IARC MONOGRAPHS ON THE EVALUATION OF CARCINOGENIC RISKS TO HUMANS

PREAMBLE

The Preamble to the *IARC Monographs* describes the objective and scope of the programme, the scientific principles and procedures used in developing a *Monograph*, the types of evidence considered and the scientific criteria that guide the evaluations. The Preamble should be consulted when reading a *Monograph* or list of evaluations.

A. GENERAL PRINCIPLES AND PROCEDURES

1. Background

Soon after IARC was established in 1965, it received frequent requests for advice on the carcinogenic risk of chemicals, including requests for lists of known and suspected human carcinogens. It was clear that it would not be a simple task to summarize adequately the complexity of the information that was available, and IARC began to consider means of obtaining international expert opinion on this topic. In 1970, the 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 IARC Governing Council adopted a resolution concerning the role of IARC in providing government authorities with expert, independent, scientific opinion on environmental carcinogenesis. As one means to that end, the Governing Council recommended that IARC should prepare monographs on the evaluation of carcinogenic risk of chemicals to man, which became the initial title of the series.

In the succeeding years, the scope of the programme broadened as *Monographs* were developed for groups of related chemicals, complex mixtures, occupational exposures, physical and biological agents and lifestyle factors. In 1988, the phrase 'of chemicals' was dropped from the title, which assumed its present form, *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*.

Through the *Monographs* programme, IARC seeks to identify the causes of human cancer. This is the first step in cancer prevention, which is needed as much today as when

IARC was established. The global burden of cancer is high and continues to increase: the annual number of new cases was estimated at 10.1 million in 2000 and is expected to reach 15 million by 2020 (Stewart & Kleihues, 2003). With current trends in demographics and exposure, the cancer burden has been shifting from high-resource countries to low- and medium-resource countries. As a result of *Monographs* evaluations, national health agencies have been able, on scientific grounds, to take measures to reduce human exposure to carcinogens in the workplace and in the environment.

The criteria established in 1971 to evaluate carcinogenic risks to humans were adopted by the Working Groups whose deliberations resulted in the first 16 volumes of the *Monographs* series. Those criteria were subsequently updated by further ad-hoc Advisory Groups (IARC, 1977, 1978, 1979, 1982, 1983, 1987, 1988, 1991; Vainio *et al.*, 1992; IARC, 2005, 2006).

The Preamble is primarily a statement of scientific principles, rather than a specification of working procedures. The procedures through which a Working Group implements these principles are not specified in detail. They usually involve operations that have been established as being effective during previous *Monograph* meetings but remain, predominantly, the prerogative of each individual Working Group.

2. Objective and scope

The objective of the programme is to prepare, with the help of international Working Groups of experts, and to publish in the form of *Monographs*, critical reviews and evaluations of evidence on the carcinogenicity of a wide range of human exposures. The *Monographs* represent the first step in carcinogen risk assessment, which involves examination of all relevant information in order to assess the strength of the available evidence that an agent could alter the age-specific incidence of cancer in humans. The *Monographs* may also indicate where additional research efforts are needed, specifically when data immediately relevant to an evaluation are not available.

In this Preamble, the term 'agent' refers to any entity or circumstance that is subject to evaluation in a *Monograph*. As the scope of the programme has broadened, categories of agents now include specific chemicals, groups of related chemicals, complex mixtures, occupational or environmental exposures, cultural or behavioural practices, biological organisms and physical agents. This list of categories may expand as causation of, and susceptibility to, malignant disease become more fully understood.

A cancer 'hazard' is an agent that is capable of causing cancer under some circumstances, while a cancer 'risk' is an estimate of the carcinogenic effects expected from exposure to a cancer hazard. The *Monographs* are an exercise in evaluating cancer hazards, despite the historical presence of the word 'risks' in the title. The distinction between hazard and risk is important, and the *Monographs* identify cancer hazards even when risks are very low at current exposure levels, because new uses or unforeseen exposures could engender risks that are significantly higher.

In the *Monographs*, an agent is termed 'carcinogenic' if it is capable of increasing the incidence of malignant neoplasms, reducing their latency, or increasing their severity or multiplicity. The induction of benign neoplasms may in some circumstances (see Part B, Section 3a) contribute to the judgement that the agent is carcinogenic. The terms 'neoplasm' and 'tumour' are used interchangeably.

