International Agency for Research on Cancer

World Health Organization

IARC Monographs on the Identification of Carcinogenic Hazards to Humans

PREAMBLE

Lyon, France
Amended January 2019

Contents

The Preamble to the *IARC Monographs* describes the objective and scope of the programme, general principles and procedures, and scientific review and evaluations. The *IARC Monographs* embody principles of scientific rigour, impartial evaluation, transparency, and consistency. The Preamble should be consulted when reading a *Monograph* or a summary of a *Monograph*'s evaluations. Separate Instructions for Authors describe the operational procedures for the preparation and publication of a volume of the *Monographs*.

A. GENERAL PRINCIPLES AND PROCEDURES

1. Background

Soon after the International Agency for Research on Cancer (IARC) was established in 1965, it started to receive frequent requests for advice on the carcinogenicity of chemicals, including requests for lists of established and suspected human carcinogens. In 1970, an IARC Advisory Committee on Environmental Carcinogenesis recommended "that a compendium on carcinogenic chemicals be prepared by experts. The biological activity and evaluation of practical importance to public health should be referenced and documented." The next year, the IARC Governing Council adopted a resolution that IARC should prepare "monographs on the evaluation of carcinogenic risk of chemicals to man", which became the initial title of the series.

In succeeding years, the scope of the programme broadened as *Monographs* were developed for complex mixtures, occupational exposures, physical agents, biological organisms, pharmaceuticals, and other exposures. In 1988, "of chemicals" was dropped from the title, and in 2019, "evaluation of carcinogenic risks" became "identification of carcinogenic hazards", in line with the objective of the programme.

Identifying the causes of human cancer is the first step in cancer prevention. The identification of a cancer hazard may have broad and profound implications. National and international authorities and organizations can and do use information on causes of cancer in support of actions to reduce exposure to carcinogens in the workplace, in the environment, and elsewhere. Cancer prevention is needed as much today as it was when IARC was established, because the global burden of cancer is high and continues to increase as a result of population growth and ageing and upward trends in some exposures, especially in low- and middle-income countries (http://publications.iarc.fr/Non-Series-Publications/World-Cancer-Reports).

IARC's process for developing *Monographs*, which has evolved over several decades, involves the engagement of international, interdisciplinary Working Groups of expert scientists, the transparent synthesis of different streams of evidence (exposure characterization, cancer in humans, cancer in experimental animals, and mechanisms of carcinogenesis), and the integration of these streams of evidence into an overall evaluation and classification according to criteria developed and refined by IARC. Since the *Monographs* programme was established, the understanding of carcinogenesis has greatly deepened. Scientific advances are incorporated into the evaluation methodology. In particular, strong mechanistic evidence has had an increasing role in the overall evaluations since 1991.

The Preamble is primarily a statement of the general principles and procedures used in developing a *Monograph*, to promote transparency and consistency across *Monographs* evaluations. In addition, IARC provides Instructions for Authors (https://monographs.iarc.fr/instructions-for-authors/), which specify more detailed working procedures. IARC routinely updates these Instructions for Authors to reflect advances in methods for cancer hazard identification and accumulated experience, including input from experts.

2. Objective and scope

The objective of the programme is to prepare, with the engagement of international, interdisciplinary Working Groups of experts, scientific reviews and evaluations of evidence on the carcinogenicity of a wide range of agents.

The *Monographs* assess the strength of the available evidence that an agent can cause cancer in humans, based on three streams of evidence: on cancer in humans (see Part B, Section 2), on cancer in experimental animals (see Part B, Section 3), and on mechanistic evidence (see Part B, Section 4). In addition, the exposure to each agent is characterized (see Part B, Section 1). In this Preamble, the term "agent" refers to any chemical, physical, or biological entity or exposure circumstance (e.g. occupation as a painter) for which evidence on the carcinogenicity is evaluated.

A cancer *hazard* is an agent that is capable of causing cancer, whereas a cancer *risk* is an estimate of the probability that cancer will occur given some level of exposure to a cancer hazard. The *Monographs* assess the strength of evidence that an agent is a cancer hazard. The distinction between hazard and risk is fundamental. The *Monographs* identify cancer hazards even when risks appear to be low in some exposure scenarios. This is because the exposure may be widespread at low levels, and because exposure levels in many populations are not known or documented.

Although the *Monographs* programme has focused on hazard identification, some epidemiological studies used to identify a cancer hazard are also used to estimate an exposure–response relationship within the range of the available data. However, extrapolating exposure–response relationships beyond the available data (e.g. to lower exposures, or from experimental animals to humans) is outside the scope of *Monographs* Working Groups (IARC, 2014). In addition, the *Monographs* programme does not review quantitative risk characterizations developed by other health agencies.

The identification of a cancer hazard should trigger some action to protect public health, either directly as a result of the hazard identification or through the conduct of a risk assessment. Although such actions are outside the scope of the programme, the *Monographs* are used by national and international authorities and organizations to inform risk assessments, formulate decisions about preventive measures, motivate effective cancer control programmes, and choose among options for public health decisions. *Monographs* evaluations are only one part of the body of information on which decisions to control exposure to carcinogens may be based. Options to prevent cancer vary from one situation to another and across geographical regions and take many factors into account, including different national priorities. Therefore, no recommendations are given in the *Monographs* with regard to regulation, legislation, or other policy approaches, which are the responsibility of individual governments or organizations. The *Monographs* programme also does not make research recommendations. However, it is important to note that *Monographs* contribute significantly to the science of carcinogenesis by synthesizing and integrating streams of evidence about carcinogenicity and pointing to critical gaps in knowledge.

3. Selection of agents for review

Since 1984, about every five years IARC convenes an international, interdisciplinary Advisory Group to recommend agents for review by the *Monographs* programme. IARC selects Advisory Group members who are knowledgeable about current research on carcinogens and public health priorities. Before an Advisory Group meets, IARC solicits nominations of agents from scientists and government agencies worldwide. Since 2003, IARC also invites nominations from the public. IARC charges each Advisory Group with reviewing nominations, evaluating exposure and hazard potential, and preparing a report that documents the Advisory Group's process for these activities and its rationale for the recommendations.

For each new volume of the *Monographs,* IARC selects the agents for review from those recommended by the most recent Advisory Group, considering the availability of pertinent research studies and current public health priorities. On occasion, IARC may select other agents if there is a need to rapidly evaluate an emerging carcinogenic hazard or an urgent need to re-evaluate a previous classification. All evaluations consider the full body of available evidence, not just information published after a previous review.

A *Monograph* may review:

(a) An agent not reviewed in a previous *Monograph*, if there is potential human exposure and there is evidence for assessing its carcinogenicity. A group of related agents (e.g. metal compounds) may be reviewed together if there is evidence for assessing carcinogenicity for one or more members of the group.

(b) An agent reviewed in a previous *Monograph*, if there is new evidence of cancer in humans or in experimental animals, or mechanistic evidence to warrant re-evaluation of the classification. In the interests of efficiency, the literature searches may build on previous comprehensive searches.

(c) An agent that has been established to be carcinogenic to humans and has been reviewed in a previous *Monograph*, if there is new evidence of cancer in humans that indicates new tumour sites where there might be a causal association. In the interests of efficiency, the review may focus on these new tumour sites.

4. The Working Group and other meeting participants

Five categories of participants can be present at *Monographs* meetings:

(i) *Working Group* members are responsible for all scientific reviews and evaluations developed in the volume of the *Monographs*. The Working Group is interdisciplinary and comprises subgroups of experts in the fields of (a) exposure characterization, (b) cancer in humans, (c) cancer in experimental animals, and (d) mechanistic evidence. IARC selects Working Group members on the basis of expertise related to the subject matter and relevant methodologies, and absence of conflicts of interest. Consideration is also given to diversity in scientific approaches and views, as well as demographic composition. Working Group members generally have published research related to the exposure or carcinogenicity of the agents being reviewed, and IARC uses literature searches to identify most experts. Since 2006, IARC also has encouraged public nominations through its Call for Experts. IARC's reliance on experts with knowledge of the subject matter and/or expertise in methodological assessment is confirmed by decades of experience documenting that there is value in specialized expertise and that the overwhelming majority of Working Group members are committed to the objective evaluation of scientific evidence and not to the narrow advancement of their own research results or a pre-determined outcome (Wild & Cogliano, 2011). Working Group members are expected to serve the public health mission of IARC, and should refrain from consulting and other activities for financial gain that are related to the agents under review, or the use of inside information from the meeting, until the full volume of the *Monographs* is published.

IARC identifies, from among Working Group members, individuals to serve as Meeting Chair and Subgroup Chairs. At the opening of the meeting, the Working Group is asked to endorse the selection of the Meeting Chair, with the opportunity to propose alternatives. The Meeting Chair and Subgroup Chairs take a leading role at all stages of the review process (see Part A, Section 7), promote open scientific discussions that involve all Working

Group members in accordance with normal committee procedures, and ensure adherence to the Preamble.

(ii) *Invited Specialists* are experts who have critical knowledge and experience but who also have a conflict of interest that warrants exclusion from developing or influencing the evaluations of carcinogenicity. Invited Specialists do not draft any section of the *Monograph* that pertains to the description or interpretation of cancer data, and they do not participate in the evaluations. These experts are invited in limited numbers when necessary to assist the Working Group by contributing their unique knowledge and experience to the discussions.

(iii) *Representatives of national and international health agencies* may attend because their agencies are interested in the subject of the meeting. They do not draft any section of the *Monograph* or participate in the evaluations.

(iv) *Observers* with relevant scientific credentials may be admitted in limited numbers. Attention is given to the balance of Observers from constituencies with differing perspectives. Observers are invited to observe the meeting and should not attempt to influence it, and they agree to respect the Guidelines for Observers at *IARC Monographs* meetings. Observers do not draft any section of the *Monograph* or participate in the evaluations.

(v) The *IARC Secretariat* consists of scientists who are designated by IARC and who have relevant expertise. The IARC Secretariat coordinates and facilitates all aspects of the evaluation and ensures adherence to the Preamble throughout development of the scientific reviews and classifications (see Part A, Sections 5 and 6). The IARC Secretariat organizes and announces the meeting, identifies and recruits the Working Group members, and assesses the declared interests of all meeting participants. The IARC Secretariat supports the activities of the Working Group (see Part A, Section 7) by searching the literature and performing title and abstract screening, organizing conference calls to coordinate the development of pre-meeting drafts and discuss cross-cutting issues, and reviewing drafts before and during the meeting. Members of the IARC Secretariat serve as meeting rapporteurs, assist the Meeting Chair and Subgroup Chairs in facilitating all discussions, and may draft text or tables when designated by the Meeting Chair and Subgroup Chairs. Their participation in the evaluations is restricted to the role of clarifying or interpreting the Preamble.

All participants are listed, with their principal affiliations, in the front matter of the published volume of the *Monographs*. Working Group members and Invited Specialists serve as individual scientists and not as representatives of any organization, government, or industry (Cogliano et al., 2004).

The roles of the meeting participants are summarized in Table 1.

