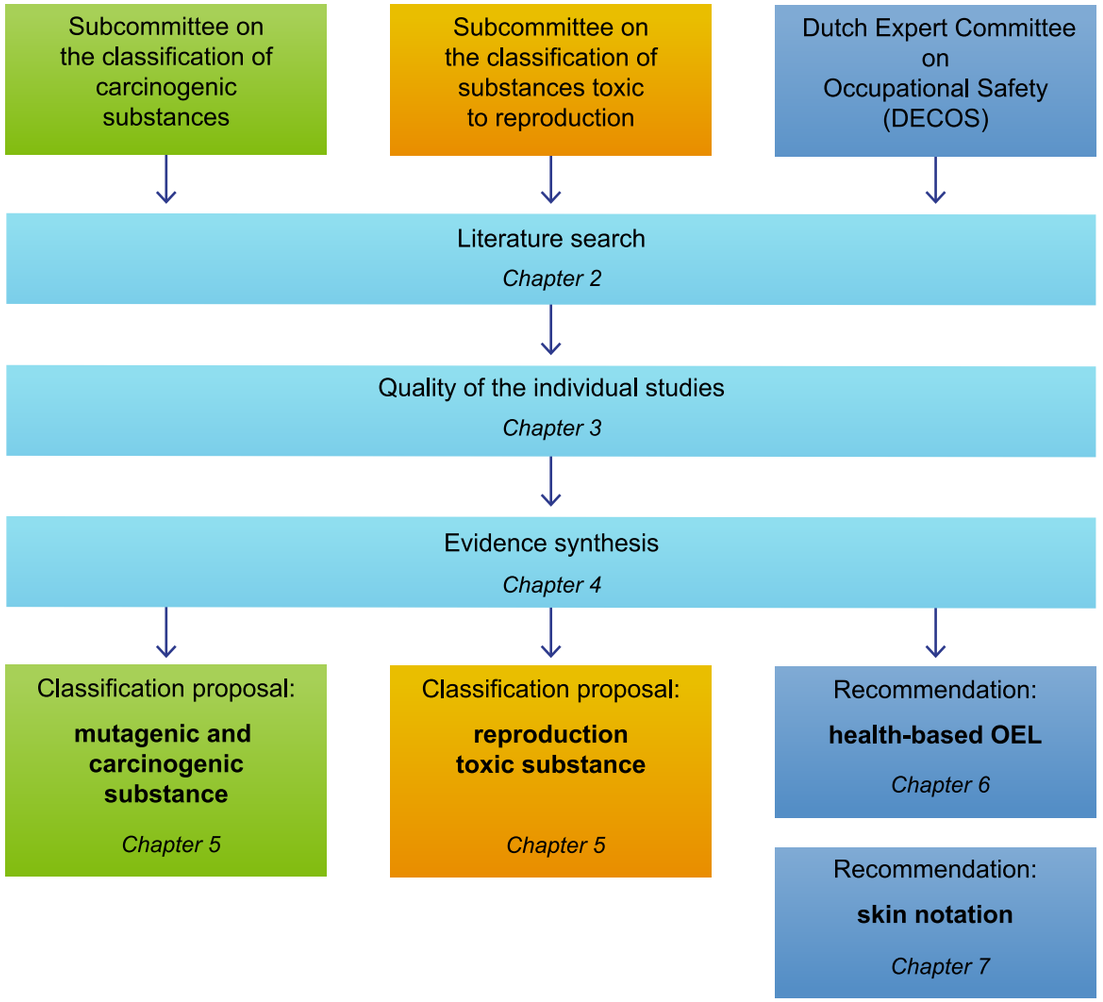
almost always present in the mixture, but also a constituent that serves as a substitute for it and is easy to monitor.

1.6 Steps towards a recommendation

To recommend classifications and health-based OELs, the committees need to identify the potential adverse effects (hazards), and to characterise the hazards by assessing the existence of a relationship between the substance under investigation and the observed adverse effects. In regulatory toxicology, the identification and characterisation of health hazards are part of a hazard assessment that by itself is part of the risk assessment, which is used to manage health hazard risks in the workplace. In risk assessment, the hazard classifications and the OELs can be used to control, and thus to manage, these risks.

To identify and characterise the potential adverse effects, the DECOS and the subcommittees start with a literature search (Chapter 2), followed by assessing the quality of the individual studies (Chapter 3), and an evidence synthesis on the existence of a relationship between exposure and hazard (adverse effect) (Chapter 4). To propose a classification, the subcommittees search for evidence of a relationship regarding the hazard under investigation (mutagen, carcinogen, or reproductive toxicant). It then use criteria on the strength of evidence to propose a classification category (Chapter 5).

Steps to a recommendation



To propose a health-based OEL, the DECOS searches for evidence of a relationship regarding all possible adverse effects (local and systemic effects; acute, short, mid and long term effects) and for studies with data on quantitative relationships or associations between exposure and identified adverse effects in order to characterise the hazard. These quantitative relationships or associations are then used to derive health-based OELs (Chapter 6) and to consider whether a skin notation should be issued if a substance can cause adverse effects upon dermal exposure (Chapter 7).

02 literature search

The recommendations of the committees are established on what is known in publicly available scientific literature; the Health Council itself conducts neither experimental nor observational research. To be able to identify and characterise potential adverse effects, data are needed to assess a toxicological profile of the substance under investigation.

2.1 Search strategy

The search starts with the search for chemical hazard and risk assessment reports that were published by other scientific organisations, such as the:

- monographs by the International Agency for Research on Cancer (IARC);
- opinions by the RAC of the European Chemicals Agency (EU);
- opinions by the Scientific Committee on Occupational Exposure Limits (EU);
- assessment reports of sister organisations (i.e. NEG, DFG, ACG, Anses, NIOSH, ATSDR);
- concise international chemical assessment documents (CICADs) by the WHO;
- risk assessment reports by the European Union.

If such reports are available, the literature search starts at the last date of the search mentioned in the relevant assessment report. In addition, the most relevant references in the reports are retrieved to be judged by the committees.

Original studies and other reviews are retrieved using PubMed and SCOPUS, and through screening the reference lists in the original studies and reviews. The keywords used to extract literature include substance identification terms and terms related to exposure, adverse health effects, toxicology, kinetics, and mode of action. For the classification proposals, the literature search is restricted to the technical terms related to carcinogenicity, genotoxicity, or reproductive toxicity.

2.2 Data extraction

To obtain a complete toxicological profile of a substance under investigation, the following information is extracted from the literature:

- Substance identification and physicochemical properties;
- Occurrence, production and use;
- Stationary and personal airborne exposure and biological exposure monitoring techniques, as well as data on occupational exposure routes, levels and characteristics;
- Substance toxicokinetics; toxic mode of action/adverse outcome pathways or pathogenesis;
- Adverse effects (including acute, short, and long term effects);
- Exposure-response relationships;

- Data on existing classifications and labelling, Occupational Exposure Limits (OELs), and Biological Limit Values (BLVs).

Information on the toxicity of substances is derived from a variety of research methods, including epidemiological studies, laboratory studies with humans, animal (*in vivo*) experiments, cellular and tissue (*in vitro* or *ex vivo*) studies, and computational methodologies and disciplines.

03 quality of the individual studies

After scientific literature has been collected, each study is assessed for quality by judging its relevance, reliability, and validity. This is necessary to prevent less relevant or low-quality studies from contributing disproportionately to the strength of evidence. The three aspects may influence each other to a certain extent. In addition, when judging the quality of a study the committees look specifically at the research design, the assessment of exposure and effect, and the practical execution of the study.

The committees are aware that several methods are developed to evaluate the quality of a study (see Section 3.2 for details). New scientific insights and technological developments ensure that there is an ongoing discussion about how best to judge or set up a study. As part of the scientific community, the committee members closely monitor these developments, and apply the generally accepted views in their recommendations.

3.1 Relevance

Regarding *relevance*, the committees judge whether the individual studies give information on causal relationships or associations between exposure and response in a qualitative (associations, mode of actions) and quantitative way (exposure-response models).

The committees consider studies with data on adverse effects and exposure-response models relevant, because these data are essential to assess causal relationships or associations, and are needed to perform quantitative risk analyses (modelling exposure-response relationships). Data on the physical and chemical properties, toxicokinetics, pathogenesis, and toxic mode of action/adverse outcome pathways are considered relevant to understand how the substance induces adverse effects, and are considered as supportive evidence for a causal relationship. Data on source, use, exposure monitoring techniques, and occupational exposure levels in the workplace are relevant to gain insight into what the substance is used for, how it is applied, and to which exposure levels people are exposed in practice.

3.1.1 Research methods

Research in humans

The committees prefer to address causal relationships or associations in controlled human exposure studies, in which subjects are exposed to exposure conditions similar to those in the workplace. Due to ethical and

practical constraints, however, such studies are limited to safe exposure levels in volunteers to assess toxicokinetics or acute or short-term effects, to establish safe dosage ranges of medicines, or to identify and monitor side effects. Therefore, occupational toxicity research in human subjects mainly depends on observational studies, in particular on observational studies with cohort, case-control, and cross-sectional study designs. Particularly useful for evaluating mid- and long term adverse effects are cohort studies, in which participants are free of the effect (disease) under investigation at the start of the study and are followed over time (in some fashion). In these studies, exposure and effect are often independently measured, with the exposure occurring before the effect is assessed. Case-control studies, in which participants already have the effect of interest (cases) or not (controls) and exposure from the relevant period is reconstructed, are useful for rare effects, because smaller numbers of participants are needed compared to cohort studies. Cross-sectional studies, in which exposure, body burden, and a measure of physiological effect can be measured at one point in time, are useful to assess exposure and body burden using up-to-date methods, but lack a clear timeline from exposure to effect. In their evaluations, the committees consider all observational designs, giving - if available - most value to prospective cohort studies, because these types of studies are closest to the designs of controlled human exposure studies.

Animal experiments

The benefit of using laboratory animals (mammals) in toxicological research is that it offers the possibility to study causal exposure-response relationships and toxicokinetics under controlled exposure conditions during a short or long term period, up to the entire life span of the animal. Despite interspecies differences, the micro-anatomy of the body and many physiological and pathophysiological processes are generally considered comparable between humans and other mammals, but not necessary identical. Animal experiments are a valuable model in hazard and risk assessment, as it allows investigators to fill the gap in knowledge of epidemiological and controlled human exposure studies.

In vitro and ex vivo test systems

To reduce animal experiments, *in vitro* and *ex vivo* test systems have been developed for hazard and risk assessment purposes. In addition, standardised *in vitro* and *ex vivo* skin and eye irritation and sensitisation tests are considered sufficient to judge the irritation and sensitising properties of a substance. Overall, *in vitro* or *ex vivo* studies could be relevant in understanding early or acute biochemical or cellular toxic effects, albeit without the complexity of interaction between cells or tissues, and other biochemical processes in the body. At present, however, no reliable and standardised test systems exist for studying systemic effects.