The Preamble continues the previous usage of the phrase 'strength of evidence' as a matter of historical continuity, although it should be understood that *Monographs* evaluations consider studies that support a finding of a cancer hazard as well as studies that do not.

Some epidemiological and experimental studies indicate that different agents may act at different stages in the carcinogenic process, and several different mechanisms may be involved. The aim of the *Monographs* has been, from their inception, to evaluate evidence of carcinogenicity at any stage in the carcinogenesis process, independently of the underlying mechanisms. Information on mechanisms may, however, be used in making the overall evaluation (IARC, 1991; Vainio *et al.*, 1992; IARC, 2005, 2006; see also Part B, Sections 4 and 6). As mechanisms of carcinogenesis are elucidated, IARC convenes international scientific conferences to determine whether a broad-based consensus has emerged on how specific mechanistic data can be used in an evaluation of human carcinogenicity. The results of such conferences are reported in IARC Scientific Publications, which, as long as they still reflect the current state of scientific knowledge, may guide subsequent Working Groups.

Although the *Monographs* have emphasized hazard identification, important issues may also involve dose–response assessment. In many cases, the same epidemiological and experimental studies used to evaluate a cancer hazard can also be used to estimate a dose–response relationship. A *Monograph* may undertake to estimate dose–response relationships within the range of the available epidemiological data, or it may compare the dose–response information from experimental and epidemiological studies. In some cases, a subsequent publication may be prepared by a separate Working Group with expertise in quantitative dose–response assessment.

The *Monographs* are used by national and international authorities to make risk assessments, formulate decisions concerning preventive measures, provide effective cancer control programmes and decide among alternative options for public health decisions. The evaluations of IARC Working Groups are scientific, qualitative judgements on the evidence for or against carcinogenicity provided by the available data. These evaluations represent only one part of the body of information on which public health decisions may be based. Public health options vary from one situation to another and from country to country and relate to many factors, including different socioeconomic and national priorities. Therefore, no recommendation is given with regard to regulation or legislation, which are the responsibility of individual governments or other international organizations.

3. Selection of agents for review

Agents are selected for review on the basis of two main criteria: (a) there is evidence of human exposure and (b) there is some evidence or suspicion of carcinogenicity. Mixed exposures may occur in occupational and environmental settings and as a result of individual and cultural habits (such as tobacco smoking and dietary practices). Chemical analogues and compounds with biological or physical characteristics similar to those of suspected carcinogens may also be considered, even in the absence of data on a possible carcinogenic effect in humans or experimental animals.

The scientific literature is surveyed for published data relevant to an assessment of carcinogenicity. Ad-hoc Advisory Groups convened by IARC in 1984, 1989, 1991, 1993, 1998 and 2003 made recommendations as to which agents should be evaluated in the *Monographs* series. Recent recommendations are available on the *Monographs* programme website (http://monographs.iarc.fr). IARC may schedule other agents for review as it becomes aware of new scientific information or as national health agencies identify an urgent public health need related to cancer.

As significant new data become available on an agent for which a *Monograph* exists, a re-evaluation may be made at a subsequent meeting, and a new *Monograph* published. In some cases it may be appropriate to review only the data published since a prior evaluation. This can be useful for updating a database, reviewing new data to resolve a previously open question or identifying new tumour sites associated with a carcinogenic agent. Major changes in an evaluation (e.g. a new classification in Group 1 or a determination that a mechanism does not operate in humans, see Part B, Section 6) are more appropriately addressed by a full review.

4. Data for the *Monographs*

Each *Monograph* reviews all pertinent epidemiological studies and cancer bioassays in experimental animals. Those judged inadequate or irrelevant to the evaluation may be cited but not summarized. If a group of similar studies is not reviewed, the reasons are indicated.

Mechanistic and other relevant data are also reviewed. A *Monograph* does not necessarily cite all the mechanistic literature concerning the agent being evaluated (see Part B, Section 4). Only those data considered by the Working Group to be relevant to making the evaluation are included.