Table 1. Roles of participants at *IARC Monographs* meetings

	Role			
Category of participant	**Prepare text, tables, and analyses**	**Participate in discussions**	**Participate in evaluations**	**Eligible to serve as Chair**
Working Group members	√	√	√	√
Invited Specialists	√[a]	√		
Representatives of health agencies		√[b]		
Observers		√[b]		
IARC Secretariat	√[c]	√	√[d]	

[a] Only for the section on exposure characterization

[b] Only at times designated by the Meeting Chair and Subgroup Chairs

[c] When needed or requested by the Meeting Chair and Subgroup Chairs

[d] Only for clarifying or interpreting the Preamble

5. Working procedures

A separate Working Group is responsible for developing each volume of the *Monographs*. A volume contains one or more *Monographs*, which can cover either a single agent or several related agents. Approximately one year before the meeting of a Working Group, a preliminary list of agents to be reviewed, together with a Call for Data and a Call for Experts, is announced on the *Monographs* programme website (http://monographs.iarc.fr).

Before a meeting invitation is extended, each potential participant, including the IARC Secretariat, completes the WHO Declaration of Interests form to report financial interests, employment and consulting (including remuneration for serving as an expert witness), individual and institutional research support, and non-financial interests such as public statements and positions related to the subject of the meeting. IARC assesses the declared interests to determine whether there is a conflict that warrants any limitation on participation (see Table 2).

Approximately two months before a *Monographs* meeting, IARC publishes the names and affiliations of all meeting participants together with a summary of declared interests, in the interests of transparency and to provide an opportunity for undeclared conflicts of interest to be brought to IARC's attention. It is not acceptable for Observers or third parties to contact other participants before a meeting or to lobby them at any time. Meeting participants are asked to report all such contacts to IARC (Cogliano et al., 2005).

The Working Group meets at IARC for approximately eight days to discuss and finalize the scientific review and to develop summaries and evaluations. At the opening of the meeting, all participants update their Declaration of Interests forms, which are then reviewed by IARC. Declared interests related to the subject of the meeting are disclosed to the meeting participants during the meeting and in the published volume (Cogliano et al., 2004). The objectives of the meeting are peer review and consensus. During the first part of the meeting, subgroup sessions (covering exposure characterization, cancer in humans, cancer in experimental animals, and mechanistic evidence) review the pre-meeting drafts, develop a joint subgroup draft, and draft subgroup summaries. During the last part of the meeting, the Working Group meets in plenary session to review the subgroup drafts and summaries and to develop the consensus evaluations. As a result, the entire volume is the joint product of the Working Group, and there are no individually authored sections. After the meeting, the master copy is verified by the IARC Secretariat and is then edited and prepared for publication. The aim is to publish the volume within approximately nine months of the Working Group meeting. A summary of the evaluations and key supporting evidence is prepared for publication in a scientific journal or is made available on the *Monographs* programme website soon after the meeting.

In the interests of transparency, IARC engages with the public throughout the process, as summarized in Table 2.

Table 2. Public engagement during *Monographs* development

Approximate timeframe	**Engagement**
Every 5 years	IARC convenes an Advisory Group to recommend high-priority agents for future review
~1 year before a *Monographs* meeting	IARC selects agents for review in a new volume of the *Monographs* IARC posts on its website: Preliminary List of Agents to be reviewed Call for Data and Call for Experts Request for Observer Status WHO Declaration of Interests form
~8 months before a *Monographs* meeting	Call for Experts closes
~4 months before a *Monographs* meeting	Request for Observer Status closes
~2 months before a *Monographs* meeting	IARC posts the names of all meeting participants together with a summary of declared interests, and a statement discouraging contact of the Working Group by interested parties
~1 month before a *Monographs* meeting	Call for Data closes
~2–4 weeks after a *Monographs* meeting	IARC publishes a summary of evaluations and key supporting evidence
~9 months after a *Monographs* meeting	IARC Secretariat publishes the verified and edited master copy of plenary drafts as a *Monographs* volume

6. Overview of the scientific review and evaluation process

The Working Group considers all pertinent epidemiological studies, cancer bioassays in experimental animals, and mechanistic evidence, as well as pertinent information on exposure in humans. In general, for cancer in humans, cancer in experimental animals, and mechanistic evidence, only studies that have been published or accepted for publication in the openly available scientific literature are reviewed. Under some circumstances, materials that are publicly available and whose content is final may be reviewed if there is sufficient information to permit an evaluation of the quality of the methods and results of the studies (see Step 1, below). Such materials may include reports and databases publicly available from government agencies, as well as doctoral theses. The reliance on published and publicly available studies promotes transparency and protects against citation of premature information.

The principles of systematic review are applied to the identification, screening, synthesis, and evaluation of the evidence related to cancer in humans, cancer in experimental animals, and mechanistic evidence (as described in Part B, Sections 2–4 and as detailed in the Instructions for Authors). Each *Monograph* specifies or references information on the conduct of the literature searches, including search terms and inclusion/exclusion criteria that were used for each stream of evidence.

In brief, the steps of the review process are as follows:

Step 1. Comprehensive and transparent identification of the relevant information: The IARC Secretariat identifies relevant studies through initial comprehensive searches of literature contained in authoritative biomedical databases (e.g. PubMed, PubChem) and through a Call for Data. These literature searches, designed in consultation with a librarian and other technical experts, address whether the agent causes cancer in humans, causes cancer in experimental systems, and/or exhibits key characteristics of established human carcinogens (in humans or in experimental systems). The Working Group provides input and advice to IARC to refine the search strategies, and identifies literature through other searches (e.g. from reference lists of past *Monographs*, retrieved articles, and other authoritative reviews).

For certain types of agents (e.g. regulated pesticides and pharmaceuticals), IARC also provides an opportunity to relevant regulatory authorities, and regulated parties through such authorities, to make pertinent unpublished studies publicly available by the date specified in the Call for Data. Consideration of such studies by the Working Group is dependent on the public availability of sufficient information to permit an independent evaluation of (a) whether there has been selective reporting (e.g. on outcomes, or from a larger set of conducted studies), (b) study quality (e.g. design, methodology, and reporting of results), and (c) study results.

Step 2. Screening, selection, and organization of the studies: The IARC Secretariat screens the retrieved literature for inclusion based on title and abstract review, according to pre-defined exclusion criteria. For instance, studies may be excluded if they were not about the agent (or a metabolite of the agent), or if they reported no original data on epidemiological or toxicological end-points (e.g. review articles). The Working Group reviews the title and abstract screening done by IARC, and performs full-text review. Any reasons for exclusion are recorded, and included studies are organized according to factors pertinent to the considerations described in Part B, Sections 2–4 (e.g. design, species, and end-point). Inclusion of a study does not imply acceptance of the adequacy of the study design or of the analysis and interpretation of the results.

Step 3. Evaluation of study quality: The Working Group evaluates the quality of the included studies based on the considerations (e.g. design, methodology, and reporting of results) described in Part B, Sections 2–4. Based on these considerations, the Working Group may accord greater weight to some of the included studies. Interpretation of the results and the strengths and limitations of a study are clearly outlined in square brackets at the end of study descriptions (see Part B).

Step 4. Report characteristics of included studies, including assessment of study quality: Pertinent characteristics and results of included studies are reviewed and succinctly described, as detailed in Part B, Sections 1–4. Tabulation of data may facilitate this reporting. This step may be iterative with Step 3.

Step 5. Synthesis and evaluation of strength of evidence: The Working Group summarizes the overall strengths and limitations of the evidence from the individual streams of evidence (cancer in humans, cancer in experimental animals, and mechanistic evidence; see Part B, Section 5). The Working Group then evaluates the strength of evidence from each stream of evidence by using the transparent methods and defined descriptive terms given in Part B, Sections 6a–c. The Working Group then develops, and describes the rationale for, the consensus classification of carcinogenicity that integrates the conclusions about the strength of evidence from studies of cancer in humans, studies of cancer in experimental animals, and mechanistic evidence (see Part B, Section 6d).

7. Responsibilities of the Working Group

The Working Group is responsible for identifying and evaluating the relevant studies and developing the scientific reviews and evaluations for a volume of the *Monographs*. The IARC Secretariat supports these activities of the Working Group (see Part A, Section 4). Briefly, the Working Group's tasks in developing the evaluation are, in sequence:

(i) Before the meeting, the Working Group ascertains that all appropriate studies have been identified and selected, and assesses the methods and quality of each individual study, as outlined above (see Part A, Section 6). The Working Group members prepare pre-meeting working drafts that present accurate tabular or textual summaries of informative studies by extracting key elements of the study design and results, and highlighting notable strengths and limitations. They participate in conference calls organized by IARC to coordinate the development of working drafts and to discuss cross-cutting issues. Pre-meeting reviews of all working drafts are generally performed by two or more subgroup members who did not participate in study identification, data extraction, or study review for the draft. Each study summary is written or reviewed by someone who is not associated with the study.

(ii) At the meeting, within subgroups, the Working Group members critically review, discuss, and revise the pre-meeting drafts and adopt the revised versions as consensus subgroup drafts. Subgroup Chairs ensure that someone who is not associated with the study leads the discussion of each study summary. A proposed classification of the strength of the evidence reviewed in the subgroup using the *IARC Monographs* criteria (see Part B, Sections 6a–c) is then developed from the consensus subgroup drafts of the evidence summaries (see Part B, Section 5).

(iii) During the plenary session, each subgroup presents its drafts for scientific review and discussion to the other Working Group members, who did not participate in study identification, data extraction, or study review for the drafts. Subgroup Chairs ensure that someone who is not associated with the study leads the discussion of each study summary. After review, discussion, and revisions as needed, the subgroup drafts are adopted as a consensus Working Group product. The summaries and classifications of the strength of the evidence, developed in the subgroup in line with the *IARC Monographs* criteria (see Part B, Sections 6a–c), are considered, revised as needed, and adopted by the full Working Group. The Meeting Chair proposes an overall evaluation using the guidance provided in Part B, Section 6d.

The Working Group strives to achieve consensus evaluations. Consensus reflects broad agreement among the Working Group, but not necessarily unanimity. The Meeting Chair may poll the Working Group to determine the diversity of scientific opinion on issues where consensus is not apparent.

Only the final product of the plenary session represents the views and expert opinions of the Working Group. The entire *Monographs* volume is the joint product of the Working Group and represents an extensive and thorough peer review of the body of evidence (individual studies, synthesis, and evaluation) by an interdisciplinary expert group. Initial working papers and subsequent revisions are not released, because they would give an incomplete and possibly misleading impression of the consensus developed by the Working Group over a full week of deliberation.

B. SCIENTIFIC REVIEW AND EVALUATION

This part of the Preamble discusses the types of evidence that are considered and summarized in each section of a *Monograph*, followed by the scientific criteria that guide the evaluations. In addition, a section of General Remarks at the front of the volume discusses the reasons the agents were scheduled for evaluation and any key issues encountered during the meeting.