Computational (in silico) toxicology

As an alternative to animal testing, computational methods have been developed to predict the toxicity of chemicals. Examples are:

- read across, in which information on toxicity of one substance is predicted by using toxicity data of one or more other related substances;
- physiologically based pharmacokinetic (PBPK) modelling, a mathematical framework for predicting toxicokinetics of a substance;
- quantitative structure-activity relationship (QSAR) modelling, a mathematical model to predict physicochemical, biological and environmental fate properties of compounds based on the knowledge of their chemical structure;
- trend analysis methods to predict toxicity of a chemical by analysing toxicity trends of other tested chemicals;
- quantitative *in vitro* to *in vivo* extrapolation (QIVIVE) of toxicity data.[14-16]

For certain methods, the OECD (Organisation for Economic Co-operation and Development) has developed guidelines, such as the OECD guidance on grouping of chemicals, and the OECD QSAR validation guidelines. As with *in vitro* and *ex vivo* test systems, the relevance for the committees lies in predicting the toxic properties and toxicokinetics, and in deciding whether chemicals can be grouped.

The availability of big data and the progress of deep learning stimulates advanced computational modelling approaches for chemical risk assessment.[17] For example, these efforts make it possible to develop adverse outcome pathway (AOP) frameworks, which describe a sequence of biological events that lead to an adverse effect by mechanistic reasoning. The committees expect that the development and use of such advanced computational methods will expand in the near future, as increasing efforts are made to reduce or replace animal experiments. They will closely follow these developments for possible future use in their evaluations.

3.1.2 Exposure assessment

To derive health-based OELs, the DECOS needs data on intensity (exposure levels), time (duration, frequency, etc.) and timing of exposure. More specifically, it searches for data from actual or historical (individual) exposure measurements performed during work activities, or for data reflecting occupational exposure conditions (intensity, time and route of exposure) as much as possible. It should be clear that the exposure preceded the induction of an adverse effect, taking into account the latency period of the health effect being studied. Moreover, the authors of human studies should address the use of protective measures to have a more accurate insight in what the actual inhaled exposure levels were. Studies providing these data are considered of high relevance.

The subcommittees on the classification of substances also need exposure data to be sure that the exposure preceded the induction of the health effect, and that exposure occurred in the relevant time window (effects on reproduction). However, data on intensity and time of exposure is less important, because classification is based on the toxic properties of a substance (hazard principle) and not on the intensity of exposure. Exposure levels are only considered to rule out the possibility that general toxicity may have influenced the adverse effect under investigation.

3.1.3 Adverse effects

The goal of classifying substances and setting OELs is to prevent diseases (any condition that impairs the normal homeostatic functioning of the body), disorders (structural or functional abnormality or disturbance) or other adverse effects in workers who are exposed to hazardous chemicals in the workplace, and in their progeny. These effects can arise after single or repeated exposure, almost immediately upon exposure, or after a short or long latency period. In addition, the committees consider any adverse effect in humans of relevance, irrespective of its reversibility, severity or specificity, or whether continuing exposure results in an adaptive response.

When considering effects, the committees distinguish adverse and non-adverse effects. Within the field of toxicology and risk assessment, however, the difference between adverse and non-adverse is still under discussion.[18-23] At the moment, the committees use the definition used by the ECETOC (2002): 'non-adverse effects can be defined as those biological effects that do not cause biochemical, behavioural, morphological or physiological changes that affect the general well-being, growth, development or life span of a human or an animal'.[21] A significant biological effect is 'a response to a stimulus in an organism or other biological system that is considered to have substantial or noteworthy effect (positive or negative) on the well-being of the biological system. The concept is to be distinguished from statistically significant effects or changes, which may or may not be meaningful to the general state of health of the system'. In addition, the committees distinguish normal and abnormal biological effects or responses in their evaluations, the first often representing a normal homeostatic reaction to a stimulus. Normal non-specific biological reactions can serve as an indication for causality, although data on these reactions are generally not used to derive health-based OELs.[24]

3.2 Reliability and validity

The judgement on *reliability* concerns the correct and consistent use of methods (precision), so that similar results can be produced under the same conditions over time, by the same and other research groups (reproducibility). Reliability also depends on how complete the reporting is.

Validity refers to whether results really represent what they are supposed to measure (accuracy). The committees make a distinction between external and internal validity. Internal validity is the extent to which a method or study design provides results as close to the truth as possible, while alternative explanations are ruled out. External validity is the extent to which the results of a study can be generalised to other situations, populations, or organisms.

3.2.1 Epidemiological studies

Observational studies feature a more limited control of the circumstances under which exposure took place than experiments. Consequently, these studies are more prone to random error or random variation in estimates. In addition, all observational studies contain some degree of systematic error (bias). Random errors concern the lack of precision (reliability), whereas systematic errors concern the lack of accuracy (validity).
The essence of judging the quality of observational studies is to what degree the investigators minimised random and systematic errors in their study through proper design, measurement methods, and conduct of the study.[25-28]

Regarding bias, it can affect the internal and external validity of a study.[27,29] It is the result of limitations in study design or conduct, or of a lack of correction for confounders, that systematically leads to an erroneously stronger, weaker, or even inverted association compared to the real association between a substance and the health effect of interest. Many potential sources of bias may occur in epidemiological studies. The committees pay special attention to the types of bias, which often occur in observational research on occupational exposure: selection bias, information bias, and confounding.

Selection bias

Selection bias may occur in cohort and case-control studies if exposed and non-exposed groups - or the selected cases and controls, respectively - are not truly comparable. In cohort studies, it is caused by differential selection or loss-to-follow-up of study participants according to their exposure status and outcome/disease status. An example is the healthy worker effect, which reflects a healthier status of the workforce compared to the general population, and may result in lower incidence or mortality rates than expected. The healthy worker effect is less likely when adverse health effects with a long latency period are considered, such as cancer.[30] Another example is a higher incidence rate than expected for

subfertility, because reproductively unhealthy workers with unsuccessful reproduction stay in the workforce. In case-control studies, the selection of controls may not be representative of the population from which the cases are derived. This may result in under- or overestimation of effect estimates. Cross-sectional studies are often hampered by self-selection of the study population or by changes in lifestyle habits as a result of the disease, which precludes conclusions about exposure-effect associations.

Information bias

Information bias is the misclassification of study participants with respect to exposure and/or effect, with the consequence that the participants are assigned to the wrong category of exposure, effect, or both. It may occur, for example, when participants are asked to recall their occupational history and related exposures, which might be different depending on whether or not they developed the adverse effect under investigation. This form of bias can play a role in case-control studies and cross-sectional studies in particular, but may also occur in cohort studies, for example when assessment of the adverse effect is not independent of exposure status.

Confounding

Confounding occurs when the effect of the exposure is mixed with an effect of one or more other risk factors (confounders) for the adverse effect of interest. The potential confounder should be associated with or precede the exposure, cause the outcome, and should not be on the causal pathway of exposure to effect. An example is a study on occupational exposure to a carcinogenic substance that causes lung cancer. Another risk factor for lung cancer is smoking (cause of outcome). When smokers are distributed disproportionally between exposed and unexposed groups (association with exposure) and smoking does not interfere with the induction of lung cancer by the substance under investigation (not on the causal pathway), the observed association between substance exposure and lung cancer may partly or completely explained by smoking. The result is over- or underestimation of the risk that the substance causes lung cancer. By correcting for confounding in statistical analyses, however, all or part of the confounding effects can be eliminated, leading to an adjusted effect estimate that is closer to the true association than the uncorrected effect estimate.

Several frameworks have been developed in an attempt to evaluate the reliability and validity of human observational studies systematically in a replicable and transparent way (e.g. IARC preamble 2019; Goldbohm et al. 2006; Money et al. 2013; Swaen 2006; Vlaanderen et al. 2008; WHO Workgroup 2000).[12,26-28,31,32] Some of these have been developed for a specific purpose. For example, Vlaanderen et al. (2008) developed a guideline for the evaluation of observational studies for quantitative risk assessment with the focus on the quality of the exposure assessment.[28] Money et al. (2013) developed a scoring system on quality for regulatory

hazard and risk assessment purposes, which is based on the criterion and rating system developed by Klimisch et al. (2007) for animal toxicity studies (see Section 3.2.2).[26] Several risk of bias tools (RoB tools) have also been developed (e.g. ROBINS-I; Newcastle-Ottawa scale; GRADE [as part of quality of evidence assessment]; Navigation Guide; OHAT; CBO), which differ to some degree in addressing and defining bias domains.[33-36] However, the usefulness of these tools for observational studies is the subject of debate. For instance, most of these tools attribute the highest weight to randomised controlled trials (RCTs), because bias is less common in RCTs. In addition, most RoB tools do not take into account the magnitude and direction of the bias, so that valuable information is lost.[37] To better judge the reliability and validity of studies, guidelines have been published to strengthen the reporting of observational studies in scientific papers (STROBE, Von Elm et al. 2008).[36]

In summary, the frameworks addressing reliability and validity issues related to the following:

- *Study design*
  Was the study design appropriate to the study objective? Were the timeframe and observation period sufficient for the study objective (in cohort and case-control studies)?
- *Study population*
  Was the study population representative and the comparison group appropriate to allow for comparison between exposure groups? What was the completeness of follow-up (in cohort studies)?
- *Exposure assessment*
  Were exposure data available? Was the exposure assessment performed independently of outcome? Were validated exposure monitoring and analysis techniques used? Was insight given in variability of exposure? Were individual or group exposure data obtained? Were occupational hygiene data (e.g. the use of preventive measurements) and exposure patterns available?
- *Effect assessment*
  What was the quality and completeness of health parameters? Was the effect assessment performed independently of exposure status? Were validated procedures used for data collection on effects?
- *Statistics*
  Was the population size (power) sufficient? Were the statistical techniques appropriate, including confounding checks through statistical adjustment and/or sensitivity analyses?
- *Reporting*
  Were methods, data, and results clearly presented in relation to the study objective to allow for evaluation? Were bias, assumptions or other aspects of the study presented ? Was STROBE used?

Sometimes the results of a number of individual epidemiological studies on the same substance and adverse effects (multi-comparable studies)

are aggregated in a meta-analysis. The goals of such analyses are to address ambiguity (random or systematic errors in design, exposure and outcome) among the individual studies, or to study a hypothesis that is difficult to address in a single study.[12,38] A special type of meta-analyses is the IPD meta-analysis, in which individual participant or patient data (IPD) from multiple studies are used, including correction for confounding, instead of aggregated data. The main concerns regarding reliability and validity of all meta-analyses are the objectiveness of the study selection, the heterogeneity among the selected studies, handling of incomplete data, publication bias, and statistical approaches.

Considering the frameworks and guidelines, the IARC preamble is mostly in line with how the committees wish to assess the observational studies for quality.[12] Therefore, the IARC preamble serves as an example, but the committees do not let their judgement depend solely on this framework. They will continue to monitor developments in the scientific literature and adjust their assessment method where necessary.

3.2.2 Controlled human exposure studies

During literature search, the committees sporadically come across controlled human exposure studies (clinical studies) to evaluate adverse effects. If the authors of the study have taken ethical issues into consideration, the committees also evaluate these studies on reliability and validity.