With regard to epidemiological studies, cancer bioassays, and mechanistic and other relevant data, only reports that have been published or accepted for publication in the openly available scientific literature are reviewed. The same publication requirement applies to studies originating from IARC, including meta-analyses or pooled analyses commissioned by IARC in advance of a meeting (see Part B, Section 2c). Data from government agency reports that are publicly available are also considered. Exceptionally,

doctoral theses and other material that are in their final form and publicly available may be reviewed.

Exposure data and other information on an agent under consideration are also reviewed. In the sections on chemical and physical properties, on analysis, on production and use and on occurrence, published and unpublished sources of information may be considered.

Inclusion of a study does not imply acceptance of the adequacy of the study design or of the analysis and interpretation of the results, and limitations are clearly outlined in square brackets at the end of each study description (see Part B). The reasons for not giving further consideration to an individual study also are indicated in the square brackets.

5. Meeting participants

Five categories of participant can be present at *Monograph* meetings.

(a) The Working Group is responsible for the critical reviews and evaluations that are developed during the meeting. The tasks of Working Group Members are: (i) to ascertain that all appropriate data have been collected; (ii) to select the data relevant for the evaluation on the basis of scientific merit; (iii) to prepare accurate summaries of the data to enable the reader to follow the reasoning of the Working Group; (iv) to evaluate the results of epidemiological and experimental studies on cancer; (v) to evaluate data relevant to the understanding of mechanisms of carcinogenesis; and (vi) to make an overall evaluation of the carcinogenicity of the exposure to humans. Working Group Members generally have published significant research related to the carcinogenicity of the agents being reviewed, and IARC uses literature searches to identify most experts. Working Group Members are selected on the basis of (a) knowledge and experience and (b) absence of real or apparent conflicts of interests. Consideration is also given to demographic diversity and balance of scientific findings and views.

(b) Invited Specialists are experts who also have critical knowledge and experience but have a real or apparent conflict of interests. These experts are invited when necessary to assist in the Working Group by contributing their unique knowledge and experience during subgroup and plenary discussions. They may also contribute text on non-influential issues in the section on exposure, such as a general description of data on production and use (see Part B, Section 1). Invited Specialists do not serve as meeting chair or subgroup chair, draft text that pertains to the description or interpretation of cancer data, or participate in the evaluations.

(c) Representatives of national and international health agencies often attend meetings because their agencies sponsor the programme or are interested in the subject of a meeting. Representatives do not serve as meeting chair or subgroup chair, draft any part of a *Monograph,* or participate in the evaluations.

(d) Observers with relevant scientific credentials may be admitted to a meeting by IARC in limited numbers. Attention will be given to achieving a balance of Observers

from constituencies with differing perspectives. They are invited to observe the meeting and should not attempt to influence it. Observers do not serve as meeting chair or subgroup chair, draft any part of a *Monograph*, or participate in the evaluations. At the meeting, the meeting chair and subgroup chairs may grant Observers an opportunity to speak, generally after they have observed a discussion. Observers agree to respect the Guidelines for Observers at *IARC Monographs* meetings (available at http://monographs.iarc.fr).

(e) The IARC Secretariat consists of scientists who are designated by IARC and who have relevant expertise. They serve as rapporteurs and participate in all discussions. When requested by the meeting chair or subgroup chair, they may also draft text or prepare tables and analyses.

Before an invitation is extended, each potential participant, including the IARC Secretariat, completes the WHO Declaration of Interests to report financial interests, employment and consulting, and individual and institutional research support related to the subject of the meeting. IARC assesses these interests to determine whether there is a conflict that warrants some limitation on participation. The declarations are updated and reviewed again at the opening of the meeting. Interests related to the subject of the meeting are disclosed to the meeting participants and in the published volume (Cogliano *et al.*, 2004).

The names and principal affiliations of participants are available on the *Monographs* programme website (http://monographs.iarc.fr) approximately two months before each meeting. 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).

All participants are listed, with their principal affiliations, at the beginning of each volume. Each participant who is a Member of a Working Group serves as an individual scientist and not as a representative of any organization, government or industry.