1. Exposure characterization

This section identifies the agent and describes its occurrence, main uses, and production locations and volumes, where relevant. It also summarizes the prevalence, concentrations in relevant studies, and relevant routes of exposure in humans worldwide. Methods of exposure measurement and analysis are described, and methods of exposure assessment used in key epidemiological studies reviewed by the Working Group are described and evaluated.

Over the course of the *Monographs* programme, concepts of exposure and dose have evolved substantially with deepening understanding of the interactions of agents and biological systems. The concept of exposure has broadened and become more holistic, extending beyond chemical, physical, and biological agents to stressors as construed generally, including psychosocial stressors (National Research Council, 2012; National Academies of Sciences, Engineering, and Medicine, 2017). Overall, this broader conceptualization supports greater integration between exposure characterization and other sections of the *Monographs*. Concepts of absorption, distribution, metabolism, and excretion are considered in the first subsection of mechanistic evidence (see Part B, Section 4a), whereas validated biomarkers of internal exposure or metabolites that are routinely used for exposure assessment are reported on in this section (see Part B, Section 1b).

(a) Identification of the agent

The agent being evaluated is unambiguously identified. Details will vary depending on the type of agent but will generally include physical and chemical properties relevant to the agent's identification, occurrence, and biological activity. If the material that has been tested in experimental animals or in vitro systems is different from that to which humans are exposed, these differences are noted.

For chemical agents, the Chemical Abstracts Service Registry Number is provided, as well as the latest primary name and other names in common use, including important trade names, along with available information on the composition of common mixtures or products containing the agent, and potentially toxic and/or carcinogenic impurities. Physical properties relevant to understanding the potential for human exposure and measures of exposure used in studies in humans are summarized. These might include physical state, volatility, aqueous and fat solubility, and half-life in the environment and/or in human tissues.

For biological agents, taxonomy and structure are described. Mode of replication, life-cycle, target cells, persistence, latency, and host responses, including morbidity and mortality through pathologies other than cancer, are also presented.

For foreign bodies, fibres and particles, composition, size range, relative dimensions, and accumulation, persistence, and clearance in target organs are summarized. Physical agents that are forms of radiation are described in terms of frequency spectrum and energy transmission.

Exposures may result from, or be influenced by, a diverse range of social and environmental factors, including components of diet, sleep, and physical activity patterns. In these instances, this section will include a description of the agent, its variability across human populations, and its composition or characteristics relevant to understanding its potential carcinogenic hazard to humans and to evaluating exposure assessments in epidemiological studies.

(b) Detection and analysis

Key methods of detection and quantification of the agent are presented, with an emphasis on those used most widely in surveillance, regulation, and epidemiological studies. Measurement methods for sample matrices that are deemed important sources of human exposure (e.g. air, drinking-water, food, residential dust) and for validated exposure biomarkers (e.g. the agent or its metabolites in human blood, urine, or saliva) are described. Information on detection and quantification limits is provided when it is available and is useful for interpreting studies in humans and in experimental animals. This is not an exhaustive treatise but is meant to help readers understand the strengths and limitations of the available exposure data and of the epidemiological studies that rely on these measurements.

(c) Production and use

Historical and geographical patterns and trends in production and use are included when they are available, to help readers understand the contexts in which exposures may occur, both within key epidemiological studies reviewed by the Working Group and in human populations generally. Industries that produce, use, or dispose of the agent are described, including their global distribution, when available. National or international listing as a high-production-volume chemical or similar classification may be included. Production processes with significant potential for occupational exposure or environmental pollution are indicated. Trends in global production volumes, technologies, and other data relevant to understanding exposure potential are summarized. Minor or historical uses with significant exposure potential or with particular relevance to key epidemiological studies are included. Particular effort may be directed towards finding data on production in low- and middle-income countries, where rapid economic development may lead to higher exposures than those in high-income countries.

(d) Exposure

A concise overview of quantitative information on sources, prevalence, and levels of exposure in humans is provided. Representative data from research studies, government reports and websites, online databases, and other citable, publicly available sources are tabulated. Data from low- and middle-income countries are sought and included to the extent feasible; information gaps for key regions are noted. Naturally occurring sources of exposure, if any, are noted. Primary exposure routes (e.g. inhalation, ingestion, skin uptake) and other considerations relevant to understanding the potential for cancer hazard from exposure to the agent are reported.

For occupational settings, information on exposure prevalence and levels (e.g. in air or human tissues) is reported by industry, occupation, region, and other characteristics (e.g. process, task) where feasible. Information on historical exposure trends, protection measures to limit exposure, and potential co-exposures to other carcinogenic agents in workplaces is provided when available.

For non-occupational settings, the occurrence of the agent is described with environmental monitoring or surveillance data. Information on exposure prevalence and levels (e.g. concentrations in human tissues) as well as exposure from and/or concentrations in food and beverages, consumer products, consumption practices, and personal microenvironments is reported by region and other relevant characteristics. Particular importance is placed on describing exposures in life stages or in states of disease or nutrition that may involve greater exposure or susceptibility.

Current exposures are of primary interest; however, information on historical exposure trends is provided when available. Historical exposures may be relevant for interpreting epidemiological studies, and when agents are persistent or have long-term effects. Information gaps for important time periods are noted. Exposure data that are not deemed to have high relevance to human exposure are generally not considered.

(e) Regulations and guidelines

Regulations or guidelines that have been established for the agent (e.g. occupational exposure limits, maximum permitted levels in foods and water, pesticide registrations) are described in brief to provide context about government efforts to limit exposure; these may be tabulated if they are informative for the interpretation of existing or historical exposure levels. Information on applicable populations, specific agents concerned, basis for regulation (e.g. human health risk, environmental considerations), and timing of implementation may be noted. National and international bans on production, use, and trade are also indicated.

This section aims to include major or illustrative regulations and may not be comprehensive, because of the complexity and range of regulatory processes worldwide. An absence of information on regulatory status should not be taken to imply that a given country or region lacks exposure to, or regulations on exposure to, the agent.

(f) Critical review of exposure assessment in key epidemiological studies

Epidemiological studies evaluate cancer hazard by comparing outcomes across differently exposed groups. Therefore, the type and quality of the exposure assessment methods used are key considerations when interpreting study findings for hazard identification. This section summarizes and critically reviews the exposure assessment methods used in the individual epidemiological studies that contribute data relevant to the *Monographs* evaluation.

Although there is no standard set of criteria for evaluating the quality of exposure assessment methods across all possible agents, some concepts are universally relevant. Regardless of the agent, all exposures have two principal dimensions: intensity (sometimes defined as concentration or dose) and time. Time considerations include duration (time from first to last exposure), pattern or frequency (whether continuous or intermittent), and windows of susceptibility. This section considers how each of the key epidemiological studies characterizes these dimensions. Interpretation of exposure information may also be informed by consideration of mechanistic evidence (e.g. as described in Part B, Section 4a), including the processes of absorption, distribution, metabolism, and excretion.

Exposure intensity and time in epidemiological studies can be characterized by using environmental or biological monitoring data, records from workplaces or other sources, expert assessments, modelled exposures, job-exposure matrices, and subject or proxy reports via questionnaires or interviews. Investigators use these data sources and methods individually or in combination to assign levels or values of an exposure metric (which may be quantitative, semi-quantitative, or qualitative) to members of the population under study.

In collaboration with the Working Group members reviewing human studies (of cancer and of mechanisms), key epidemiological studies are identified. For each selected study, the exposure assessment approach, along with its strengths and limitations, is summarized using text and tables. Working Group members identify concerns about exposure assessment methods and their impacts on overall quality for each study reviewed (see Part B, Sections 2d and 4d). In situations where the information provided in the study is inadequate to properly consider the exposure assessment, this is indicated. When adequate information is available, the likely direction of bias due to error in exposure measurement, including misclassification (overestimated effects, underestimated effects, or unknown) is discussed.

2. Studies of cancer in humans

This section includes all pertinent epidemiological studies (see Part B, Section 2b) that include cancer as an outcome. These studies encompass certain types of biomarker studies, for example, studies with biomarkers as exposure metrics (see Part B, Section 2) or those evaluating histological or tumour subtypes and molecular signatures in tumours consistent with a given exposure (Alexandrov et al., 2016). Studies that evaluate early biological effect biomarkers are reviewed in Part B, Section 4.

(a) Types of study considered

Several types of epidemiological studies contribute to the assessment of carcinogenicity in humans; they typically include cohort studies (including variants such as case–cohort and nested case–control studies), case–control studies, ecological studies, and intervention studies. Rarely, results from randomized trials may be available. Exceptionally, case reports and case series of cancer in humans may also be reviewed. In addition to these designs, innovations in epidemiology allow for many other variants that may be considered in any given *Monographs* evaluation.

Cohort and case–control studies typically have the capacity to relate individual exposures under study to the occurrence of cancer in individuals, and provide an estimate of effect (such as relative risk) as the main measure of association. Well-conducted cohort and case–control studies provide most of the evidence of cancer in humans evaluated by Working Groups. Intervention studies are much less common, but when available can provide strong evidence for making causal inferences.

In ecological studies, the units of investigation are usually whole populations (e.g. in particular geographical areas or at particular times), and cancer frequency is related to a summary measure of the exposure in the population under study. In ecological studies, data on individual exposure and outcome are not available, which renders this type of study more prone to confounding and exposure misclassification. In some circumstances, however, ecological studies may be informative, especially when the unit of exposure is most accurately measured at the population level (see, for example, the *Monograph* on arsenic in drinking-water; IARC, 2004).

Exceptionally, case reports and case series may provide compelling evidence about the carcinogenicity of an agent. In fact, many of the early discoveries of occupational cancer hazards came about because of observations by workers and their clinicians, who noted a high frequency of cancer in workers who share a common occupation or exposure. Such observations may be the starting point for more structured investigations, but in exceptional circumstances, when the risk is high enough, the case series may in itself provide compelling evidence. This would be especially warranted in situations where the exposure circumstance is fairly unusual, as it was in the example of plants containing aristolochic acid (IARC, 2012a).

The uncertainties that surround the interpretation of case reports, case series, and ecological studies typically make them inadequate, except in rare instances as described above, to form the sole basis for inferring a causal relationship. However, when considered together with cohort and case–control studies, these types of study may support the judgement that a causal relationship exists.

Epidemiological studies of benign neoplasms, pre-neoplastic lesions, malignant precursors, and other end-points are also reviewed when they relate to the agents reviewed. On occasion they can strengthen inferences drawn from studies of cancer itself. For example, benign brain tumours may share common risk factors with those that are malignant, and benign neoplasms (or those of uncertain behaviour) may be part of the causal path to malignancies (e.g. myelodysplastic syndromes, which may progress to acute myeloid leukaemia).

(b) Identification of eligible studies of cancer in humans

Relevant studies of cancer in humans are identified by using systematic review principles as described in Part A, further elaborated in the Instructions for Authors, and as detailed below. Eligible studies include all studies in humans of exposure to the agent of interest with cancer as an outcome. Multiple publications on the same study population are identified so that the number of independent studies is accurately represented. Multiple publications may result, for example, from successive follow-ups of a single cohort, from analyses focused on different aspects of an exposure–disease association, or from inclusion of overlapping populations. Usually in such situations, only the most recent, most comprehensive, or most informative report is reviewed in detail.