The considerations that the committees take into account to judge reliability and validity of these studies are similar to those for observational studies (see Section 3.2.1) and animal experiments (see Section 3.2.3) and include correctness of design to allow a conclusion on the study objective; control of variables that could influence the occurrence of effects (e.g. selection of volunteers, objectiveness in assigning volunteers in exposure groups; size of study groups, exposure frequency and levels); completeness in presenting design, data, methods, results and bias and other limitations of the study.

3.2.3 Animal experiments

For animal toxicology many internationally accepted study guidelines exist, which are considered reliable and valid. Examples are those prepared by the OECD, the International Organization for Standardization (ISO), and the NIOSH. The committees consider animal studies that were performed according to these guidelines, or were closely related or comparable to these guidelines, of high value.

For regulatory hazard and risk assessment purposes, tools have been developed to evaluate the reliability and internal validity of animal toxicity tests systematically. One such tool was developed by Klimisch et al. (1997).[39] This is a rating system with four scores of reliability, which is widely used by many authorities and organisations. It mainly assesses whether the study was performed in accordance with validated guidelines

(preferably in accordance with Good Laboratory Practice) and which source of information was used (primary or secondary literature, full report or abstracts, detailed description).[40-43] Because of the lack of detail in assigning data quality, the European Centre for the Validation of Alternative Methods developed the Toxicological data Reliability Assessment Tool (ToxRTool).[44] This Excel-based tool provides comprehensive criteria and guidance for reliability evaluations of *in vivo* and *in vitro* studies, and leads to the assignment of Klimisch scores. The committees may use the criteria of these tools in combination with expert judgement.

In summary, the frameworks address reliability and validity issues related to the following:

- *Study design*
  Was the study design appropriate to the study objective? Were reliable and validated study designs used according to internationally accepted standards?
- *Animal species*
  Were the animal species well selected and defined in relation to the study objective? Was the number of animals sufficient to take into account intra-species variability? What controls were used (e.g. sham controls, in which animals are treated the same way as exposed animals, but without the actual use of the test substance; positive controls, in which animals are exposed to a substance that is known to induce the adverse effect under investigation; negative controls, in which animals are not given the substance under investigation; historical controls, if no changes in housing, diet, and microbiota composition have taken place in de animal facility) and were these controls sufficiently defined to allow for comparisons with exposure groups?
- *Effect monitoring*
  Were reliable and validated effect monitoring techniques used? Were data presented on general indicators of toxicity, such as changes in food consumption, body weight, and early mortality, to allow for considering secondary effects? Did the study include extensive pathological analysis?
- *Exposure assessment*
  Were the duration, timing and level of exposure appropriate, and was the observation period long enough to observe an adverse effect in relation to the study objective? Were sufficiently low exposure levels included to allow for no adverse effects? Were exposure range finding studies performed? Were data presented on exposure-response relationships?
- *Statistical analysis methods*
  Were these correctly applied? What data on effect size are available? What is the statistical significance?
- *Reporting*
  Is full insight given into the study design, methods and analyses,

results, possible flaws and deficiencies? Would the study be reproducible? Was the study peer-reviewed?

Mammals (rats, mice, dogs etc.) are used in toxicity testing, because their body composition, metabolism and other biological processes are assumed to be very similar to humans. However, they can differ in anatomy, physiology, vulnerability and biological rhythm. This means that there is always uncertainty in external validity due to species differences. The committees assess whether the observations in animal experiments are relevant for humans. In doing so, they take into account the findings of studies with other animal species (if available) and knowledge about the toxicological action mechanisms and biological differences between the laboratory animals and humans.

Another issue with external validity is that uncertainty can be introduced by large differences in exposure conditions (exposure levels, duration, frequency, route of exposure etc.) between animal testing and the exposure conditions in the workplace. This could be an issue when animal testing is used to derive a health-based OEL for humans. The DECOS addresses this on a case-by-case basis. As the classification of substances is based on the toxic properties, the exposure conditions are less relevant for classification.

3.4.2 *In vitro* and *ex vivo* test systems

Much that has been written about the reliability and internal validity of animal experiments also applies to studies using *in vitro* and *ex vivo* test systems. For *in vitro* and *ex vivo* testing, internationally accepted standards and guidelines (e.g. OECD, ISO, NIOSH) exist as well.
In addition, the ToxRTool is used to assess reliability and internal validity, as well as the OECD guidance document on Good In Vitro Method Practices (GIVIMP) and the template of cell-based toxicological test methods (Krebs et al. 2019).[44-46] The committees may make use of these tools, in combination with expert judgement.

3.3 Expert judgement

The evaluation of study quality is inherently reliant on expert judgement. This may lead to differences in conclusions among experts and among expert committees. To obtain scientifically robust hazard and risk assessments, the Health Council strives to use a systematic and transparent evaluation, so that third parties can find out how the quality of the individual studies was addressed. This means that the advisory reports contain objective summaries of the available studies, accompanied by a comprehensive evaluation of the quality of the study.

04 strength of evidence for causality

Based on the available data and the evaluation on the quality of the individual studies, the DECOS judges whether there is a causal relationship or association between exposure to the substance under investigation and the adverse effects observed.

4.1 Weighing the strength of evidence

Causality is a basic philosophical concept, in which a factor, such as exposure to the substance, activates a mechanism that could result in an adverse effect in an organism.[29] The committees are aware that the induction of an adverse effect is rarely caused by just one factor, in the sense that, for instance, genetic constitution, lifestyle and environmental factors together result in an adverse effect. Therefore, when evaluating causality, they rather weigh the strength of evidence that exposure could has caused an adverse effect.

For each potential adverse effect, the committees search for several lines of evidence, which they then integrate to determine the overall weight of evidence for each effect observed. This is done to evaluate consistencies/inconsistencies and uncertainties in the patterns of effects and strength of evidence across the lines of evidence, which represent different disciplines within evidence research, e.g.: human (observational) studies, animal experiments, mechanistic studies (mode of action) and studies on toxicokinetics. For classifying substances, the subcommittees focus on genotoxic and carcinogenic effects or on adverse effects on reproduction, whereas the DECOS takes all possible adverse effects into account. In addition, the relevance and quality of a study determine the extent to which the data in that study can contribute to weighing the strength of evidence.

4.2 Tools to assess strength of evidence

Some considerations and tools have been presented in the literature to evaluate the strength of evidence for causality in a transparent and consistent way.

4.2.1 Bradford Hill's considerations

In 1965 Bradford Hill published a list of considerations for causality in human studies (Bradford Hill, 1965) that is still applied today.[47] His considerations are:

- consistency (repeated observation of an association in different populations and circumstances);
- strength of association (strong associations are more likely to be causal than weak associations);
- temporality (exposure precedes disease in time, which means that study designs, such as cohort and case-control studies, provide stronger evidence for causality than cross-sectional studies);
- specificity (a cause leads to a single effect, not multiple effects);
- coherence (association does not conflict with what is known of biology or disease);

- biological plausibility;
- biological gradient (presence of a unidirectional dose-response curve);
- experimental evidence;
- analogy.

Bradford Hill's considerations were never intended as a checklist, or to be rigid criteria for causality assessment. Instead, the considerations are fundamental concepts in supporting evidence synthesis, which should be discussed in the light of the current knowledge. If not all considerations are met, the substance can still be a cause, as the evidence may not yet be complete or the measurements too imprecise.

Over the years, critical comments have been made about the utility of some of Bradford Hill's considerations.[12,29,37] For instance, 'specificity' as originally meant by Bradford Hill is no longer considered by epidemiologists, although specificity in another sense still is. Furthermore, he did not assign values to his considerations, whereas nowadays some considerations are given more weight in the evidence synthesis than others. In addition, in modern epidemiology, the use of Directed Acyclic Graphs (DAGs) to identify causal pathways is increasing rapidly.
The committees monitor these developments and adjust their assessments accordingly.

4.2.2 GRADE framework

A more recently developed tool to evaluate the strength of evidence from individual epidemiological studies is the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework, which is widely used by many organisations.[18] GRADE was developed to assess the certainty in evidence (quality of evidence) and the strength of recommendations. It was predominantly developed for recommendations on interventions in clinical health care. This has raised the question, also posed by the committees, to what extent it could be applied, with or without adjustments, for recommendations in environmental and occupational health.[48] These fields often make use of other types of data and studies than in health care, such as more observational studies, animal experiments, *in vitro* test systems, and mechanistic data.
All of these types of data and studies add to the weight of evidence. Furthermore, reliable quantitative exposure-response data are needed to recommend on standards, such as OELs. Because of the previous remarks, the committees have not used yet the GRADE framework. However, they are aware that efforts are being made to adapt the GRADE framework to environmental and occupational health. The committees are monitoring these developments for possible future use.

4.2.3 Triangulation and expert judgement

Bradford Hill's considerations focus on individual epidemiological studies. The committees, however, want to weigh the strength of evidence by

integrating the findings of all the individual epidemiological studies combined (a form of triangulation). That is why the committees support the concept of triangulation. Others have suggested such an approach for environmental and occupational observational epidemiology.[37,49] It extends Bradford Hill's approach by considering evidence from all relevant epidemiological studies, irrespective of the design or context, so that not only the overall strength of evidence can be evaluated, but also the overall direction of various possible biases.

4.2.4 Integration of various lines of evidence

The foregoing focuses on the line of evidence from epidemiological studies, whereas the ultimate goal for the committees is to integrate various lines of evidence to come to an overall judgement on the strength of evidence for a particular effect. Such a judgement is based on the committees' conclusion about the quality of the individual studies, and the extent to which the various lines of evidence complement each other. Essential to the judgement is the consideration of uncertainty in the overall evidence due to uncertainties in the values of exposure levels and effect estimates, or because data may not be available to cover all aspects. Because of the complexity of integrating lines of evidence, it is difficult to capture this process in detail in a framework or tool. It is, therefore, inevitable that the judgement on the strength of evidence depends on expert judgement.

The US EPA developed an Integrated Science Assessment (ISA) tool, and used it for a number of environmental contaminants, to serve as a scientific foundation for the review of U.S. ambient air quality standards.[50] The ISA provides a concise review, synthesis and integration of evidence that is helpful to judge causality. It is not a protocol, but describes the aspects that need to be taken into account to arrive at a conclusion on the overall strength of evidence. The latter is expressed in five descriptors (likelihood of causal relationships) regarding health effects in humans, and ecological and other welfare effects. The committees will also consider the aspects that are described in the ISA.

In addition, the scientific criteria for classifying a substance as mutagenic, carcinogenic, or toxic to reproduction in humans, are to a certain extent based on the integration of different lines of evidence (see Chapter 5).

05 classification of substances

The criteria for the classification categories are based on the Globally Harmonized System, which has been incorporated into the system and guideline used by the European Union for the classification, labelling, and packaging of substances and mixtures (Regulation EC 1272/2008: Section 3.6, Carcinogenicity; Section 3.5, Germ cell mutagenicity; and Section 3.7, Reproductive toxicity). Although the criteria mentioned in the EU Regulation are set for substances that are evaluated according to the CLP regulation, the Health Council also considers the criteria useful in recommending classifications for individual substances, mixtures and emissions for which the Regulation does not apply.