6. Working procedures

A separate Working Group is responsible for developing each volume of *Monographs*. A volume contains one or more *Monographs*, which can cover either a single agent or several related agents. Approximately one year in advance of the meeting of a Working Group, the agents to be reviewed are announced on the *Monographs* programme website (http://monographs.iarc.fr) and participants are selected by IARC staff in consultation with other experts. Subsequently, relevant biological and epidemiological data are collected by IARC from recognized sources of information on carcinogenesis, including data storage and retrieval systems such as PubMed. Meeting participants who are asked to prepare preliminary working papers for specific sections are expected to supplement the IARC literature searches with their own searches.

For most chemicals and some complex mixtures, the major collection of data and the preparation of working papers for the sections on chemical and physical properties, on analysis, on production and use, and on occurrence are carried out under a separate

contract funded by the US National Cancer Institute. Industrial associations, labour unions and other knowledgeable organizations may be asked to provide input to the sections on production and use, although this involvement is not required as a general rule. Information on production and trade is obtained from governmental, trade and market research publications and, in some cases, by direct contact with industries. Separate production data on some agents may not be available for a variety of reasons (e.g. not collected or made public in all producing countries, production is small). Information on uses may be obtained from published sources but is often complemented by direct contact with manufacturers. Efforts are made to supplement this information with data from other national and international sources.

Six months before the meeting, the material obtained is sent to meeting participants to prepare preliminary working papers. The working papers are compiled by IARC staff and sent, prior to the meeting, to Working Group Members and Invited Specialists for review.

The Working Group meets at IARC for seven to eight days to discuss and finalize the texts and to formulate the evaluations. The objectives of the meeting are peer review and consensus. During the first few days, four subgroups (covering exposure data, cancer in humans, cancer in experimental animals, and mechanistic and other relevant data) review the working papers, develop a joint subgroup draft and write summaries. Care is taken to ensure that each study summary is written or reviewed by someone not associated with the study being considered. During the last few days, the Working Group meets in plenary session to review the subgroup drafts and develop the evaluations. As a result, the entire volume is the joint product of the Working Group, and there are no individually authored sections.

IARC Working Groups strive to achieve a consensus evaluation. Consensus reflects broad agreement among Working Group Members, but not necessarily unanimity. The chair may elect to poll Working Group Members to determine the diversity of scientific opinion on issues where consensus is not readily apparent.

After the meeting, the master copy is verified by consulting the original literature, edited and prepared for publication. The aim is to publish the volume within six months of the Working Group meeting. A summary of the outcome is available on the *Monographs* programme website soon after the meeting.

B. SCIENTIFIC REVIEW AND EVALUATION

The available studies are summarized by the Working Group, with particular regard to the qualitative aspects discussed below. In general, numerical findings are indicated as they appear in the original report; units are converted when necessary for easier comparison. The Working Group may conduct additional analyses of the published data and use them in their assessment of the evidence; the results of such supplementary analyses are given in square brackets. When an important aspect of a study that directly impinges

on its interpretation should be brought to the attention of the reader, a Working Group comment is given in square brackets.

The scope of the *IARC Monographs* programme has expanded beyond chemicals to include complex mixtures, occupational exposures, physical and biological agents, lifestyle factors and other potentially carcinogenic exposures. Over time, the structure of a *Monograph* has evolved to include the following sections:

1. Exposure data
2. Studies of cancer in humans
3. Studies of cancer in experimental animals
4. Mechanistic and other relevant data
5. Summary
6. Evaluation and rationale

In addition, a section of General Remarks at the front of the volume discusses the reasons the agents were scheduled for evaluation and some key issues the Working Group encountered during the meeting.

This part of the Preamble discusses the types of evidence considered and summarized in each section of a *Monograph*, followed by the scientific criteria that guide the evaluations.

1. Exposure data

Each *Monograph* includes general information on the agent: this information may vary substantially between agents and must be adapted accordingly. Also included is information on production and use (when appropriate), methods of analysis and detection, occurrence, and sources and routes of human occupational and environmental exposures. Depending on the agent, regulations and guidelines for use may be presented.

(*a*) *General information on the agent*

For chemical agents, sections on chemical and physical data are included: the Chemical Abstracts Service Registry Number, the latest primary name and the IUPAC systematic name are recorded; other synonyms are given, but the list is not necessarily comprehensive. Information on chemical and physical properties that are relevant to identification, occurrence and biological activity is included. A description of technical products of chemicals includes trade names, relevant specifications and available information on composition and impurities. Some of the trade names given may be those of mixtures in which the agent being evaluated is only one of the ingredients.