(c) Assessment of study quality and informativeness

Epidemiological studies are potentially susceptible to several different sources of error, summarized briefly below. Qualities of individual studies that address these issues are also described below.

Study quality is assessed as part of the structured expert review process undertaken by the Working Group. A key aspect of quality assessment is consideration of the possible roles of chance and bias in the interpretation of epidemiological studies. Chance, which is also called random variation, can produce misleading study results. This variability in study results is strongly influenced by the sample size: smaller studies are more likely than larger studies to have effect estimates that are imprecise. Confidence intervals around a study's point estimate of effect are used routinely to indicate the range of values of the estimate that could easily be produced by chance alone.

Bias is the effect of factors in study design or conduct that lead an association to erroneously appear stronger or weaker than the association that really exists between the agent and the disease. Biases that require consideration are varied but are usually categorized as selection bias, information bias (e.g. error in measurement of exposure and diseases), and confounding (or confounding bias), (Rothman et al., 2008). Selection bias in an epidemiological study occurs when inclusion of participants from the eligible population or their follow-up in the study is influenced by their exposure or their outcome (usually disease occurrence). Under these conditions, the measure of association found in the study will not accurately reflect the association that would otherwise have been found in the eligible population (Hernán et al., 2004). Information bias results from inaccuracy in exposure or outcome measurement. Both can cause an association between hypothesized cause and effect to appear stronger or weaker than it really is. Confounding is a mixing of extraneous effects with the effects of interest (Rothman et al., 2008). An association between the purported causal factor and another factor that is associated with an increase or decrease in incidence of disease can lead to a spurious association or absence of a real association of the presumed causal factor with the disease. When either of these occurs, confounding is present.

In assessing study quality, the Working Group consistently considers the following aspects:

- **Study description:** Clarity in describing the study design and its implementation, and the completeness of reporting of all other key information about the study and its results.
- **Study population:** Whether the study population was appropriate for evaluating the association between the agent and cancer. Whether the study was designed and carried out to minimize selection bias. Cancer cases in the study population must have been identified in a way that was independent of the exposure of interest, and exposure assessed in a way that was not related to disease (outcome) status. In these respects, completeness of recruitment into the study from the population of interest and completeness of follow-up for the outcome are essential measures.
- **Outcome measurement:** The appropriateness of the cancer outcome measure (e.g. mortality vs incidence) for the agent and cancer type under consideration, outcome ascertainment methodology, and the extent to which outcome misclassification may have led to bias in the measure(s) of association.
- **Exposure measurement:** The adequacy of the methods used to assess exposure to the agent, and the likelihood (and direction) of bias in the

measure(s) of association due to error in exposure measurement, including misclassification (as described in Part B, Section 1f).

- **Assessment of potential confounding:** To what extent the authors took into account in the study design and analysis other variables (including co-exposures, as described in Part B, Section 1d) that can influence the risk of disease and may have been related to the exposure of interest. Important sources of potential confounding by such variables should have been addressed either in the design of the study, such as by matching or restriction, or in the analysis, by statistical adjustment. In some instances, where direct information on confounders is unavailable, use of indirect methods to evaluate the potential impact of confounding on exposure–disease associations is appropriate (e.g. Axelson & Steenland, 1988; Richardson et al., 2014).
- **Other potential sources of bias:** Each epidemiological study is unique in its study population, its design, its data collection, and, consequently, its potential biases. All possible sources of bias are considered for their possible impact on the results. The possibility of reporting bias (i.e. selective reporting of some results and the suppression of others) should be explored.
- **Statistical methodology:** Adequacy of the statistical methods used and their ability to obtain unbiased estimates of exposure–outcome associations, confidence intervals, and test statistics for the significance of measures of association. Appropriateness of methods used to investigate confounding, including adjusting for matching when necessary and avoiding treatment of probable mediating variables as confounders. Detailed analyses of cancer risks in relation to summary measures of exposure such as cumulative exposure, or temporal variables such as age at first exposure or time since first exposure, are reviewed and summarized when available.

For the sake of economy and simplicity, in this Preamble the list of possible sources of error is referred to with the phrase "chance, bias, and confounding", but it should be recognized that this phrase encompasses a comprehensive set of concerns pertaining to study quality.

These sources of error do not constitute and should not be used as a formal checklist of indicators of study quality. The judgement of experienced experts is critical in determining how much weight to assign to different issues in considering how all of these potential sources of error should be integrated and how to rate the potential for error related to each of these considerations.

The informativeness of a study is its ability to show a true association, if there is one, between the agent and cancer, and the lack of an association, if no association exists. Key determinants of informativeness include: having a study population of sufficient size to obtain precise estimates of effect; sufficient elapsed time from exposure to measurement of outcome for an effect, if present, to be observable; presence of an adequate exposure contrast (intensity, frequency, and/or duration); biologically relevant definitions of exposure; and relevant and well-defined time windows for exposure and outcome.

(d) Meta-analyses and pooled analyses

Independent epidemiological studies of the same agent may lead to inconsistent results that are difficult to interpret or reconcile. Combined analyses of data from multiple studies may be conducted as a means to address this ambiguity. There are two types of combined analysis. The first involves combining summary statistics such as relative risks from individual studies (meta-analysis), and the second involves a pooled analysis of the raw data from the individual studies (pooled analysis) (Greenland & O'Rourke, 2008).

The strengths of combined analyses are increased precision because of increased sample size and, in the case of pooled analyses, the opportunity to better control for potential confounders and to explore in more detail interactions and modifying effects that may explain heterogeneity among studies. A disadvantage of combined analyses is the possible lack of comparability of data from various studies, because of differences in population characteristics, subject recruitment, procedures of data collection, methods of measurement, and effects of unmeasured covariates that may differ among studies. These differences in study methods and quality can influence results of either meta-analyses or pooled analyses. If published meta-analyses are to be considered by the Working Group, their adequacy needs to be carefully evaluated, including the methods used to identify eligible studies and the accuracy of data extracted from the individual studies.

The Working Group may conduct ad hoc meta-analyses during the course of a *Monographs* meeting, when there are sufficient studies of an exposure–outcome association to contribute to the Working Group's assessment of the association. The results of such unpublished original calculations, which would be specified in the text by presentation in square brackets, might involve updates of previously conducted analyses that incorporate the results of more recent studies, or de novo analyses.

Irrespective of the source of data for the meta-analyses and pooled analyses, the following key considerations apply: the same criteria for data quality must be applied as for individual studies; sources of heterogeneity among studies must be carefully considered; and the possibility of publication bias should be explored.

(e) Considerations in assessing the body of epidemiological evidence

The ability of the body of epidemiological evidence to inform the Working Group about the carcinogenicity of the agent is related to both the quantity and the quality of the evidence. There is no formulaic answer to the question of how many studies of cancer in humans are needed from which to draw inferences about causality, although more than a single study in a single population will almost always be needed. The number will depend on the considerations relating to evidence described below.

After the quality of individual epidemiological studies of cancer has been assessed and the informativeness of the various studies on the association between the agent and cancer has been evaluated, a judgement is made about the strength of evidence that the agent in question is carcinogenic to humans. In making its judgement, the Working Group considers several aspects of the body of evidence (e.g. Hill, 1965; Rothman et al., 2008; Vandenbroucke et al., 2016).

A strong association (e.g. a large relative risk) is more likely to indicate causality than is a weak association, because it is more difficult for confounding to falsely create a strong association. However, it is recognized that estimates of effect of small magnitude do not imply lack of causality and may have impact on public health if the disease or exposure is common. Estimates of effect of small magnitude could also contribute useful information to the assessment of causality if level of risk is commensurate with level of exposure when compared with risk estimates from populations with higher exposure (e.g. as seen in residential radon studies compared with studies of radon from uranium mining).

Associations that are consistently observed in several studies of the same design, or in studies that use different epidemiological approaches, or under different circumstances of exposure are more likely to indicate a causal relationship than are isolated observations from single studies. If there are inconsistent results among investigations, possible reasons are sought (e.g. differences in study informativeness because of latency, exposure levels, or assessment methods). Results of studies that are judged to be of high quality and informativeness are given more weight than those of studies judged to be methodologically less sound or less informative.

Temporality of the association is an essential consideration: that is, the exposure must precede the outcome.

An observation that cancer risk increases with increasing exposure is considered to be a strong indication of causality, although the absence of a graded response is not necessarily evidence against a causal relationship, and there are several reasons why the shape of the exposure–response association may be non-monotonic (e.g. Stayner et al., 2003). The demonstration of a decline in risk after cessation of or reduction in exposure in individuals or in whole populations also supports a causal interpretation of the findings.

Confidence in a causal interpretation of the evidence from studies of cancer in humans is enhanced if it is coherent with physiological and biological knowledge, including information about exposure to the target organ, latency and timing of the exposure, and characteristics of tumour subtypes.

The Working Group considers whether there are subpopulations with increased susceptibility to cancer from the agent. For example, molecular epidemiology studies that identify associations between genetic polymorphisms and inter-individual differences in cancer susceptibility to the agent(s) being evaluated may contribute to the identification of carcinogenic hazards to humans. Such studies may be particularly informative if polymorphisms are found to be modifiers of the exposure–response association, because evaluation of polymorphisms may increase the ability to detect an effect in susceptible subpopulations.

When, in the process of evaluating the studies of cancer in humans, the Working Group identifies several high-quality, informative epidemiological studies that clearly show either no positive association or an inverse association between an exposure and a specific type of cancer, a judgement may be made that, in the aggregate, they suggest evidence of lack of carcinogenicity for that cancer type. Such a judgement requires, first, that the studies strictly meet the standards of design and analysis described above. Specifically, the possibility that bias, confounding, or misclassification of exposure or outcome could explain the observed results should be considered and ruled out with reasonable confidence. In addition, all studies that are judged to be methodologically sound should (a) be consistent with an estimate of relative effect of unity (or below unity) for any observed level of exposure, (b) when considered together, provide a combined estimate of relative risk that is at or below unity, and (c) have a narrow confidence interval. Moreover, neither any individual well-designed and well-conducted study nor the pooled results of all the studies should show any consistent tendency that the relative risk of cancer increases with increasing level of exposure. It must be noted that evidence of lack of carcinogenicity obtained from several epidemiological studies can apply only to the type(s) of cancer studied, to the exposure levels reported and the timing and route of exposure studied, to the intervals between first exposure and disease onset observed in these studies, and to the general population(s) studied (i.e. there may be susceptible subpopulations or life stages). Experience from studies of cancer in humans indicates that the period from first exposure to the development of clinical cancer is sometimes longer than 20 years; therefore, latency periods substantially shorter than about 30 years cannot provide evidence of lack of carcinogenicity. Furthermore, there may be critical windows of exposure, for example, as with diethylstilboestrol and clear cell adenocarcinoma of the cervix and vagina (IARC, 2012a).