5.1 Carcinogenic and mutagenic substances

Classification proposals for carcinogens and germ cell mutagens are prepared by the Subcommittee on the Classification of Carcinogenic Substances. The proposals are based on a hazard assessment, in which the carcinogenic properties, the genotoxic properties, and the carcinogenic mode of action are extensively evaluated on study quality and strength of evidence.

5.1.1 Classification for germ cell mutagenicity

The strength of evidence for germ cell mutagenicity is expressed in the following EU classification categories:

- Category 1A
  Substances known to induce heritable mutations in the germ cells of humans (H340)
- Category 1B
  Substances to be regarded as if they induce heritable mutations in the germ cells of humans (H340)
- Category 2
  Substances which cause concern for humans owing to the possibility that they may induce heritable mutations in the germ cells of humans (H341)
- No classification for germ cell mutagenicity

EU hazard statement codes:

- H340
  May cause genetic defects
- H341
  Suspected of causing genetic effects

In proposing a classification, the subcommittee takes into account that many *in vitro* and *in vivo* genotoxicity tests have been developed over the years (OECD guidelines) and used to classify a substance for its genotoxic properties. The OECD guidelines are periodically updated to include new scientific insights. As a result, the OECD has withdrawn some of these tests or downgraded them to indicator tests in recent years, because questions has arisen on their validity. Examples are tests for unscheduled DNA synthesis, sister chromatid exchanges, and the mouse

heritable translocation test. The subcommittee still includes studies using these tests in its evaluation, but gives them a lower weight in the evidence synthesis for genotoxicity.[51-53]

5.1.2 Classification for carcinogenicity

The strength of evidence for carcinogenicity is expressed in the following (EU) classification categories:

- Category 1A
  Known to be carcinogenic to humans (H350)
- Category 1B
  Presumed to be carcinogenic to humans (H350)
- Category 2
  Suspected to be carcinogenic to humans (H351)
- No classification for carcinogenicity

EU Hazard statement codes:

- H350
  May cause cancer
- H351
  Suspected of causing cancer

The criteria for carcinogenicity are described in detail in the *Guideline to the classification of carcinogenic compounds* (Health Council, 2010).[2] In addition to the EU categories, the subcommittee takes two subcategories for 'no classification for carcinogenicity' into account. When a substance is not classified, these categories explain the reason for not classifying:

- The available data are insufficient to evaluate the carcinogenic properties of the substance (category 3)
- The substance is probably not carcinogenic to humans (category 4)

In addition, for substances classified in categories 1A or 1B, the subcommittee judges the mode of action to aid future decisions on the health-based OEL for the substance regarding whether this OEL should be risk- or threshold based. The modes of actions are as follows:

- *Direct-acting genotoxic carcinogen (stochastic genotoxic carcinogen)*.
  This includes substances that either in their unchanged form or as reactive metabolites interact directly with DNA, causing DNA damage (adducts, single and double-strand breaks) resulting in irreversible gene mutations. These carcinogens do not show a threshold exposure level below which no risk for a carcinogenic effect exists. As they act by a stochastic process, a risk-based approach to deriving a health-based OEL is appropriate.
- *Indirect acting genotoxic carcinogen (non-stochastic genotoxic carcinogen)*. This includes substances that do not interact directly with DNA, but can ultimately damage DNA indirectly (e.g. by inhibition of DNA repair, affecting spindle apparatus, and inhibition of topoisomerases). As a threshold could be identified for these mechanisms, a threshold-based approach to deriving a health-based

OEL is appropriate.
- *Non-genotoxic* carcinogen. This includes substances that are capable of promoting various phases of the cancer process without damaging DNA, either directly or indirectly. Such compounds are known as tumour promoters. For these carcinogens, a threshold-based approach to deriving a health-based OEL is appropriate.
- *Genotoxicity has been investigated insufficiently*.

In proposing a classification, the subcommittee notes that if a substance induces malignant tumours, it will usually constitute sufficient evidence of carcinogenicity, whereas induction of benign tumours usually constitutes a lower strength of evidence. However, some benign tumours may have the potential to progress to malignant tumours, or may in themselves be of concern. Based on current knowledge and using expert judgement the subcommittee will judge whether the benign tumours observed should be given more weight in the evidence synthesis on carcinogenicity than is normally done.

5.2 Reproduction toxic substances

Classification and labelling proposals for reproduction toxic substances are prepared by the Subcommittee on the Classification of Substances Toxic to Reproduction. The proposals are based on a hazard assessment, in which adverse effects on fertility and offspring development, and adverse effects on or via lactation, are evaluated on study quality and weight of evidence.

Classification for effects on fertility and offspring development

The classification of substances with adverse effects on fertility and offspring development is the result of an integrated assessment of the nature of all parental and developmental effects observed, with their specificity and adversity.

5.2.1 Classification for effects on fertility and offspring development

The weight of evidence for effects on fertility (F/f) and offspring development (D/d) is expressed in the following EU classification categories:

- Category 1
  Known or presumed human reproductive toxicant (H360(F/D))
- Category 1A
  Known human reproductive toxicant
- Category 1B
  Presumed human reproductive toxicant
- Category 2
  Suspected human reproductive toxicant (H361(f/d))
- No classification for effects on fertility or development

EU Hazard statement codes:

- H360F
  May damage fertility
- H360D
  May damage the unborn child
- H361f
  Suspected of damaging fertility
- H361d
  Suspected of damaging the unborn child
- H360FD
  May damage fertility. May damage the unborn child
- H361fd
  Suspected of damaging fertility. Suspected of damaging the unborn child
- H360Fd
  May damage fertility. Suspected of damaging the unborn child
- H360Df
  May damage the unborn child. Suspected of damaging fertility

With regard to the EU guideline, the subcommittee takes into account a number of additional considerations for the relevance of the available literature. Concerning male fertility, for instance, the subcommittee considers data on parameters related to fertility, such as spermatozoa concentration and motility. The subcommittee excludes publications containing only data on sex hormone levels, because the relationship between these hormone levels and functional fertility (the ability to conceive children) is uncertain.

In proposing a classification, the subcommittee takes the following additional considerations to the EU criteria into account:

- If sufficient evidence is available to establish a causal relationship between human exposure to the substance and impaired fertility or developmental toxic effects in the offspring, the compound will be classified in category 1A, irrespective of the general toxic effects (see Regulation (EC) 1272/2008, 3.7.2.2.1.).
- Adverse reproductive effects, reported without information on the paternal or maternal toxicity, may lead to a classification other than category 1B, if the effects on reproduction occur at dosage levels in which general toxicity studies showed severe toxicity in other organs.
- Clear adverse reproductive effects will not be disregarded on the basis of reversibility per se. The committee does not only use guideline studies (studies performed according to OECD standard protocols) for the classification of compounds, but also non-guideline studies.

5.2.2 Classification for effects on or via lactation

The EU criteria for classifying substances for effects on or via lactation dictate that substances that are absorbed by women and have been shown to interfere with lactation, or which may be present (including metabolites) in breast milk in amounts sufficient to cause concern for the

health of a breastfed child, shall be classified and labelled. Unlike the classification of substances for fertility and developmental effects, which is based only on hazard identification (largely independent of dosage), the labelling for effects on or via lactation is based on risk characterisation. Therefore, it also includes a consideration of the level of exposure of the breastfed child. Consequently, a substance is labelled for effects on or via lactation when it is likely that the substance would be present in breast milk at potentially toxic levels. The subcommittee considers a concentration of a compound as potentially toxic to the breastfed child when this concentration exceeds the exposure limit for children, or, if that level is unknown, when it exceeds the exposure limit for the general population, i.e. the acceptable daily intake (ADI).

Classification for lactation:

- Effects on or via lactation (H362)
- No labelling for lactation

EU Hazard statement code:

- H362

  May cause harm to breast-fed children

5.3 CMR list

The Minister of Social Affairs and Employment considers the recommendations of the subcommittees to decide on the inclusion of the substance in the official Carcinogens, Mutagens and Reproduction toxic substances (CMR) list in the Dutch Working Conditions Act. For carcinogens and mutagens, the CMR list includes substances (and processes) classified in category 1A or 1B, as indicated in Annex I of the EU Regulation 1272/2008 of the European Parliament for the classification, labelling and packaging of substances and mixtures (Annex VI, CLP-REACH Regulation).[54] In addition, the CMR list includes a non-limitative list of reproduction toxic substances for which additional registration obligations apply (classification categories 1A, 1B, and 2, plus a mark [letter 'x'] for effects on or via lactation).

06 health-based occupational exposure limits

The DECOS recommends health-based OELs for hazardous substances to which workers can be exposed in the workplace. These OELs are average concentrations of a substance in the air of a workplace that should offer sufficient protection against possible adverse effects. The OELs are obtained by a quantitative risk analysis in which the DECOS has to make decisions on:

- critical adverse effect(s);
- the type of OELs to be set;
- choosing a threshold- or risk-based approach in deriving OELs;
- the choice of the key study or studies as point of departure in deriving OELs;
- the derivation method;
- the use of adjustment factors and defaults.

6.1 Critical adverse effects

A health-based OEL is an exposure level below which no or almost no significant adverse effects are expected in the course of and after working life, or in workers' offspring. It does not matter whether the effect is local or systemic, or whether the effects occur immediately upon exposure or in the medium or long term. Which adverse effects are considered in deriving a health-based OEL depends on the outcome of the quality evaluation of the individual studies and whether the evidence for a causal relationship or association is sufficiently demonstrated. In addition, a health-based OEL is derived from data on the adverse effect that occurs first at increasing exposure, also called the critical effect, assuming that other adverse effects are then also prevented.

Regularly, the Minister of Social Affairs and Employment specifically asks the Health Council to advise health-based OELs for known carcinogens. In this case, the DECOS derives a health-based OEL to prevent cancer, and evaluates whether that OEL also protects against other possible adverse effects. If the data shows that other adverse effects could occur below this limit, the DECOS decides not to derive a health-based OEL from cancer data, but from the first adverse effect apparent at increasing exposure.

6.2 Types of exposure limits

The DECOS differentiates several types of exposure limits.

8-hour time-weighted average concentration (8-h TWA)

The most commonly used health-based recommended OEL is an 8-hour time-weighted average concentration, which reflects an average working day of 8 hours.

Short-term exposure limit (STEL)

In addition, the DECOS assesses whether workers should also be protected from adverse effects caused by peak exposure. A STEL represents a mean concentration measured over an arbitrary period of

15 minutes that should not be exceeded (15-minute TWA). Peak exposure may cause adverse effects that occur immediately or in the short- or long-term.

Ceiling value

For substances that may immediately cause life-threatening health effects, the DECOS can derive a ceiling value. In such cases, exposure should never exceed the ceiling value at any time.

Biological limit values (BLV)

There are situations in which the air-based exposure limits do not provide sufficient protection. For example, when a substance accumulates in the body (for years up to the total life-time, irrespective of the route of exposure; by other exposure routes), it can induce effects after inhalation exposure has been lowered or discontinued, because it is still present in the body. In those cases, a BLV is derived, indicating the concentration of a substance in a biological medium (e.g. blood, urine).