For biological agents, taxonomy, structure and biology are described, and the degree of variability is indicated. Mode of replication, life cycle, target cells, persistence, latency, host response and clinical disease other than cancer are also presented.

For physical agents that are forms of radiation, energy and range of the radiation are included. For foreign bodies, fibres and respirable particles, size range and relative dimensions are indicated.

For agents such as mixtures, drugs or lifestyle factors, a description of the agent, including its composition, is given.

Whenever appropriate, other information, such as historical perspectives or the description of an industry or habit, may be included.

(*b*) *Analysis and detection*

An overview of methods of analysis and detection of the agent is presented, including their sensitivity, specificity and reproducibility. Methods widely used for regulatory purposes are emphasized. Methods for monitoring human exposure are also given. No critical evaluation or recommendation of any method is meant or implied.

(*c*) *Production and use*

The dates of first synthesis and of first commercial production of a chemical, mixture or other agent are provided when available; for agents that do not occur naturally, this information may allow a reasonable estimate to be made of the date before which no human exposure to the agent could have occurred. The dates of first reported occurrence of an exposure are also provided when available. In addition, methods of synthesis used in past and present commercial production and different methods of production, which may give rise to different impurities, are described.

The countries where companies report production of the agent, and the number of companies in each country, are identified. Available data on production, international trade and uses are obtained for representative regions. It should not, however, be inferred that those areas or nations are necessarily the sole or major sources or users of the agent. Some identified uses may not be current or major applications, and the coverage is not necessarily comprehensive. In the case of drugs, mention of their therapeutic uses does not necessarily represent current practice nor does it imply judgement as to their therapeutic efficacy.

(*d*) *Occurrence and exposure*

Information on the occurrence of an agent in the environment is obtained from data derived from the monitoring and surveillance of levels in occupational environments, air, water, soil, plants, foods and animal and human tissues. When available, data on the generation, persistence and bioaccumulation of the agent are also included. Such data may be available from national databases.

Data that indicate the extent of past and present human exposure, the sources of exposure, the people most likely to be exposed and the factors that contribute to the exposure are reported. Information is presented on the range of human exposure, including occupational and environmental exposures. This includes relevant findings from both developed and developing countries. Some of these data are not distributed widely and may be available from government reports and other sources. In the case of mixtures, industries, occupations or processes, information is given about all agents known to be

present. For processes, industries and occupations, a historical description is also given, noting variations in chemical composition, physical properties and levels of occupational exposure with date and place. For biological agents, the epidemiology of infection is described.

(*e*) *Regulations and guidelines*

Statements concerning regulations and guidelines (e.g. occupational exposure limits, maximal levels permitted in foods and water, pesticide registrations) are included, but they may not reflect the most recent situation, since such limits are continuously reviewed and modified. The absence of information on regulatory status for a country should not be taken to imply that that country does not have regulations with regard to the exposure. For biological agents, legislation and control, including vaccination and therapy, are described.

2. Studies of cancer in humans

This section includes all pertinent epidemiological studies (see Part A, Section 4). Studies of biomarkers are included when they are relevant to an evaluation of carcinogenicity to humans.

(*a*) *Types of study considered*

Several types of epidemiological study contribute to the assessment of carcinogenicity in humans — cohort studies, case–control studies, correlation (or ecological) studies and intervention studies. Rarely, results from randomized trials may be available. Case reports and case series of cancer in humans may also be reviewed.

Cohort and case–control studies 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. Intervention studies may provide strong evidence for making causal inferences, as exemplified by cessation of smoking and the subsequent decrease in risk for lung cancer.

In correlation 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 of the population to the agent under study. In correlation studies, individual exposure is not documented, which renders this kind of study more prone to confounding. In some circumstances, however, correlation studies may be more informative than analytical study designs (see, for example, the *Monograph* on arsenic in drinking-water; IARC, 2004).