3. Studies of cancer in experimental animals

Most human carcinogens that have been studied adequately for carcinogenicity in experimental animals have produced positive results in one or more animal species. For some agents, carcinogenicity in experimental animals was demonstrated before epidemiological studies identified their carcinogenicity in humans. Although this observation cannot establish that all agents that cause cancer in experimental animals also cause cancer in humans, it is biologically plausible that agents for which there is *sufficient evidence of carcinogenicity* in experimental animals (see Part B, Section 6b) present a carcinogenic hazard to humans. Accordingly, in the absence of additional scientific information, such as strong evidence that a given agent causes cancer in experimental animals through a species-specific mechanism that does not operate in humans (see Part B, Sections 4 and 6; Capen et al., 1999; IARC, 2003), these agents are considered to pose a potential carcinogenic hazard to humans. The inference of potential carcinogenic hazard to humans does not imply tumour site concordance across species (Baan et al., 2019).

(a) Types of studies considered

Relevant studies of cancer in experimental animals are identified by using systematic review principles as described in Part A, further elaborated in the Instructions for Authors, and as detailed below. Consideration is given to all available long-term studies of cancer in experimental animals with the agent under review (or possibly metabolites or derivatives of the agent) (see Part A, Section 7) after a thorough evaluation of the study features (see Part B, Section 3b). Those studies that are judged to be irrelevant to the evaluation or judged to be inadequate (e.g. too short a duration, too few animals, poor survival; see below) may be omitted. Guidelines for conducting long-term carcinogenicity experiments have been published (e.g. OECD, 2018).

In addition to conventional long-term bioassays, alternative studies (e.g. in genetically engineered mouse models) may be considered in assessing carcinogenicity in experimental animals, also after a critical evaluation of the study features. For studies of certain exposures, such as viruses that typically only infect humans, use of such specialized experimental animal models may be particularly important; models include genetically engineered mice with targeted expression of viral genes to tissues from which human cancers arise, as well as humanized mice implanted with the human cells usually infected by the virus.

Other types of studies can provide supportive evidence. These include: experiments in which the agent was administered in the presence of factors that modify carcinogenic effects (e.g. initiation–promotion studies); studies in which the end-point was not cancer but a defined precancerous lesion; and studies of cancer in non-laboratory animals (e.g. companion animals) exposed to the agent.

(b) Study evaluation

Considerations of importance in the interpretation and evaluation of a particular study include: (i) whether the agent was clearly characterized, including the nature and extent of impurities and contaminants and the stability of the agent, and, in the case of mixtures, whether the sample characterization was adequately reported; (ii) whether the dose was monitored adequately, particularly in inhalation experiments; (iii) whether the doses, duration and frequency of treatment, duration of observation, and route of exposure were appropriate; (iv) whether appropriate experimental animal species and strains were evaluated; (v) whether there were adequate numbers of animals per group; (vi) whether animals were allocated randomly to groups; (vii) whether the body weight, food and water consumption, and survival of treated animals were affected by any factors other than the test agent; (viii) whether the histopathology review was adequate; and (ix) whether the data were reported and analysed adequately.

(c) Outcomes and statistical analyses

An assessment of findings of carcinogenicity in experimental animals involves consideration of (i) study features such as route, doses, schedule and duration of exposure, species, strain (including genetic background where applicable), sex, age, and duration of follow-up; (ii) the spectrum of neoplastic response, from pre-neoplastic lesions and benign tumours to malignant neoplasms; (iii) the incidence, latency, severity, and multiplicity of neoplasms and pre-neoplastic lesions; (iv) the consistency of the results for a specific target organ or organs across studies of similar design; and (v) the possible role of modifying factors (e.g. diet, infection, stress).

Key factors for statistical analysis include: (i) number of animals studied and number examined histologically, (ii) number of animals with a given tumour type or lesion, and (iii) duration of survival.

Benign tumours may be combined with malignant tumours in the assessment of tumour incidence when (a) they occur together with and originate from the same cell type as malignant tumours in an organ or tissue in a particular study and (b) they appear to represent a stage in the progression to malignancy (Huff et al., 1989). The occurrence of lesions presumed to be pre-neoplastic may in certain instances aid in assessing the biological plausibility of any neoplastic response observed.

Evidence of an increased incidence of neoplasms with increasing level of exposure strengthens the inference of a causal association between the exposure and the development of neoplasms. The form of the dose–response relationship can vary widely, including non-linearity, depending on the particular agent under study and the target organ. The dose–response relationship can also be affected by differences in survival among the treatment groups.

The statistical methods used should be clearly stated and should be the generally accepted techniques refined for this purpose (Peto et al., 1980; Gart et al., 1986; Portier & Bailer, 1989; Bieler & Williams, 1993). The choice of the most appropriate statistical method requires consideration of whether there are differences in survival among the treatment groups; for example, reduced survival because of non-tumour-related mortality can preclude the occurrence of tumours later in life and a survival-adjusted analysis would be warranted. When detailed information on survival is not available, comparisons of the proportions of tumour-bearing animals among the effective number of animals (alive at the time that the first tumour was discovered) can be useful when significant differences in survival occur before tumours appear. The lethality of the tumour also requires consideration: for rapidly fatal tumours, the time of death provides an indication of the time of tumour onset and can be assessed using life-table methods; non-fatal or incidental tumours that do not affect survival can be assessed using methods such as the Mantel–Haenszel test for changes in tumour prevalence. Because tumour lethality is often difficult to determine, methods such as the poly-*k* test that do not require such information can also be used. When results are available on the number and size of tumours seen in experimental animals (e.g. papillomas on mouse skin, liver tumours observed through nuclear magnetic resonance tomography), other, more complicated statistical procedures may be needed (Sherman et al., 1994; Dunson et al., 2003).

The concurrent control group is generally the most appropriate comparison group for statistical analysis; however, for uncommon tumours, the analysis may be improved by considering historical control data, particularly when between-study variability is low. Historical controls should be selected to resemble the concurrent controls as closely as possible with respect to species, sex, and strain, as well as other factors, such as basal diet and general laboratory environment, which may affect tumour response rates in control animals (Haseman et al., 1984; Fung et al., 1996; Greim et al., 2003). It is generally not appropriate to discount a tumour response that is significantly increased compared with concurrent controls by arguing that it falls within the range of historical controls.

Meta-analyses and pooled analyses may be appropriate when the experimental protocols are sufficiently similar.

4. Mechanistic evidence

Mechanistic data may provide evidence of carcinogenicity and may also help in assessing the relevance and importance of findings of cancer in experimental animals and in humans (Guyton et al., 2009; Parkkinen et al., 2018) (see Part B, Section 6). Mechanistic studies have gained in prominence, increasing in their volume, diversity, and relevance to cancer hazard evaluation, whereas studies pertinent to other streams of evidence evaluated in the *Monographs* (i.e. studies of cancer in humans and lifetime cancer bioassays in rodents) may only be available for a fraction of agents to which humans are currently exposed (Guyton et al., 2009, 2018). Mechanistic studies and data are identified, screened, and evaluated for quality and importance to the evaluation by using systematic review principles as described in Part A, further elaborated in the Instructions for Authors, and as detailed below.

The Working Group's synthesis reflects the extent of available evidence, summarizing groups of included studies with an emphasis on characterizing consistencies or differences in results within and across experimental designs. Greater emphasis is given to informative mechanistic evidence from human-related studies than to that from other experimental test systems, and gaps are identified. Tabulation of data may facilitate this review. The specific topics addressed in the evidence synthesis are described below.

(a) Absorption, distribution, metabolism, and excretion

Studies of absorption, distribution, metabolism, and excretion in mammalian species are addressed in a summary fashion; exposure characterization is addressed in Part B, Section 1. The Working Group describes the metabolic fate of the agent in mammalian species, noting the metabolites that have been identified and their chemical reactivity. A metabolic schema may indicate the relevant metabolic pathways and products and whether supporting evidence is from studies in humans and/or studies in experimental animals. Evidence on other adverse effects that indirectly confirm absorption, distribution, and/or metabolism at tumour sites is briefly summarized when direct evidence is sparse.

(b) Evidence relevant to key characteristics of carcinogens

A review of Group 1 human carcinogens classified up to and including *IARC Monographs* Volume 100 revealed several issues relevant to improving the evaluation of mechanistic evidence for cancer hazard identification (Smith et al., 2016). First, it was noted that human carcinogens often share one or more characteristics that are related to the multiple mechanisms by which agents cause cancer. Second, different human carcinogens may exhibit a different spectrum of these key characteristics and operate through distinct mechanisms. Third, for many carcinogens evaluated before Volume 100, few data were available on some mechanisms of recognized importance in carcinogenesis, such as epigenetic alterations (Herceg et al., 2013). Fourth, there was no widely accepted method to search systematically for relevant mechanistic evidence, resulting in a lack of uniformity in the scope of mechanistic topics addressed across *IARC Monographs* evaluations.

To address these challenges, the key characteristics of human carcinogens were introduced to facilitate systematic consideration of mechanistic evidence in *IARC Monographs* evaluations (Smith et al., 2016; Guyton et al., 2018). The key characteristics described by Smith et al. (2016) (see Table 3), such as "is genotoxic", "is immunosuppressive", or "modulates receptor-mediated effects", are based on empirical observations of the chemical and biological properties associated with the human carcinogens identified by the *IARC Monographs* programme up to and including Volume 100. The list of key characteristics and associated end-points may evolve, based on the experience of their application and as new human carcinogens are identified. Key characteristics are distinct from the "hallmarks of cancer", which relate to the properties of cancer cells (Hanahan & Weinberg, 2000, 2011). Key characteristics are also distinct from hypothesized mechanistic pathways, which describe a sequence of biological events postulated to occur during carcinogenesis. As such, the evaluation approach based on key characteristics, outlined below, "avoids a narrow focus on specific pathways and hypotheses and provides for a broad, holistic consideration of the mechanistic evidence" (National Academies of Sciences, Engineering, and Medicine, 2017).

Table 3. The key characteristics of carcinogens described by Smith et al. (2016)

Ten key characteristics of carcinogens
1. Is electrophilic or can be metabolically activated to an electrophile
2. Is genotoxic
3. Alters DNA repair or causes genomic instability
4. Induces epigenetic alterations
5. Induces oxidative stress
6. Induces chronic inflammation
7. Is immunosuppressive
8. Modulates receptor-mediated effects
9. Causes immortalization
10. Alters cell proliferation, cell death, or nutrient supply

Studies in exposed humans and in human primary cells or tissues that incorporate end-points relevant to key characteristics of carcinogens are emphasized when available. For each key characteristic with adequate evidence for evaluation, studies are grouped according to whether they involve (a) humans or human primary cells or tissues or (b) experimental systems; further organization (as appropriate) is by end-point (e.g. DNA damage), duration, species, sex, strain, and target organ as well as strength of study design. Studies investigating susceptibility related to key characteristics of carcinogens (e.g. of genetic polymorphisms, or in genetically engineered animals) can be highlighted and may provide additional support for conclusions on the strength of evidence. Findings relevant to a specific tumour type may be noted.