6.3 Threshold- and risk-based approach

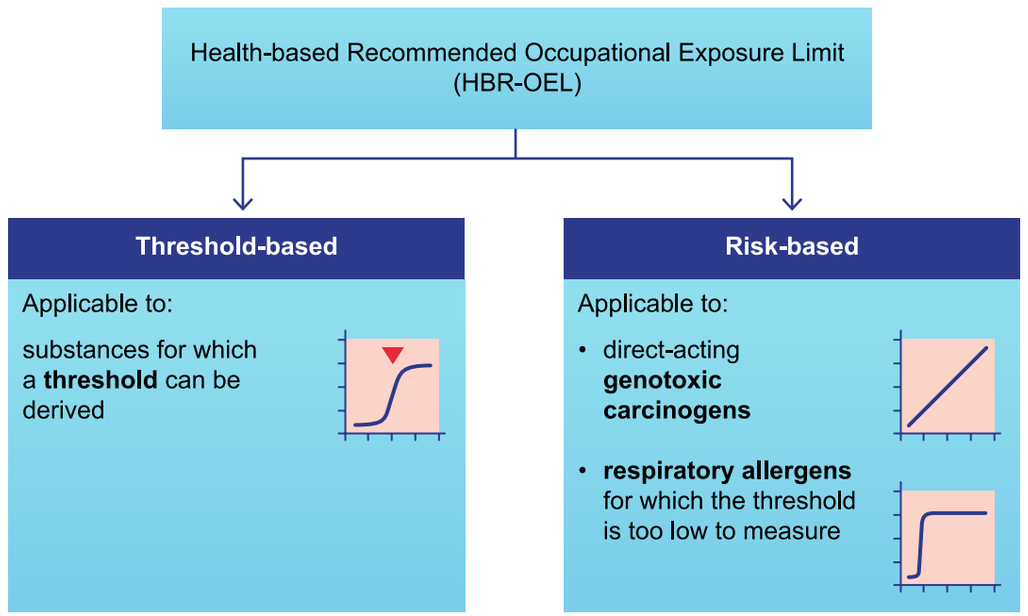
Whether a safe OEL can be set depends on the mode of action by which the substance induces the critical effect, and the reversibility/irreversibility of the key event(s) leading to this critical effect. Based on this information, the DECOS uses one of the two approaches to derive health-based OELs:

- a threshold-based approach;
- a risk-based approach.

6.3.1 Threshold-based approach

For many adverse health effects, such as irritation and fibrosis, a threshold exposure level exists at and below which no significant effect is observed. It can be mechanistically explained that a cut-off point exists. The threshold-based approach implies that the risk analysis is focussed on estimating the threshold exposure level from the available data, and that the DECOS uses this threshold level as the point of departure in deriving a health-based OEL.

HBR-OELs are derived using either a threshold approach or a risk-based approach, depending on the substance's critical effect



6.3.2 Risk-based approach

Sometimes, a threshold-based OEL cannot be assigned when exposure always poses a certain health risk, or when the threshold is so low that it cannot be detected by currently available monitoring tools. This is the case for direct-acting genotoxic carcinogens and certain allergens, respectively. The DECOS prefers then a risk-based approach, in which it derives a health-based OEL by estimating a concentration level that correspondents with a predefined extra risk level. The extra risk level is a disease risk due to occupational exposure, in relation to the disease risk when not occupationally exposed (background risk of the general population). The extra risk levels are predefined by the Minister of Social Affairs and Employment.

Direct-acting genotoxic carcinogens

The property of a substance to induce gene mutations by direct interaction with the DNA is considered a one-hit (stochastic) occasion. Once a pre-mutagenic lesion is missed by the repair and defence mechanisms of the body, it can be irreversibly fixed as a mutation. This means that one molecule of the substance is theoretically capable of causing cancer.
As a consequence, any exposure involves a risk of getting cancer and no threshold exposure level can be derived below which there is no cancer risk.

For direct-acting genotoxic carcinogens, two predefined extra cancer risk levels are set:

1. *Target risk level*: four additional cases of cancer due to occupational exposure over a 40-year working life period per 100,000 cases from all causes. This equates to a risk of $4x10^{-5}$ associated with exposure throughout a person's working life.
2. *Prohibition risk level*: four additional cases of cancer due to occupational exposure over a 40-year working life period per 1,000 cases from all causes. This equates to a risk of $4x10^{-3}$ associated with exposure throughout a person's working life.

At the time that the extra cancer risk levels were predefined, cancer was considered a fatal disease with very low 5-year survival rates. Therefore, target and prohibition risk levels were mainly based on mortality data. Due to improved early detection and cancer treatment, current 5-year survival rates for many cancer types have increased and epidemiological research is increasingly based on cancer incidence. Working conditions policy is primarily focused on protecting employees from the occurrence of disease, in this case cancer, regardless of associated mortality. Therefore, the DECOS currently prefers incidence statistics over mortality statistics in deriving cancer risk-based OELs. An additional argument is that registration of incidence data is generally more reliable than registration of mortality data.

Allergic respiratory disorders

Exposure to allergens may lead to allergic respiratory disorders.

The Health Council considers allergic sensitisation the critical event in developing allergies, because once sensitisation has occurred, it is irreversible and always poses a risk of allergic symptoms upon continuing exposure. Based on the scientific knowledge on allergic immunological mechanisms, therefore, the Health Council has concluded that it is plausible that a threshold level for these substances exists below which no allergic sensitisation may be expected.[3] However, the council has also concluded that the threshold level for preventing allergic sensitisation may be so low that a very small amount of an allergen is needed to provoke an allergic sensitisation. This implies that, for certain allergens, the threshold will be too low to be detectable with available exposure monitoring techniques. Consequently, when the DECOS derives a health-based OEL for allergic substances, it will first consider whether exposure-response data are available that show a clear threshold level in the lowest exposure range (threshold-based approach). If that is not the case, the DECOS will use a risk-based approach. For the risk-based approach, the Minister of Social Affairs and Employment has set an additional risk level of 1% due to occupational exposure, which adds to the background risk level of developing allergic sensitisation to the substance when not occupationally exposed (the general population). See the advisory report of the Health Council for a more detailed description of the additional risk levels for allergens.[3]

6.4 Key study or studies as point of departure in deriving health-based OELs

Based on the quality evaluation of the individual studies, and the results from the evidence synthesis, the DECOS decides on the critical effect, and assesses which study or studies that have described this effect can be used to derive a health-based OEL. Only those studies are selected that reported quantitative exposure and response data to model exposure-response relationships. The DECOS uses these models to determine the point of departure in deriving a health-based OEL. Sometimes there are multiple studies with such data. In that case, the DECOS gives preference to studies with:

- data from observational studies (cohorts and cross-sectional studies) performed in workplaces, because these reflect actual exposure and working conditions. If these are not available or of low quality, the committee assesses whether case-control studies or animal experiments are available;
- data of adverse effects that could become manifest after a long latency period (even after reaching retirement), for instance cancer, because OELs should protect against hazardous health effects during and after the entire working life. This is under the assumption that possible acute, short or mid-term effects will also be prevented if a health-based OEL is derived from these data;
- data on exposure by inhalation, because OELs are average concentrations of substances in the air. Exposure data on inhalation

concentrations are not always available. To derive a health-based OEL, the DECOS decides whether data from oral exposure can be used as second best. In rare cases and in the absence of inhalation or oral intake studies, the DECOS evaluates whether studies with dermal application can be used. Animal experiments in which animals are exposed to a substance by intraperitoneal or intravenous injections, or by intratracheal installation, are not used, because the route of exposure is not relevant for workers, and questions may arise on kinetics and systemic availability;

- exposure data of several exposure groups, preferably with at least one exposure group in the exposure range close to the intended OEL and down to virtually no exposure.

If multiple suitable observational studies are available, the DECOS decides case by case whether a meta-analysis can serve as a point of departure. In addition, if multiple animal experiments of comparable suitability are available, the study resulting in the lowest point of departure is considered the key study to derive a health-based OEL.

6.5 Derivation methods

The DECOS currently uses two methods to derive health-based OELs, namely the benchmark-dose method (BMD method) and the survival analysis.

6.5.1 Benchmark-dose method (BMD)

The BMD method involves deriving a best fit of the shape of the exposure-response relationship and calculating a benchmark dose (BMD). The BMD is a dose or concentration that corresponds to a response, the benchmark response (BMR); the BMDL is the lower limit of the 95% confidence interval of the BMD, the BMDU the upper limit. The BMD and BMDL/BMDU depend on all available data and on the shape of the curve that best describes the exposure-response relationship. A few scientific bodies, such as the European Food Safety Authority (EFSA) and the US Environmental Protection Agency (US EPA) have published guidelines on the BMD method for quantitative risk assessment purposes.[55-58] Earlier, the Health Council also published an advisory report on the use of the BMD method instead of the no-observed-adverse-effect-level or lowest-observed-adverse-effect-level (NOAEL/LOAEL) approach.[59] In this respect, the DECOS broadly concurs with EFSA's working method.[57] Therefore, the DECOS uses the EFSA guideline for the application of the BMD method, and refers to this guideline (and manual) for a detailed explanation of the technical implementation.

Besides using the EFSA guideline, the DECOS has made a number of additional considerations:

- *BMD or BMDL as point of departure*
  The position of the BMDL in relation to the BMD reflects the uncertainty about the BMD due to the number of exposure groups chosen, group size, and distribution of data within a group. Because the uncertainty of data is discounted in the BMDL, and in line with the EFSA (and US EPA) the DECOS considers the BMDL to be the best starting point to derive threshold-based OELs. However, for risk-based OELs, the DECOS considers the BMD as the best estimate, because the method of the risk-based approach results in a conservative health-based OEL, so that the statistical uncertainties need not be taken into account anymore.
- *Benchmark response (BMR)*
  The BMR (or critical effect size [CES]) is the boundary between what is considered a physiological or biological response and what is considered an adverse health effect. Partly based on scientific knowledge and insights, this has to be considered for each effect endpoint. There are no schemes worldwide that assign the best BMR per effect, although the EFSA has set a global guideline.[56,57] The DECOS follows the EFSA guideline, but deviates from these BMR values if there are grounds for this. For instance, it considers body weight increases or decreases by more than 10% instead of 5% (continuous data) as a relevant effect. Additional factors that the DECOS takes into account in assessing the BMR includes the size and the power of the study.
- *Observational studies*
  Although the BMD method was primarily developed for experimental research data, the DECOS also considers it suitable for data from observational studies. However, the BMD analysis of observational exposure-response data can be more complicated than that of typical experimental data. For instance: group exposure categories are sometimes imbalanced in group size; sometimes 'time to development of a disease or death in a population' (lag times) should be taken into account; and exposure categories are sometimes defined in ranges instead of one average exposure value, for example 'lower than 1 mg/m^3', 'between 1 and 5 mg/m^3', and 'higher than 5 mg/m^3'.
  The DECOS addresses all these factors in the BMD analysis.

The DECOS uses available BMD software programmes (PROAST; US EPA BMDS) to analyse the mathematical model (curve) to characterise the shape of the exposure-response relationship, and to calculate BMDs, BMDLs and BMDUs. In the case of data from observational studies, other more sophisticated programmes (e.g. SAS, R, or STATA) are used.