In some instances, case reports and case series have provided important information about the carcinogenicity of an agent. These types of study generally arise from a suspicion, based on clinical experience, that the concurrence of two events — that is, a particular exposure and occurrence of a cancer — has happened rather more frequently

than would be expected by chance. Case reports and case series usually lack complete ascertainment of cases in any population, definition or enumeration of the population at risk and estimation of the expected number of cases in the absence of exposure.

The uncertainties that surround the interpretation of case reports, case series and correlation studies make them inadequate, except in rare instances, to form the sole basis for inferring a causal relationship. When taken together with case–control and cohort studies, however, these types of study may add materially to the judgement that a causal relationship exists.

Epidemiological studies of benign neoplasms, presumed preneoplastic lesions and other end-points thought to be relevant to cancer are also reviewed. They may, in some instances, strengthen inferences drawn from studies of cancer itself.

(*b*) *Quality of studies considered*

It is necessary to take into account the possible roles of bias, confounding and chance in the interpretation of epidemiological studies. Bias is the effect of factors in study design or execution that lead erroneously to a stronger or weaker association than in fact exists between an agent and disease. Confounding is a form of bias that occurs when the relationship with disease is made to appear stronger or weaker than it truly is as a result of an association between the apparent causal factor and another factor that is associated with either an increase or decrease in the incidence of the disease. The role of chance is related to biological variability and the influence of sample size on the precision of estimates of effect.

In evaluating the extent to which these factors have been minimized in an individual study, consideration is given to a number of aspects of design and analysis as described in the report of the study. For example, when suspicion of carcinogenicity arises largely from a single small study, careful consideration is given when interpreting subsequent studies that included these data in an enlarged population. Most of these considerations apply equally to case–control, cohort and correlation studies. Lack of clarity of any of these aspects in the reporting of a study can decrease its credibility and the weight given to it in the final evaluation of the exposure.

Firstly, the study population, disease (or diseases) and exposure should have been well defined by the authors. Cases of disease in the study population should have been identified in a way that was independent of the exposure of interest, and exposure should have been assessed in a way that was not related to disease status.

Secondly, the authors should have taken into account — in the study design and analysis — other variables that can influence the risk of disease and may have been related to the exposure of interest. Potential confounding by such variables should have been dealt with either in the design of the study, such as by matching, or in the analysis, by statistical adjustment. In cohort studies, comparisons with local rates of disease may or may not be more appropriate than those with national rates. Internal comparisons of frequency of disease among individuals at different levels of exposure are also desirable

in cohort studies, since they minimize the potential for confounding related to the difference in risk factors between an external reference group and the study population.

Thirdly, the authors should have reported the basic data on which the conclusions are founded, even if sophisticated statistical analyses were employed. At the very least, they should have given the numbers of exposed and unexposed cases and controls in a case–control study and the numbers of cases observed and expected in a cohort study. Further tabulations by time since exposure began and other temporal factors are also important. In a cohort study, data on all cancer sites and all causes of death should have been given, to reveal the possibility of reporting bias. In a case–control study, the effects of investigated factors other than the exposure of interest should have been reported.

Finally, the statistical methods used to obtain estimates of relative risk, absolute rates of cancer, confidence intervals and significance tests, and to adjust for confounding should have been clearly stated by the authors. These methods have been reviewed for case–control studies (Breslow & Day, 1980) and for cohort studies (Breslow & Day, 1987).

(*c*) *Meta-analyses and pooled analyses*

Independent epidemiological studies of the same agent may lead to results that are difficult to interpret. Combined analyses of data from multiple studies are a means of resolving this ambiguity, and well-conducted analyses can be considered. 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, 1998).

The advantages of combined analyses are increased precision due to increased sample size and the opportunity to explore potential confounders, interactions and modifying effects that may explain heterogeneity among studies in more detail. A disadvantage of combined analyses is the possible lack of compatibility of data from various studies due to differences in subject recruitment, procedures of data collection, methods of measurement and effects of unmeasured co-variates that may differ among studies. Despite these limitations, well-conducted combined analyses may provide a firmer basis than individual studies for drawing conclusions about the potential carcinogenicity of agents.

IARC may commission a meta-analysis or pooled analysis that is pertinent to a particular *Monograph* (see Part A, Section 4). Additionally, as a means of gaining insight from the results of multiple individual studies, ad-hoc calculations that combine data from different studies may be conducted by the Working Group during the course of a *Monograph* meeting. The results of such 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, it is important that the same criteria for data quality be applied as those that would be applied to individual studies and to ensure also that sources of heterogeneity between studies be taken into account.