(c) Other relevant evidence

Other informative evidence may be described when it is judged by the Working Group to be relevant to an evaluation of carcinogenicity and to be of sufficient importance to affect the overall evaluation. Quantitative structure–activity information, such as on specific chemical and/or biological features or activities (e.g. electrophilicity, molecular docking with receptors), may be informative. In addition, evidence that falls outside of the recognized key characteristics of carcinogens, reflecting emerging knowledge or important novel scientific developments on carcinogen mechanisms, may also be included. Available evidence relevant to criteria provided in authoritative publications (e.g. Capen et al., 1999; IARC, 2003) on thyroid, kidney, urinary bladder, or other tumours in experimental animals induced by mechanisms that do not operate in humans is also described.

(d) Study quality and importance to the evaluation

Based on formal considerations of the quality of the studies (e.g. design, methodology, and reporting of results), the Working Group may give greater weight to some included studies.

For observational and other studies in humans, the quality of study design, exposure assessment, and assay accuracy and precision are considered, in collaboration with the Working Group members reviewing exposure characterization and studies of cancer in humans, as are other important factors, including those described above for evaluation of epidemiological evidence (García-Closas et al., 2006, 2011; Vermeulen et al., 2018) (Part B, Sections 1 and 2).

In general, in experimental systems, studies of repeated doses and of chronic exposures are accorded greater importance than are studies of a single dose or time point. Consideration is also given to factors such as the suitability of the dosing range, the extent of concurrent toxicity observed, and the completeness of reporting of the study (e.g. the source and purity of the agent, the analytical methods, and the results). Route of exposure is generally considered to be a less important factor in the evaluation of experimental studies, recognizing that the exposures and target tissues may vary across experimental models and in exposed human populations. Non-mammalian studies can be synthetically summarized when they are considered to be supportive of evidence in humans or higher organisms.

In vitro test systems can provide mechanistic insights, but important considerations include the limitations of the test system (e.g. in metabolic capabilities) as well as the suitability of a particular test article (i.e. because of physical and chemical characteristics) (Hopkins et al., 2004). For studies on some end-points, such as for traditional studies of mutations in bacteria and in mammalian cells, formal guidelines, including those from the Organisation for Economic Co-operation and Development, may be informative in conducting the quality review (OECD, 1997, 2016a, b). However, existing guidelines will not generally cover all relevant assays, even for genotoxicity. Possible considerations when evaluating the quality of in vitro studies encompass the methodology and design (e.g. the end-point and test method, the number of replicate samples, the suitability of the concentration range, the inclusion of positive and negative controls, and the assessment of cytotoxicity) as well as reporting (e.g. of the source and purity of the agent, and of the analytical methods and results). High-content and high-throughput in vitro data can serve as an additional or supportive source of mechanistic evidence (Chiu et al., 2018; Guyton et al., 2018), although large-scale screening programmes measuring a variety of end-points were designed to evaluate large chemical libraries in order to prioritize chemicals for additional toxicity testing rather than to identify the hazard of a specific chemical or chemical group.

The synthesis is focused on the evidence that is most informative for the overall evaluation. In this regard, it is of note that some human carcinogens exhibit a single or primary key characteristic, evidence of which has been influential in their cancer hazard classifications. For instance, ethylene oxide is genotoxic (IARC, 1994), 2,3,7,8-tetrachlorodibenzo-*para*-dioxin modulates receptor-mediated effects (IARC, 1997), and etoposide alters DNA repair (IARC, 2012a). Similarly, oncogenic viruses cause immortalization, and certain drugs are, by design, immunosuppressive (IARC, 2012a, b). Because non-carcinogens can also induce oxidative stress, this key characteristic should be interpreted with caution unless it is found in combination with other key characteristics (Guyton et al., 2018). Evidence for a group of key characteristics can strengthen mechanistic conclusions (e.g. “induces oxidative stress” together with “is electrophilic or can be metabolically activated to an electrophile”, “induces chronic inflammation”, and “is immunosuppressive”); see, for example, 1-bromopropane (IARC, 2018).

5. Summary of data reported

(a) Exposure characterization

Exposure data are summarized to identify the agent and describe its production, use, and occurrence. Information on exposure prevalence and intensity in different settings, including geographical patterns and time trends, may be included. Exposure assessment methods used in key epidemiological studies reviewed by the Working Group are described and evaluated.

(b) Cancer in humans

Results of epidemiological studies pertinent to an evaluation of carcinogenicity in humans are summarized. The overall strengths and limitations of the epidemiological evidence base are highlighted to indicate how the evaluation was reached. The target organ(s) or tissue(s) in which a positive association between the agent and cancer was observed are identified. Exposure–response and other quantitative data may be summarized when available. When the available epidemiological studies pertain to a mixed exposure, process, occupation, or industry, the Working Group seeks to identify the specific agent considered to be most likely to be responsible for any excess risk. The evaluation is focused as narrowly as the available data permit.

(c) Cancer in experimental animals

Results pertinent to an evaluation of carcinogenicity in experimental animals are summarized to indicate how the evaluation was reached. For each animal species, study design, and route of administration, there is a statement about whether an increased incidence, reduced latency, or increased severity or multiplicity of neoplasms or pre-neoplastic lesions was observed, and the tumour sites are indicated. Special conditions resulting in tumours, such as prenatal exposure or single-dose experiments, are mentioned. Negative findings, inverse relationships, dose–response patterns, and other quantitative data are also summarized.

(d) Mechanistic evidence

Results pertinent to an evaluation of the mechanistic evidence on carcinogenicity are summarized to indicate how the evaluation was reached. The summary encompasses the informative studies on absorption, distribution, metabolism, and excretion; on the key characteristics with adequate evidence for evaluation; and on any other aspects of sufficient importance to affect the overall evaluation, including on whether the agent belongs to a class of agents for which one or more members have been classified as carcinogenic or probably carcinogenic to humans, and on criteria with respect to tumours in experimental animals induced by mechanisms that do not operate in humans. For each topic addressed, the main supporting findings are highlighted from exposed humans, human cells or tissues, experimental animals, or in vitro systems. When mechanistic studies are available in exposed humans, the tumour type or target tissue studied may be specified. Gaps in the evidence are indicated (i.e. if no studies were available in exposed humans, in in vivo systems, etc.). Consistency or differences of effects across different experimental systems are emphasized.

6. Evaluation and rationale

Consensus evaluations of the strength of the evidence of cancer in humans, the evidence of cancer in experimental animals, and the mechanistic evidence are made using transparent criteria and defined descriptive terms. The Working Group then develops a consensus overall evaluation of the strength of the evidence of carcinogenicity for each agent under review.

An evaluation of the strength of the evidence is limited to the agents under review. When multiple agents being evaluated are considered by the Working Group to be sufficiently closely related, they may be grouped together for the purpose of a single and unified evaluation of the strength of the evidence.

The framework for these evaluations, described below, may not encompass all factors relevant to a particular evaluation of carcinogenicity. After considering all relevant scientific findings, the Working Group may exceptionally assign the agent to a different category than a strict application of the framework would indicate, while providing a clear rationale for the overall evaluation.

When there are substantial differences of scientific interpretation among the Working Group members, the overall evaluation will be based on the consensus of the Working Group. A summary of the alternative interpretations may be provided, together with their scientific rationale and an indication of the relative degree of support for each alternative.

The categories of the classification refer to the strength of the evidence that an exposure is carcinogenic and not to the risk of cancer from particular exposures. The terms *probably carcinogenic* and *possibly carcinogenic* have no quantitative significance and are used as descriptors of different strengths of evidence of carcinogenicity in humans; *probably carcinogenic* signifies a greater strength of evidence than *possibly carcinogenic*.

(a) Carcinogenicity in humans

Based on the principles outlined in Part B, Section 2, the evidence relevant to carcinogenicity from studies in humans is classified into one of the following categories:

***Sufficient evidence of carcinogenicity*:** A causal association between exposure to the agent and human cancer has been established. That is, a positive association has been observed in the body of evidence on exposure to the agent and cancer in studies in which chance, bias, and confounding were ruled out with reasonable confidence.

***Limited evidence of carcinogenicity*:** A causal interpretation of the positive association observed in the body of evidence on exposure to the agent and cancer is credible, but chance, bias, or confounding could not be ruled out with reasonable confidence.

***Inadequate evidence regarding carcinogenicity*:** The available studies are of insufficient quality, consistency, or statistical precision to permit a conclusion to be drawn about the presence or the absence of a causal association between exposure and cancer, or no data on cancer in humans are available. Common findings that lead to a determination of inadequate evidence of carcinogenicity include: (a) there are no data available in humans; (b) there are data available in humans, but they are of poor quality or informativeness; and (c) there are studies of sufficient quality available in humans, but their results are inconsistent or otherwise inconclusive.

***Evidence suggesting lack of carcinogenicity*:** There are several high-quality studies covering the full range of levels of exposure that humans are known to encounter, which are mutually consistent in not showing a positive association between exposure to the agent and the studied cancers at any observed level of exposure. The results from these studies alone or combined should have narrow confidence intervals with an upper limit below or close to the null value (e.g. a relative risk of unity). Bias and confounding were ruled out with reasonable confidence, and the studies were considered informative. A conclusion of *evidence suggesting lack of carcinogenicity* is limited to the cancer sites, populations and life stages, conditions and levels of exposure, and length of observation covered by the available studies. In addition, the possibility of a very small risk at the levels of exposure studied can never be excluded.

When there is *sufficient evidence*, a separate sentence identifies the target organ(s) or tissue(s) for which a causal interpretation has been established. When there is *limited evidence*, a separate sentence identifies the target organ(s) or tissue(s) for which a positive association between exposure to the agent and the cancer(s) was observed in humans. When there is *evidence suggesting lack of carcinogenicity*, a separate sentence identifies the target organ(s) or tissue(s) where evidence of lack of carcinogenicity was observed in humans. Identification of a specific target organ or tissue as having *sufficient evidence* or *limited evidence* or *evidence suggesting lack of carcinogenicity* does not preclude the possibility that the agent may cause cancer at other sites.

(b) Carcinogenicity in experimental animals

The evidence relevant to carcinogenicity from studies in experimental animals is classified into one of the following categories:

***Sufficient evidence of carcinogenicity*:** A causal relationship has been established between exposure to the agent and cancer in experimental animals based on an increased incidence of malignant neoplasms or of an appropriate combination of benign and malignant neoplasms in (a) two or more species of animals or (b) two or more independent studies in one species carried out at different times or in different laboratories and/or under different protocols. An increased incidence of malignant neoplasms or of an appropriate combination of benign and malignant neoplasms in both sexes of a single species in a well-conducted study, ideally conducted under Good Laboratory Practices (GLP), can also provide *sufficient evidence*.

Exceptionally, a single study in one species and sex may be considered to provide *sufficient evidence of carcinogenicity* when malignant neoplasms occur to an unusual degree with regard to incidence, site, type of tumour, or age at onset, or when there are marked findings of tumours at multiple sites.