The use of the BMD in the risk-based approach

The DECOS follows several steps to derive a risk-based OEL from animal data (if no suitable epidemiological data are available). These steps include calculating the lowest BMD, followed by estimating the carcinogenic activity (expressed as incidence per unit of the daily dose [oral exposure], or per unit of air concentration [inhalation]) in the animals, and extrapolating the carcinogenic activity by a linear model to estimate the cancer risk in humans. Details on the estimation method is given in the Guideline calculating cancer risk values, published by the Health Council.[1] A number of assumptions are made in the calculation formulas about, for example, the average weight and life expectancy of the animals from which the data originated, the average weight of the healthy worker, tidal volumes, duration of exposure, and working life period. These are expressed in default values, unless data on these factors are presented in the selected studies. Regarding humans, the default values are as follows:

- The duration of life time exposure is 75 years of exposure, 24 hours per day, 7 days per week, 52 weeks per year, with an average inhalation volume of 18 m^3 per 24 hours.
- The duration of working life time exposure is 40 years of exposure, 8 hours per day, 5 days per week, 48 weeks per year, with an average inhalation volume of 10 m^3 per 8-hour working day (moderate working activities).
- For body weight, a default value of 70 kilograms is used.

The DECOS considers it important that the preconditions and principles of the applied BMD analysis are reported properly, so that third parties are able to ascertain how the analysis came about. It therefore adds the data from the analysis in an appendix to the advisory reports.

In the event that the exposure and response data are insufficient to perform a proper BMD analysis, the DECOS will assess whether a health-based OEL (threshold- or risk-based) can be derived using the traditional NOAEL or LOAEL approach.

6.5.2 Survival analysis

For carcinogenic substances for which no threshold can be established, but suitable data from observational studies are available, the DECOS uses a survival analysis to derive a risk-based OEL. An advantage of this method is that diseases or other causes of death can be accounted for. In a group of workers exposed to a carcinogen, other causes of death will lead to a reduction of the population at risk over time, and therefore to a lower number of additional cancer cases by the substance in absolute terms. If this is not corrected for, the probability of overestimating the risk is high. Furthermore, by using life tables it is possible to take into account time- and age-dependent factors in the development of cancer.

In the Health Council guideline to derive risk-based OELs, survival analyses are used by applying life tables.[1] In the guideline, stepwise

details are given on the method and the factors to take into account when performing a survival analysis. For instance, it is important to extend the life table to an age at which the mortality burden of occupational exposure to a carcinogenic compound is negligible, compared to the mortality due to other causes. For this purpose, the DECOS currently adheres to the age of 100 years, but in the literature survival analyses are also performed with life tables up to the age of 75, 80 or 90 years.[31] There is no wide consensus on the maximum age, but the most important consideration is that the life table ends near the maximum lifetime, or somewhere near the mean or median lifetime. Using life tables also makes it possible to take into account time- and age-dependent factors in the development of cancer, and cancer risks that increase or decrease after a certain period following exposure. The survival analysis results in cumulative exposure levels corresponding to one of the two predefined additional risk levels, concerning a working life period of 40 years (see Section 6.2.2). Technical details on the survival analysis are given in an appendix of the respective advisory reports.

6.6 Route-to-route extrapolation

Route-to-route extrapolation is applied when no exposure data on inhalation are available and the committee relies on oral exposure data or in rare cases on dermal exposure data. The DECOS only applies route-to-route extrapolation if the critical adverse effect is systemic in nature, and no critical differences in effect exist between routes (e.g. first-pass effect). This follows the proposals by other scientific bodies, such as the ECHA, ECETOC and the German BAuA.[60-62]

6.7 Adjustment factors

Adjustment factors (also called uncertainty, assessment or safety factors in the literature), expressed as numbers, are often used in toxicological risk assessment to account for the uncertainty in differences between experimental animal species and humans (interspecies differences); differences in sensitivity between humans (intraspecies differences); imperfections and uncertainties in data and design; and differences between the experimental exposure conditions and the exposure conditions in the workplace (e.g. actual duration, frequency and pattern of exposure; route of exposure). The use of adjustment factors is the last step to derive a health-based OEL, in which the health-based OEL is derived by dividing the BMDL by the product of all adjustment factors.

The factors should represent realistic values, but there are no universal harmonised adjustment factors available. The use and their values may differ between scientific bodies because they reflect the uncertainty in scientific knowledge, and therefore expert judgement plays a great role in choosing the height of the values. On the other hand, when no substance and species specific data are available, there has been some consensus over the years about what the levels of default adjustment factors should be (e.g. ANSES, BAuA, ECETOC, ECHA, RAC, SCOEL, US EPA,

WHO).[10,13,57,60,62-65] The main adjustments considered by the DECOS are summarised below.

Interspecies differences

Regarding systemic adverse effects, interspecies differences in body size and other species-specific differences (toxicokinetic and toxicodynamic differences) are taken into account to extrapolate animal data to humans. The adjustment factor for interspecies differences is calculated by multiplying the factor for body size differences with the factor for the remaining differences. The DECOS considers it inappropriate to adjust for differences in body size for local skin and respiratory tract effects. For the adjustment of differences in body size, see Section 6.8.2 on allometric scaling.

For the extrapolation of the remaining species-specific differences, the committee uses a default factor of 2.5, unless data indicate otherwise. The value of this default is based on (cumulative) ratio distribution profiles from empirical research. A default value of 2 is recommended when the BMD(L) serves as the starting point in deriving a health-based-OEL.[63]

Intraspecies differences

In assessing health standards for the general population, an adjustment factor of 10 is used to account for differences in response between people of all ages. The variability in response is due to differences in biological, life-style and environmental factors. The worker population is considered more homogeneous, because it does not include children, elderly, and (generally) persons with a weak health. Therefore, the DECOS uses a lower default value for the worker population, namely a value of 5, unless data indicate otherwise. This value holds for systemic and local effects. The ECETOC proposes a lower value (factor of 3), whereas the ECHA and the Anses also use a factor of 5 by default.[8,60,61] In 2019, the German BAuA analysed the coverage of the distribution for intraspecies extrapolation for these default values, and calculated that for the scenario of 5% incidence in the general population, the value of 3 would cover 36.6% of cases and the value of 5 would cover 73.4% of cases.[62]

In the case of embryotoxic and teratogenic effects, no distinction can be made between progeny of the general and occupational population. In such cases, therefore, the DECOS applies a factor of 10 for intraspecies differences.

Duration of exposure

The DECOS uses adjustment factors to extrapolate data from subacute to subchronic/chronic and for subchronic to chronic exposure, to take into account that other and more serious adverse health effects may appear with increasing exposure time (during or after exposure has been discontinued). The adjustment factor is derived from the whole toxicity profile. It ranges between a factor of 1 and 5 for extrapolation of subacute

exposure to subchronic exposure, and the same range for extrapolation of subchronic to chronic exposure. The ECETOC (2003) and the ECHA (2012) have recommended using a default factor of 6 extrapolation of subacute to chronic exposure, and a factor of 2 for extrapolating subchronic to chronic exposure.[60,61] The DECOS uses these default factors, unless data indicate otherwise. The extrapolation for duration of exposure is only applied for systemic effects and for local effects in the respiratory tract.

Other adjustment factors

Other uncertainties that the DECOS takes into account include extrapolation from a LOAEL to a NOAEL; the health consequence (biological significance) of the adverse health effect on which a health-based OEL will be derived; the reliability of the steepness of the slope of the exposure-response relationship; and the confidence in data used to derive a health-based OEL. The use and weight of the adjustment factors to adjust for these issues depends on the reliability and availability of the data. The DECOS uses a factor of 1 (no adjustment needed) for each of these issues as standard, except when extrapolating a LOAEL to a NOAEL (factor of 3), unless data indicate otherwise.

Overall adjustment factor

An overall adjustment factor is established by multiplying the separate factors. The DECOS is aware that it is not possible to distinguish all these factors, because some factors are not independent of each other. Therefore, straightforward multiplication may lead to an unreasonably high overall adjustment factor. To establish a justifiable overall factor, the DECOS discusses and weighs the individual factors case by case.

6.8 Default values for physiological and morphological parameters

When certain data are not reported, the DECOS uses default values for several parameters, such as for inhalation volumes, food and water consumption, body weight, surface areas, and for conversion of doses to concentrations (and vice versa). Based on the experiences and (scientific) findings of others, the DECOS describes what its starting points are for using these default values below. [60,63-66]

6.8.1 Workers

To extrapolate animal data to humans, the DECOS uses default values that reflect an average of a normal healthy worker and working conditions. The default worker weighs 70 kilograms, has a body surface of 2 m^2, inhales 10 m^3 during an 8-hour working day (1.25 m^3 per hour) under light working activities, and works on average 8 hours per day, 5 days per week, 48 weeks per year, for 40 years. The use of defaults is necessary to ensure that the health-based OELs for each substance under investigation have the same starting points. The default values for age, body weight, body surface, inhalation volumes for animals and humans, and water and

feeding data for animals were calculated by the US EPA by using and comparing data in the literature. Because default values must be easy to use, these calculated values are rounded off to a higher or lower value, taking worst-case scenarios into account.[66,67] For example, in 1996 the US EPA recommended a default body weight of 70 kilograms for a worker (man [77 kg] and women [62 kg] combined), which is based on a large population in the USA. The DECOS has applied this calculated default value for its occupational risk analyses, because it does not expect US figures to differ much from European or Dutch figures due to comparable living and working conditions.[66,67] Another example is the inhalation rate of 1.25 m^3 per hour for workers. The inhalation rate depends on age, sex and working activities. Taking into account these differences, the literature assumes an average inhalation volume of 10 m^3 per working day to be a relevant inhalation volume for light working activities. Since it is assumed that workers will not endure moderate (estimated average inhalation volume of 1.7 m^3/hour) and heavy working activities (estimated average inhalation volume of 2.8 m^3/hour) for an entire working day, the DECOS also uses the inhalation volume for light activities as the default value in its risk analyses.[66,67]

Health-based OELs are based on a group approach and under average working conditions. The DECOS is aware that these conditions (e.g. working hours, heaviness of physical work, working life period) and job history (change jobs) vary between individual workers. Workers also differ individually in physiological and biological status (e.g. body weight, metabolism, biological defence mechanisms) and in life style and health conditions. Consequently, one worker may be more vulnerable or sensitive to developing an exposure-related disease than the other. If there are indications of such differences in vulnerability and sensitivity, the DECOS will comment on this in its advisory report (groups at extra risk). To a certain extent, differences in sensitivity are taken into account by applying an adjustment factor for intraspecies differences.

6.8.2 Species-specific allometric relationships

Allometric relationships arise when you compare several animal species on certain morphological characteristics, for example body size, body surface area, inhalation volume and rates, and water and food consumption. It often turns out that such a characteristic does not increase in direct proportion to body weight. In the literature on occupational risk analyses, proposals are made and equations are set for allometric scaling of these morphological characteristics from animal species to humans. The DECOS also uses these proposals (defaults) and equations, for instance in assessing a skin notation (see Chapter 7) and for route-to-route extrapolation. In addition, allometric adjustments are made for differences in body size by taking into account differences in caloric requirements. Caloric requirements are proportional to the allometric body weight scaling, with a factor of around 0.7.[63,66] Examples of default adjustment factors for body size are 7 for mice and 4 for rats.

Differences in body size are not taken into account when it concerns inhalation studies, because the extrapolation is based on the toxicological equivalence of a concentration of a substance that animals and humans breathe at a rate depending on their caloric requirements. If the DECOS is of the opinion that allometric relationships are applicable for deriving health-based OELs, it will provide a detailed explanation of the methodology with references to the sources.

07
skin notation

When the DECOS recommends a health-based OEL, it also evaluates whether a skin notation (specified by the letter H for 'huid' [skin in Dutch]) for a substance should be recommended. Such a skin notation indicates a substantial contribution of dermal exposure to systemic adverse effects on which a health-based OEL is derived. A skin notation is given next to a health-based OEL, but is not a substitute for it.

Some chemical substances cause adverse dermal effects upon dermal exposure. If the mode of action or the pathogenesis make it plausible that these local dermal effects can ultimately lead to systemic effects, the DECOS will consider recommending a skin notation.

To recommend a skin notation, data on dermal (percutaneous) absorption and physicochemical characteristics (e.g. physical form) are required. Validated OECD methods exist to investigate dermal absorption. Organisations, such as the EFSA, ECETOC and NIOSH have set up guidelines or strategies to assess dermal absorption.[68-71] Overall, the DECOS uses the approach described by the ECETOC, which bases its approach on a report by the Dutch Medical Biological Laboratory of TNO on the assignment of a skin notation in the list of MAC-values (1989; at the request of the DECOS).[68,72] A skin notation is considered necessary when the amount absorbed by both hands and forearms in 1 hour could amount to more than 10% of the amount that can be absorbed by the lungs on exposure to the recommended health-based occupational exposure limit for 8 hours. The dermal absorption is estimated for intact, undamaged skin.

The absorption rate of a substance can be estimated from *in vivo*, *ex vivo*, and *in vitro* models. Preference is given to human *in vivo* studies.
The validity of the model used is assessed on a case-by-case basis, and by checking whether the absorption study was performed according to the OECD guidelines. As a default, the committee assumes that the area of the hands and forearms is 2,000 cm^2, dermal exposure is 1 hour per working day, a volume of 10 m^3 of air is inhaled in an 8 hour-working day, and 50% of the substance that is inhaled is absorbed by the lungs, unless data indicate otherwise. The DECOS notes that NIOSH uses other default values regarding the dermal area (two hand palms, 360 cm^2), dermal exposure duration (8 hours) and the substance absorption rate by the lungs (75%).[71] However, in the end both calculations do not differ significantly from each other (ratio of skin to inhalation (SI) dose: $SI\ ratio_{ECETOC} = 1.04 \times SI\ ratio_{NIOSH}$).

references

1 Health Council of the Netherlands. *Guideline for the calculation of risk values for carcinogenic compouds*. The Hague, the Netherlands: Health Council of the Netherlands; publication no. 2021/16E, 2012.

2 Health Council of the Netherlands. *Guidline to the classification of carcinogenic compounds*. The Hague, The Netherlands: Health Council of the Netherlands; publication no. A10/07E, 2010.

3 Health Council of the Netherlands. *Prevention of work-related airway allergies. Recommended occupational exposure limits and periodic screening*. The Hague: Health Council of the Netherlands; publication no. 2008/03E, 2008.

4 Health Council of the Netherlands. *Standards for dermal exposure at the workplace*. The Hague: Health Council of the Netherlands; publication no. 2001/28, 2001.

5 Health Council of the Netherlands. *Adjustment of occupational exposure limits in case of unusual work schedules*. The Hague: Health Council of the Netherlands; publication no. 2001/06OSH, 2001.

6 Health Council of the Netherlands. *Afleiding gezondheidskundige advieswaaren. Toxicologische advieswaarden voor blootstelling aan stoffen. [Toxicology-based recommende exposure limits]*. Rijswijk: Health Council of the Netherlands; publication no. 1996/12, 1996.

7 Werkgroep van Deskundingen. *Gezondheidskundige aspecten van het begrip blootstelling en van het meten/schatten ervan*. Advies van de Werkgroep van Deskundigen ter vaststelling van MAC-waarden, autheur R.L. Zielhuis, Voorburg, Nederlands, Directoraat-Generaal van de Arbeid van het Ministerie van Sociale Zaken en Werkgelegenheid, rapportnummer RA 8/90, 1990.

8 Agence nationale de sécurité sanitaire de l'alimentation dleedt. *Expert appraisal on recommending occupational exposure limits for chemical agents. Reference document for the derivation and the measurement of exposure limit values for chemical agents in the workplace (OELs)*. Expert Committee (OEL Committee), Agence nationale de sécurité sanitaire de l'alimentation, de l' environnement et du travail (Anses); Request No. 2009-SA-0339, 2014; OEL permanent mission Request n°2009-SA-0339.

9 European Chemicals Agency. *Guidance on information requirements and chemical safety assessment Chapter R.4: Evaluation of available information*. Helsinki, Finland: European Chemicals Agency (ECHA) Reference: ECHA-2011-G-13-EN, 2011.

10 European Chemicals Agency. *Scientific aspects and methodologies related to the exposure of chemicals at the workplace*. Helsinki, Finland: Joint task force: Committee for Risk Assessment (RAC) and Scientific Committee on Occupational Exposure Limits (SCOEL); European Commission and European Chemicals Agency, 2017.

11 European Chemicals Agency. *Guidance on information requirements and chemical safety assessment. Appendix to Chapter R.8: Guidance for preparing a scientific report for health-based exposure limits at the workplace*. Helsinki, Finland: European Chemicals Agency (ECHA): Reference: ECHA-19-G-05-EN, 2019.

[12] International Agency for Research on Cancer. *IARC Monographs on the identification of carcinogenic hazards to humans. Preamble*. Lyon, France: IARC, World Health Organization, 2019.

[13] Scientific Committee on Occupational Exposure Limits. *Methodology for derivation of occupational exposure limits of chemical agents. The General Decision-Making Framework of the Scientific Committee on Occupational Exposure Limits (SCOEL).* Brussels, Belgium: European Commission, Directorate-General for Employment, Social Affairs and Inclusion, 2017; ISBN: 978-92-79-85304-3; doi: 10.2767/435199.

[14] Hemmerich J, Ecker GF. *In silico toxicology: From structure–activity relationships towards deep learning and adverse outcome pathways*. WIREs Computational Molecular Science 2020; 10(e1475): 1-23.

[15] Raies AB, Bajic VB. *In silico toxicology: computational methods for the prediction of chemical toxicity*. Wiley Interdiscip Rev Comput Mol Sci 2016; 6(2): 147-172.

[16] Zhu H. *From QSAR to QSIIR: searching for enhanced computational toxicology models*. Methods Mol Biol 2013; 930: 53-65.

[17] Ciallella HL, Zhu H. *Advancing Computational Toxicology in the Big Data Era by Artificial Intelligence: Data-Driven and Mechanism-Driven Modeling for Chemical Toxicity*. Chem Res Toxicol 2019; 32(4): 536-547.

[18] Gopinath C, Mowat V. *Toxicologic Pathology Forum*: Opinion on Designation of Adverse and Nonadverse Histopathological Findings in Toxicity Studies: The Pathologist's Dilemma*. Toxicologic Pathology 2019; 47(5): 564-573.

[19] Lewis RW, Billington R, Debryune E, Gamer A, Lang B, Carpanini F. *Recognition of Adverse and Nonadverse Effects in Toxicity Studies*. Toxicologic Pathology 2002; 30(1): 66-74.

[20] Peters A, Nawrot TS, Baccarelli AA. *Hallmarks of environmental insults*. Cell 2021; 184(6): 1455-1468.

[21] European Centre for Ecotoxicology and Toxicology of Chemicals (Ecetoc). *Recognition of, and differentiation between, adverse and non-adverse effects in toxicology studies*. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium; Technical Report No. 85, 2002.

[22] Karbe E, Williams GM, Lewis RW, Kimber I, Foster PM. *Distinguishing between adverse and non-adverse effects. Session summary*. Exp Toxicol Pathol 2002; 54(1): 51-55.

[23] Pandiri AR, Kerlin RL, Mann PC, Everds NE, Sharma AK, Myers LP, et al. *Is It Adverse, Nonadverse, Adaptive, or Artifact?* Toxicol Pathol 2017; 45(1): 238-247.

[24] European Food Safety Authority, Hardy A, Benford D, Halldorsson T, Jeger M, Knutsen HK, et al. *Clarification of some aspects related to genotoxicity assessment*. European Food Safety Authority (EFSA), European Union; EFSA Journal ISSN: 1831-4732; Volume 15(12):5113, 2017; 1831-4732. https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/j.efsa.2017.5113.

[25] Carlson MD, Morrison RS. *Study design, precision, and validity in observational studies*. J Palliat Med 2009; 12(1): 77-82.

[26] Money CD, Tomenson JA, Penman MG, Boogaard PJ, Jeffrey Lewis R. *A systematic approach for evaluating and scoring human data*. Regul Toxicol Pharmacol 2013; 66(2): 241-247.

[27] Swaen GM. *A framework for using epidemiological data for risk assessment*. Hum Exp Toxicol 2006; 25(3): 147-155.

[28] Vlaanderen J, Vermeulen R, Heederik D, Kromhout H, Ecnis Integrated Risk Assessment Group EUNOE. *Guidelines to evaluate human observational studies for quantitative risk assessment*. Environ Health Perspect 2008; 116(12): 1700-1705.

[29] Rothman KJ, Greenland S. *Causation and causal inference in epidemiology*. Am J Public Health 2005; 95 Suppl 1: S144-150.

[30] Kirkeleit J, Riise T, Bjorge T, Christiani DC. *The healthy worker effect in cancer incidence studies*. Am J Epidemiol 2013; 177(11): 1218-1224.

[31] Goldbohm RA, Tielemans EL, Heederik D, Rubingh CM, Dekkers S, Willems MI, et al. *Risk estimation for carcinogens based on epidemiological data: a structured approach, illustrated by an example on chromium*. Regul Toxicol Pharmacol 2006; 44(3): 294-310.

[32] World Health Organization. *Evaluation and use of epidemiological evidence for environmental health risk assessment: WHO guideline document*. Environ Health Perspect 2000; 108(10): 997-1002.

[33] Everdingen van JJE. *Van consensus naar CBO-richtlijn*. Nederlands Tijdschrift voor Geneeskunde 1999; 143(42): 2086-2089.

[34] Guyatt GHO, A.D.; Schünemann, H.J.; Tugwell, P.; Knottnerus, A. *GRADE-guidelines: a new series of articles on the Journal of Clinical Epidemiology*. Journal of Clinical Epidemiology 2011; 64: 380-382.

[35] National Toxicology Program. *Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration*. U.S. Department of Health and Human Services; 2015.

[36] von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. *[The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies]*. Rev Esp Salud Publica 2008; 82(3): 251-259.

[37] Steenland K, Schubauer-Berigan MK, Vermeulen R, Lunn RM, Straif K, Zahm S, et al. *Risk of Bias Assessments and Evidence Syntheses for Observational Epidemiologic Studies of Environmental and Occupational Exposures: Strengths and Limitations*. Environ Health Perspect 2020; 128(9): 95002.