Limited evidence of carcinogenicity**:** The data suggest a carcinogenic effect but are limited for making a definitive evaluation because, for example, (a) the evidence of carcinogenicity is restricted to a single experiment and does not meet the criteria for *sufficient eviden*ce; (b) the agent increases the incidence only of benign neoplasms or lesions of uncertain neoplastic potential; (c) the agent increases tumour multiplicity or decreases tumour latency but does not increase tumour incidence; (d) the evidence of carcinogenicity is restricted to initiation–promotion studies; (e) the evidence of carcinogenicity is restricted to observational studies in non-laboratory animals (e.g. companion animals); or (f) there are unresolved questions about the adequacy of the design, conduct, or interpretation of the available studies.

Inadequate evidence regarding carcinogenicity**:** The studies cannot be interpreted as showing either the presence or the absence of a carcinogenic effect because of major qualitative or quantitative limitations, or no data are available on cancer in experimental animals.

Evidence suggesting lack of carcinogenicity**:** Well-conducted studies (e.g. conducted under GLP) involving both sexes of at least two species are available showing that, within the limits of the tests used, the agent was not carcinogenic. The conclusion of *evidence suggesting lack of carcinogenicity* is limited to the species, tumour sites, age at exposure, and conditions and levels of exposure covered by the available studies.

(c) Mechanistic evidence

Based on the principles outlined in Part B, Section 4, the mechanistic evidence is classified into one of the following categories:

Strong mechanistic evidence**:** Results in several different experimental systems are consistent, and the overall mechanistic database is coherent. Further support can be provided by studies that demonstrate experimentally that the suppression of key mechanistic processes leads to the suppression of tumour development. Typically, a substantial number of studies on a range of relevant end-points are available in one or more mammalian species. Quantitative structure–activity considerations, in vitro tests in non-human mammalian cells, and experiments in non-mammalian species may provide corroborating evidence but typically do not in themselves provide strong evidence. However, consistent findings across a number of different test systems in different species may provide strong evidence.

Of note, “strong” relates not to potency but to strength of evidence. The classification applies to three distinct topics:

(a) Strong evidence that the agent belongs, based on mechanistic considerations, to a class of agents for which one or more members have been classified as carcinogenic or probably carcinogenic to humans. The considerations can go beyond quantitative structure–activity relationships to incorporate similarities in biological activity relevant to common key characteristics across dissimilar chemicals (e.g. based on molecular docking, –omics data).

(b) Strong evidence that the agent exhibits key characteristics of carcinogens. In this case, three descriptors are possible:

(1) The strong evidence is in exposed humans. Findings relevant to a specific tumour type may be informative in this determination.

(2) The strong evidence is in human primary cells or tissues. Specifically, the strong findings are from biological specimens obtained from humans (e.g. ex vivo exposure), from human primary cells, and/or, in some cases, from other humanized systems (e.g. a human receptor or enzyme).

(3) The strong evidence is in experimental systems. This may include one or a few studies in human primary cells and tissues.

(c) Strong evidence that the mechanism of carcinogenicity in experimental animals does not operate in humans. Certain results in experimental animals (see Part B, Section 6b) would be discounted, according to relevant criteria and considerations in authoritative publications (e.g. Capen et al., 1999; IARC, 2003). Typically, this classification would not apply when there is strong mechanistic evidence that the agent exhibits key characteristics of carcinogens.

***Limited mechanistic evidence*:** The evidence is suggestive, but, for example, (a) the studies cover a narrow range of experiments, relevant end-points, and/or species; (b) there are unexplained inconsistencies in the studies of similar design; and/or (c) there is unexplained incoherence across studies of different end-points or in different experimental systems.

***Inadequate mechanistic evidence*:** Common findings that lead to a determination of inadequate mechanistic evidence include: (a) few or no data are available; (b) there are unresolved questions about the adequacy of the design, conduct, or interpretation of the studies; (c) the available results are negative.

(d) Overall evaluation

Finally, the bodies of evidence included within each stream of evidence are considered as a whole, in order to reach an overall evaluation of the carcinogenicity of the agent to humans. The three streams of evidence are integrated and the agent is classified into one of the following categories (see Table 4), indicating that the Working Group has established that:

The agent is *carcinogenic to humans* (Group 1)

This category applies whenever there is *sufficient evidence of carcinogenicity* in humans.

In addition, this category may apply when there is both *strong evidence in exposed humans that the agent exhibits key characteristics of carcinogens* and *sufficient evidence of carcinogenicity* in experimental animals.

The agent is *probably carcinogenic to humans* (Group 2A)

This category generally applies when the Working Group has made at least *two of the following* evaluations, *including at least one* that involves either exposed humans or human cells or tissues:

- *Limited evidence of carcinogenicity* in humans,
- *Sufficient evidence of carcinogenicity* in experimental animals,
- *Strong evidence that the agent exhibits key characteristics of carcinogens*.

If there is *inadequate evidence regarding carcinogenicity* in humans, there should be *strong evidence in human cells or tissues that the agent exhibits key characteristics of carcinogens*. If there is *limited evidence of carcinogenicity in humans*, then the second individual evaluation may be from experimental systems (i.e. s*ufficient evidence of carcinogenicity* in experimental animals or *strong evidence in experimental systems that the agent exhibits key characteristics of carcinogens).*

Additional considerations apply when there is *strong evidence that the mechanism of carcinogenicity in experimental animals does not operate in humans* for one or more tumour sites. Specifically, the remaining tumour sites should still support an evaluation of *sufficient evidence in experimental animals* in order for this evaluation to be used to support an overall classification in Group 2A.

Separately, this category generally applies if there is *strong evidence that the agent belongs, based on mechanistic considerations, to a class of agents for which one or more members have been classified in Group 1 or Group 2A.*

The agent is *possibly carcinogenic to humans* (Group 2B)

This category generally applies when only one of the following evaluations has been made by the Working Group:

- *Limited evidence of carcinogenicity* in humans,
- *Sufficient evidence of carcinogenicity* in experimental animals,
- *Strong evidence that the agent exhibits key characteristics of carcinogens.*

Because this category can be based on evidence from studies in experimental animals alone, there is **no** requirement that the strong mechanistic evidence be in exposed humans or in human cells or tissues. This category may be based on *strong evidence in experimental systems that the agent exhibits key characteristics of carcinogens.*

As with Group 2A, additional considerations apply when there is *strong evidence that the mechanism of carcinogenicity in experimental animals does not operate in humans* for one or more tumour sites. Specifically, the remaining tumour sites should still support an evaluation of *sufficient evidence in experimental animals* in order for this evaluation to be used to support an overall classification in Group 2B.

The agent is *not classifiable as to its carcinogenicity to humans* (Group 3)

Agents that do not fall into any other group are generally placed in this category.

This includes the case when there is *strong evidence that the mechanism of carcinogenicity in experimental animals does not operate in humans* for one or more tumour sites in experimental animals, the remaining tumour sites do not support an evaluation of *sufficient evidence in experimental animals*, and other categories are not supported by data from studies in humans and mechanistic studies.

An evaluation in Group 3 is not a determination of non-carcinogenicity or overall safety. It often means that the agent is of unknown carcinogenic potential and that there are significant gaps in research.

If the evidence suggests that the agent exhibits no carcinogenic activity, either through *evidence suggesting lack of carcinogenicity* in both humans and experimental animals, or through *evidence suggesting lack of carcinogenicity* in experimental animals complemented by strong negative mechanistic evidence in assays relevant to human cancer, then the Working Group may add a sentence to the evaluation to characterize the agent as well-studied and without evidence of carcinogenic activity.

(e) Rationale

The reasoning that the Working Group used to reach its evaluation is summarized so that the basis for the evaluation offered is transparent. This section integrates the major findings from studies of cancer in humans, cancer in experimental animals, and mechanistic evidence. It includes concise statements of the principal line(s) of argument that emerged in the deliberations of the Working Group, the conclusions of the Working Group on the strength of the evidence for each stream of evidence, an indication of the body of evidence that was pivotal to these conclusions, and an explanation of the reasoning of the Working Group in making its evaluation.

Table 4. Integration of streams of evidence in reaching overall classifications (the evidence in *bold italic* represents the basis of the overall evaluation)

Stream of evidence			**Classification based on strength of evidence**
Evidence of cancer in humans[a]	**Evidence of cancer in experimental animals**	**Mechanistic evidence**	
Sufficient	Not necessary	Not necessary	**Carcinogenic to humans (Group 1)**
Limited or Inadequate	***Sufficient***	***Strong (b)(1) (exposed humans)***	
Limited	***Sufficient***	Strong (b)(2–3), Limited, or Inadequate	**Probably carcinogenic to humans (Group 2A)**
Inadequate	***Sufficient***	***Strong (b)(2) (human cells or tissues)***	
Limited	Less than Sufficient	***Strong (b)(1–3)***	
Limited or Inadequate	Not necessary	***Strong (a) (mechanistic class)***	
Limited	Less than Sufficient	Limited or Inadequate	**Possibly carcinogenic to humans (Group 2B)**
Inadequate	***Sufficient***	Strong (b)(3), Limited, or Inadequate	
Inadequate	Less than Sufficient	***Strong b(1–3)***	
Limited	***Sufficient***	***Strong (c) (does not operate in humans)***[b]	
Inadequate	***Sufficient***	***Strong (c) (does not operate in humans)***[b]	**Not classifiable as to its carcinogenicity to humans (Group 3)**
All other situations not listed above			

[a] Human cancer(s) with highest evaluation

[b] The *strong evidence that the mechanism of carcinogenicity in experimental animals does not operate in humans* must specifically be for the tumour sites supporting the classification of *sufficient evidence in experimental animals*.

References

Alexandrov LB, Ju YS, Haase K, Van Loo P, Martincorena I, Nik-Zainal S, et al. (2016). Mutational signatures associated with tobacco smoking in human cancer. Science. 354(6312):618–22. https://doi.org/10.1126/science.aag0299 PMID:27811275

Axelson O, Steenland K (1988). Indirect methods of assessing the effects of tobacco use in occupational studies. Am J Ind Med. 13(1):105–18. https://doi.org/10.1002/ajim.4700130107 PMID:3344750

Baan RA, Stewart BW, Straif K, editors (2019). Tumour site concordance and mechanisms of carcinogenesis (IARC Scientific Publication No. 165). Lyon, France: International Agency for Research on Cancer. Available from: http://publications.iarc.fr/578.