[38] Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T, Friedenreich C. *Traditional reviews, meta-analyses and pooled analyses in epidemiology*. Int J Epidemiol 1999; 28(1): 1-9.

[39] Klimisch HJ, Andreae M, Tillmann U. *A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data*. Regul Toxicol Pharmacol 1997; 25(1): 1-5.

[40] Borgert CJ, Becker RA, Carlton BD, Hanson M, Kwiatkowski PL, Sue Marty M, et al. *Does GLP enhance the quality of toxicological evidence for regulatory decisions?* Toxicol Sci 2016; 151(2): 206-213.

[41] Ingre-Khans E, Agerstrand M, Beronius A, Ruden C. *Reliability and relevance evaluations of REACH data*. Toxicol Res (Camb) 2019; 8(1): 46-56.

[42] Ingre-Khans E, Agerstrand M, Beronius A, Ruden C. *Toxicity studies used in registration, evaluation, authorisation and restriction of chemicals (REACH): How accurately are they reported?* Integr Environ Assess Manag 2019; 15(3): 458-469.

[43] Zoeller RT, Vandenberg LN. *Assessing dose-response relationships for endocrine disrupting chemicals (EDCs): a focus on non-monotonicity*. Environ Health 2015; 14: 42.

[44] Schneider K, Schwarz M, Burkholder I, Kopp-Schneider A, Edler L, Kinsner-Ovaskainen A, et al. *"ToxRTool", a new tool to assess the reliability of toxicological data*. Toxicology Letters 2009; 189(2): 138-144.

[45] Organisation for Economic Co-operation and Development. *Guidance Document on Good In Vitro Method Practices (GIVIMP)*. OECD, Paris, France, https://www.oecd.org/, 2018. https://www.oecd-ilibrary.org/content/publication/9789264304796-en.

[46] Krebs A, Waldmann T, Wilks MF, Van Vugt-Lussenburg BMA, Van der Burg B, Terron A, et al. *Template for the description of cell-based toxicological test methods to allow evaluation and regulatory use of the data*. ALTEX 2019; 36(4): 682-699.

[47] Hill AB. *The Environment and Disease: Association or Causation?* Proc R Soc Med 1965; 58: 295-300.

[48] Morgan RL, Thayer KA, Bero L, Bruce N, Falck-Ytter Y, Ghersi D, et al. *GRADE: Assessing the quality of evidence in environmental and occupational health*. Environ Int 2016; 92-93: 611-616.

[49] Savitz DA, Wellenius GA, Trikalinos TA. *The Problem With Mechanistic Risk of Bias Assessments in Evidence Synthesis of Observational Studies and a Practical Alternative: Assessing the Impact of Specific Sources of Potential Bias*. Am J Epidemiol 2019; 188(9): 1581-1585.

[50] Environmental Protection Agency. *Preamble to the Integrated Science Assessments*. Washington D.C.: U.S. EPA, www.epa.gov/isa, Report EPA/600/R-15/067, 2015.

[51] Organisation for Economic Co-operation and Development. *OECD test guidelines for testing of chemicals: Introduction to the OECD guidelines on genetic toxicology testing*. OECD, Paris, France, https://www.oecd.org/; Draft, 2014.

[52] Organisation for Economic Co-operation and Development. *Guidance document on revisions to OECD genetic toxicology test guidelines*. OECD, Paris, France, https://www.oecd.org/, 2015.

[53] Organisation for Economic Co-operation and Development. *Overview of the set of OECD genetic toxicology test guidelines and updats*

performed in 2014-2015. OECD, Paris, France https://www.oecd.org/; Series on Testing & Assessment, No. 238 - 2nd edition, 2017.

54 European Commission. *Regulation (EC) No 1272/2008 of the European Parliamant and of the Council on classification, labelling and packaging of substances and mixtures*. 2009.

55 Davis JA, Gift JS, Zhao QJ. *Introduction to benchmark dose methods and U.S. EPA's benchmark dose software (BMDS) version 2.1.1*. Toxicol Appl Pharmacol 2011; 254(2): 181-191.

56 Environmental Protection Agency. *Benchmark dose software (BMDS), user manual*. United States Environmental Protection Agency; https://www.epa.goc/bmds; last visited March 2021, 2016.

57 European Food Safety Authority. *Update: use of the benchmark dose approach in risk assessment*. European Food Safety Authority (EFSA), European Union; EFSA Journal ISSN: 1831-4732; Volume 15 (1): 4658, 2017.

58 Haber LT, Dourson ML, Allen BC, Hertzberg RC, Parker A, Vincent MJ, et al. *Benchmark dose (BMD) modeling: current practice, issues, and challenges*. Crit Rev Toxicol 2018; 48(5): 387-415.

59 Health Council of the Netherlands. *Benchmark-dose method: health-based recommended exposure limits in new persective*. The Hague, the Netherlands: Health Council of the Netherlands; publication no. 2003/06, 2003.

60 European Centre for Ecotoxicology and Toxicology of Chemicals (Ecetoc). *Derivation of assessment factors for human health risk assessment*. European Centre for Ecotoxicology and Toxicology of Chemicals (Ecetoc), Brussels, Belgium**:** Technical Report No. 86, 2003.

61 European Chemicals Agency. *Guidance on information requirements and chemical safety assessment. Chapter R.8: Characterisation of dose [concentration]-response for human health*. European Chemicals Agency (ECHA), http://echa.europa.eu; Reference: ECHA-2010-G-19-EN, 2012.

62 Ferderal Insititute for Occupational Safety and Health (BAuA). *Synthesis report: modelling of distributions of assessment factors, comparison woth current methods and discussion of protection goals*. Freiburg, Germany: Reaerch project F2437: Derivation of occupational exposure limits for airborne chemicals - comparison of methods and protection levels. BAuA; research project F2437; prepared by K. Schneider, M. Dilger and E. Kaiser, 2020.

63 Bokkers BG, Slob W. *Deriving a data-based interspecies assessment factor using the NOAEL and the benchmark dose approach*. Crit Rev Toxicol 2007; 37(5): 355-373.

64 Gevel van de IA, Hakkert BC. *Methods for interspecies extrapolation: an evaluation*. Zeist, the Netherlands: TNO Nutrition and Food Research Institute; TNO-report V97.529, 1998.

65 International Programme on Chemical Safety. *Chemical-specific adjustment factors for interspecies differences and human variability: guidance document for use of data in dose/concentration-response*

assessment. Geneva, Switzerland: IPCS, World Health Organization: Harmonization project document no. 2, 2005.

[66] Paulussen JJC, Mahieu CM, Bos PMJ. *Default values in occupational risk assessment*. Zeist, the Netherlands: TNO Nutrition and Food Research Institute; TNO report V98.390, 1998.

[67] Environmental Protection Agency. *Reference physiological parameters in pharmacokinetic modeling*. Washington DC,: United States Environmental Protection Agency, report no. 600/6-88/004, 1988.

[68] European Centre for Ecotoxicology and Toxicology of Chemicals (Ecetoc). *Strategy for assessing a "skin notation"*. European Centre for Ecotoxicology and Toxicology of Chemicals (Ecetoc), Brussels, Belgium; Document No. 31 (revised), 1993.

[69] European Food Safety Authority. *Scientific opinion: Guidance on dermal absorption*. Parma, Italy: EFSA Panel on plant products and their residues (PPR); European Food Safety Authority, European Union; EFSA Journal 2012; 10(4): 2665, 2017.

[70] European Food Safety Authority. *Guidance on dermal absorption*. EFSA Journal 2017; 15(6): 4873.

[71] National Institute for Occupational Safety and Health. *A strategy for assigning new NIOSH skin notations*. USA: Current Intelligence Bulletin 61; NIOSH, Centers for Disease Control and Prevention, Dept. of Health and Human Services: DHHS (NIOSH) publication no. 2009-147, 2009.

[72] Eick van AJE, R.M. *Criteria voor toekenning "H"-indicatie aan stoffen in de MAC lijst [in Dutch]*. Rijswijk: Medisch Biologisch Laboratorium, TNO, 1989.

Abbreviations

ACGIH	American Conference of Governmental Industrial Hygienists
AGS	German Ausschuss für Gefahrstoffe
ANSES	French National Agency for Food, Environmental and Occupational Health & Safety
AOP	Adverse Outcome Pathway
ATSDR	Agency for Toxic Substances and Disease Registry
BAuA	German Bundesanstalt für Arbeitsschutz und Arbeitsmedizin
BLV	Biological Limit Value
BMD	Benchmark Dose
BMR	Benchmark Response
CBO	Dutch Institute for Healthcare Improvement
CLP	Classification, Labelling and Packaging
BLV	Biological Limit Value
CICAD	Concise International Chemical Assessment documents
CMR	Carcinogens, Mutagens and Reproduction toxic substances
DECOS	Dutch Expert Committee on Occupational Safety
DFG	Deutsche Forschungsgemeinschaft
ECETOC	European Centre for Toxicology and Toxicology of Chemicals
ECHA	European Chemicals Agency
ECVAM	European Centre for the Validation of Alternative Methods
EFSA	European Food Safety Authority
EPA	U.S. Environmental Protection Agency
GRADE	Grading of Recommendations Assessment, Development and Evaluation
IPD	Individual Patient Data
IARC	International Agency for Cancer Research
ISO	International Standard Organization
NEG	Nordic Expert Group
NIOSH	U.S. National Institute for Occupational Safety and Health
NOAEL	No Observed Adverse Effect Level
LOAEL	Lowest Observed Adverse Effect Level
OECD	Organisation for Economic Co-operation and Development
OEL	Occupational Exposure Limit
OSHA	Occupational Safety and Health Administration
PBPK	Physiologically Based Pharmacokinetic modelling
QSAR	Quantitative Structure-Activity Relationship
QVIVE	Quantitative *in vitro* to *in vivo* extrapolation
RAC	Committee for Risk Assessment
RCT	Randomized Controlled Trial
REACH	Registration, Evaluation, Authorisation and Restriction of Chemicals
RoB	Risk of bias
SCOEL	Scientific Committee on Occupational Exposure Limits
STEL	Short Term Exposure Limit
TWA	Time Weighted Average
WHO	World Health Organization

The Health Council of the Netherlands, established in 1902, is an independent scientific advisory body. Its remit is “to advise the government and Parliament on the current level of knowledge with respect to public health issues and health (services) research...” (Section 22, Health Act).
The Health Council receives most requests for advice from the Ministers of Health, Welfare and Sport, Infrastructure and Water Management, Social Affairs and Employment, and Agriculture, Nature and Food Quality. The Council can publish advisory reports on its own initiative. It usually does this in order to ask attention for developments or trends that are thought to be relevant to government policy.
Most Health Council reports are prepared by multidisciplinary committees of Dutch or, sometimes, foreign experts, appointed in a personal capacity.
The reports are available to the public.

This publication can be downloaded from www.healthcouncil.nl.

Preferred citation:
Health Council of the Netherlands. Guidance for recommending classifications and health-based occupational exposure limits.
The Hague: Health Council of the Netherlands, 2021.

All rights reserved