Bieler GS, Williams RL (1993). Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity. Biometrics. 49(3):793–801. https://doi.org/10.2307/2532200 PMID:8241374

Capen CC, Dybing E, Rice JM, Wilbourn JD, editors (1999). Species differences in thyroid, kidney and urinary bladder carcinogenesis (IARC Scientific Publication No. 147). Lyon, France: IARC Press. Available from: http://publications.iarc.fr/302 PMID:10627184

Chiu WA, Guyton KZ, Martin MT, Reif DM, Rusyn I (2018). Use of high-throughput in vitro toxicity screening data in cancer hazard evaluations by IARC Monograph Working Groups. ALTEX. 35(1):51–64. https://doi.org/10.14573/altex.1703231 PMID:28738424

Cogliano V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, et al. (2005). Transparency in IARC Monographs. Lancet Oncol. 6(10):747. https://doi.org/10.1016/S1470-2045(05)70380-6

Cogliano VJ, Baan RA, Straif K, Grosse Y, Secretan MB, El Ghissassi F, et al. (2004). The science and practice of carcinogen identification and evaluation. Environ Health Perspect. 112(13):1269–74. https://doi.org/10.1289/ehp.6950 PMID:15345338

Dunson DB, Chen Z, Harry J (2003). A Bayesian approach for joint modeling of cluster size and subunit-specific outcomes. Biometrics. 59(3):521–30. https://doi.org/10.1111/1541-0420.00062 PMID:14601753

Fung KY, Krewski D, Smythe RT (1996). A comparison of tests for trend with historical controls in carcinogen bioassay. Can J Stat. 24(4):431–54. https://doi.org/10.2307/3315326

Garcia-Closas M, Vermeulen R, Cox D, Lan Q, Caporaso N, Hunter D, et al. (2011). Population-based study designs in molecular epidemiology. In: Rothman N, Hainaut P, Schulte P, Smith M, Boffetta P, Perera FP, editors. Molecular epidemiology: principles and practices. Lyon, France: International Agency for Research on Cancer; pp. 241–59.

García-Closas M, Vermeulen R, Sherman ME, Moore LE, Smith MT, Rothman N (2006). Application of biomarkers in cancer epidemiology. In: Schottenfeld D, Fraumeni JF Jr, editors. Cancer epidemiology and prevention. 3rd ed. New York (NY), USA: Oxford University Press; pp. 70–88. https://doi.org/10.1093/acprof:oso/9780195149616.003.0006

Gart JJ, Krewski D, Lee PN, Tarone RE, Wahrendorf J (1986). Statistical methods in cancer research, Volume III: The design and analysis of long-term animal experiments (IARC Scientific Publication No. 79). Lyon, France: International Agency for Research on Cancer.

Greenland S, O'Rourke K (2008). Meta-analysis. In: Rothman KJ, Greenland S, Lash TL, editors. Modern epidemiology. 3rd ed. Philadelphia (PA), USA: Wolters Kluwer Health/Lippincott Williams & Wilkins; pp. 652–682.

Greim H, Gelbke H-P, Reuter U, Thielmann HW, Edler L (2003). Evaluation of historical control data in carcinogenicity studies. Hum Exp Toxicol. 22(10):541–9. https://doi.org/10.1191/0960327103ht394oa PMID:14655720

Guyton KZ, Kyle AD, Aubrecht J, Cogliano VJ, Eastmond DA, Jackson M, et al. (2009). Improving prediction of chemical carcinogenicity by considering multiple mechanisms and applying toxicogenomic approaches. Mutat Res. 681(2–3):230–40. https://doi.org/10.1016/j.mrrev.2008.10.001 PMID:19010444

Guyton KZ, Rusyn I, Chiu WA, Corpet DE, van den Berg M, Ross MK, et al. (2018). Application of the key characteristics of carcinogens in cancer hazard identification. Carcinogenesis. 39(4):614–22. https://doi.org/10.1093/carcin/bgy031 PMID:29562322

Hanahan D, Weinberg RA (2000). The hallmarks of cancer. Cell. 100(1):57–70. https://doi.org/10.1016/S0092-8674(00)81683-9 PMID:10647931

Hanahan D, Weinberg RA (2011). Hallmarks of cancer: the next generation. Cell. 144(5):646–74. https://doi.org/10.1016/j.cell.2011.02.013 PMID:21376230

Haseman JK, Huff J, Boorman GA (1984). Use of historical control data in carcinogenicity studies in rodents. Toxicol Pathol. 12(2):126–35. https://doi.org/10.1177/019262338401200203 PMID:11478313

Herceg Z, Lambert MP, van Veldhoven K, Demetriou C, Vineis P, Smith MT, et al. (2013). Towards incorporating epigenetic mechanisms into carcinogen identification and evaluation. Carcinogenesis. 34(9):1955–67. https://doi.org/10.1093/carcin/bgt212 PMID:23749751

Hernán MA, Hernández-Díaz S, Robins JM (2004). A structural approach to selection bias. Epidemiology. 15(5):615–25. https://doi.org/10.1097/01.ede.0000135174.63482.43 PMID:15308962

Hill AB (1965). The environment and disease: association or causation? Proc R Soc Med. 58:295–300. PMID:14283879

Hopkins AL, Groom CR, Alex A (2004). Ligand efficiency: a useful metric for lead selection. Drug Discov Today. 9(10):430–1. https://doi.org/10.1016/S1359-6446(04)03069-7 PMID:15109945

Huff JE, Eustis SL, Haseman JK (1989). Occurrence and relevance of chemically induced benign neoplasms in long-term carcinogenicity studies. Cancer Metastasis Rev. 8(1):1–22. https://doi.org/10.1007/BF00047055 PMID:2667783

IARC (1994). Some industrial chemicals. IARC Monogr Eval Carcinog Risks Hum. 60:1–560. Available from: http://publications.iarc.fr/78 PMID:7869568

IARC (1997). Polychlorinated dibenzo-*para*-dioxins and polychlorinated dibenzofurans. IARC Monogr Eval Carcinog Risks Hum. 69:1–631. Available from: http://publications.iarc.fr/87 PMID:9379504

IARC (2003). Predictive value of rodent forestomach and gastric neuroendocrine tumours in evaluating carcinogenic risks to humans (IARC Technical Publication No. 39). Lyon, France: IARC Press.

IARC (2004). Some drinking-water disinfectants and contaminants, including arsenic. IARC Monogr Eval Carcinog Risks Hum. 84:1–477. Available from: http://publications.iarc.fr/102 PMID:15645577

IARC (2012a). Pharmaceuticals. IARC Monogr Eval Carcinog Risks Hum. 100A:1–437. Available from: http://publications.iarc.fr/118 PMID:23189749

IARC (2012b). Biological agents. IARC Monogr Eval Carcinog Risks Hum. 100B:1–441. Available from: http://publications.iarc.fr/119 PMID:23189750

IARC (2014). Report of the IARC Advisory Group to Recommend on Quantitative Risk Characterization (IARC Internal Report 14/001). Available from: https://monographs.iarc.fr/wp-content/uploads/2018/06/14-001.pdf.

IARC (2018). Some industrial chemicals. IARC Monogr Eval Carcinog Risks Hum. 115:1–292. Available from: http://publications.iarc.fr/563

National Academies of Sciences, Engineering, and Medicine (2017). Using 21st century science to improve risk-related evaluations. Washington (DC), USA: National Academies Press. https://doi.org/10.17226/24635 PMID:28267305

National Research Council (2012). Exposure science in the 21st century: a vision and a strategy. Washington (DC), USA: National Academies Press. https://doi.org/10.17226/13507 PMID:24901193

OECD (1997). Test no. 471: Bacterial reverse mutation test, OECD Guidelines for the Testing of Chemicals, Section 4. Paris, France: Organisation for Economic Co-operation and Development. https://doi.org/10.1787/9789264071247-en

OECD (2016a). Test no. 476: *In vitro* mammalian cell gene mutation tests using the *Hprt* and *xprt* genes, OECD Guidelines for the Testing of Chemicals, Section 4. Paris, France: Organisation for Economic Co-operation and Development. https://dx.doi.org/10.1787/9789264264809-en

OECD (2016b). Test no. 490: *In vitro* mammalian cell gene mutation tests using the thymidine kinase gene, OECD Guidelines for the Testing of Chemicals, Section 4. Paris, France: Organisation for Economic Co-operation and Development. https://doi.org/10.1787/9789264264908-en

OECD (2018). Test no. 451: Carcinogenicity studies, OECD Guidelines for the Testing of Chemicals, Section 4: Health effects. Paris, France: Organisation for Economic Co-operation and Development. https://doi.org/10.1787/9789264071186-en

Parkkinen VP, Wallmann C, Wilde M, Clarke B, Illari P, Kelly MP, et al. (2018). Evaluating evidence of mechanisms in medicine: principles and procedures. Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-94610-8

Peto R, Pike MC, Day NE, Gray RG, Lee PN, Parish S, et al. (1980). Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal experiments. In: Long-term and short-term screening assays for carcinogens: a critical appraisal. IARC Monogr Eval Carcinog Risk Chem Hum Suppl. 2 Suppl. Lyon, France: International Agency for Research on Cancer; pp. 311–426. Available from: http://publications.iarc.fr/134 PMID:6935176

Portier CJ, Bailer AJ (1989). Testing for increased carcinogenicity using a survival-adjusted quantal response test. Fundam Appl Toxicol. 12(4):731–7. https://doi.org/10.1016/0272-0590(89)90004-3 PMID:2744275

Richardson DB, Laurier D, Schubauer-Berigan MK, Tchetgen Tchetgen E, Cole SR (2014). Assessment and indirect adjustment for confounding by smoking in cohort studies using relative hazards models. Am J Epidemiol. 180(9):933–40. https://doi.org/10.1093/aje/kwu211 PMID:25245043

Rothman KJ, Greenland S, Poole C, Lash TL (2008). Causation and causal inference. In: Rothman KJ, Greenland S, Lash TL, editors. Modern epidemiology. 3rd ed. Philadelphia (PA), USA: Wolters Kluwer Health/Lippincott Williams & Wilkins; pp. 5–31.

Sherman CD, Portier CJ, Kopp-Schneider A (1994). Multistage models of carcinogenesis: an approximation for the size and number distribution of late-stage clones. Risk Anal. 14(6):1039–48. https://doi.org/10.1111/j.1539-6924.1994.tb00074.x PMID:7846311

Smith MT, Guyton KZ, Gibbons CF, Fritz JM, Portier CJ, Rusyn I, et al. (2016). Key characteristics of carcinogens as a basis for organizing data on mechanisms of carcinogenesis. Environ Health Perspect. 124(6):713–21. https://doi.org/10.1289/ehp.1509912 PMID:26600562

Stayner L, Steenland K, Dosemeci M, Hertz-Picciotto I (2003). Attenuation of exposure-response curves in occupational cohort studies at high exposure levels. Scand J Work Environ Health. 29(4):317–24. https://doi.org/10.5271/sjweh.737 PMID:12934726

Vandenbroucke JP, Broadbent A, Pearce N (2016). Causality and causal inference in epidemiology: the need for a pluralistic approach. Int J Epidemiol. 45(6):1776–86. https://doi.org/10.1093/ije/dyv341 PMID:26800751

Vermeulen R, Bell DA, Jones DP, Garcia-Closas M, Spira A, Wang TW, et al. (2018). Application of biomarkers in cancer epidemiology. In: Thun MJ, Linet MS, Cerhan JR, Haiman CA, Schottenfeld D, editors. Schottenfeld and Fraumeni cancer epidemiology and prevention. 4th ed. New York (NY), USA: Oxford University Press; pp. 77–96.

Wild CP, Cogliano VJ (2011). A plea on behalf of expert evaluation and the experts involved. Int J Epidemiol. 40(1):253–61. https://doi.org/10.1093/ije/dyq038 PMID:20223796