(*d*) *Temporal effects*

Detailed analyses of both relative and absolute risks in relation to temporal variables, such as age at first exposure, time since first exposure, duration of exposure, cumulative exposure, peak exposure (when appropriate) and time since cessation of exposure, are reviewed and summarized when available. Analyses of temporal relationships may be useful in making causal inferences. In addition, such analyses may suggest whether a carcinogen acts early or late in the process of carcinogenesis, although, at best, they allow only indirect inferences about mechanisms of carcinogenesis.

(*e*) *Use of biomarkers in epidemiological studies*

Biomarkers indicate molecular, cellular or other biological changes and are increasingly used in epidemiological studies for various purposes (IARC, 1991; Vainio *et al.*, 1992; Toniolo *et al.*, 1997; Vineis *et al.*, 1999; Buffler *et al.*, 2004). These may include evidence of exposure, of early effects, of cellular, tissue or organism responses, of individual susceptibility or host responses, and inference of a mechanism (see Part B, Section 4b). This is a rapidly evolving field that encompasses developments in genomics, epigenomics and other emerging technologies.

Molecular epidemiological data that identify associations between genetic polymorphisms and interindividual differences in susceptibility to the agent(s) being evaluated may contribute to the identification of carcinogenic hazards to humans. If the polymorphism has been demonstrated experimentally to modify the functional activity of the gene product in a manner that is consistent with increased susceptibility, these data may be useful in making causal inferences. Similarly, molecular epidemiological studies that measure cell functions, enzymes or metabolites that are thought to be the basis of susceptibility may provide evidence that reinforces biological plausibility. It should be noted, however, that when data on genetic susceptibility originate from multiple comparisons that arise from subgroup analyses, this can generate false-positive results and inconsistencies across studies, and such data therefore require careful evaluation. If the known phenotype of a genetic polymorphism can explain the carcinogenic mechanism of the agent being evaluated, data on this phenotype may be useful in making causal inferences.

(*f*) *Criteria for causality*

After the quality of individual epidemiological studies of cancer has been summarized and assessed, a judgement is made concerning the strength of evidence that the agent in question is carcinogenic to humans. In making its judgement, the Working Group considers several criteria for causality (Hill, 1965). A strong association (e.g. a large relative risk) is more likely to indicate causality than a weak association, although it is recognized that estimates of effect of small magnitude do not imply lack of causality and may be important if the disease or exposure is common. Associations that are replicated in several studies of the same design or that use different epidemiological approaches or under

different circumstances of exposure are more likely to represent a causal relationship than isolated observations from single studies. If there are inconsistent results among investigations, possible reasons are sought (such as differences in exposure), and results of studies that are judged to be of high quality are given more weight than those of studies that are judged to be methodologically less sound.

If the risk increases with the exposure, this is considered to be a strong indication of causality, although the absence of a graded response is not necessarily evidence against a causal relationship. 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.

A number of scenarios may increase confidence in a causal relationship. On the one hand, an agent may be specific in causing tumours at one site or of one morphological type. On the other, carcinogenicity may be evident through the causation of multiple tumour types. Temporality, precision of estimates of effect, biological plausibility and coherence of the overall database are considered. Data on biomarkers may be employed in an assessment of the biological plausibility of epidemiological observations.

Although rarely available, results from randomized trials that show different rates of cancer among exposed and unexposed individuals provide particularly strong evidence for causality.

When several epidemiological studies show little or no indication of an association between an exposure and cancer, a judgement may be made that, in the aggregate, they show evidence of lack of carcinogenicity. Such a judgement requires firstly that the studies meet, to a sufficient degree, 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 excluded with reasonable certainty. In addition, all studies that are judged to be methodologically sound should (a) be consistent with an estimate of effect of unity for any observed level of exposure, (b) when considered together, provide a pooled estimate of relative risk that is at or near to unity, and (c) have a narrow confidence interval, due to sufficient population size. Moreover, no individual 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 is important to note that evidence of lack of carcinogenicity obtained from several epidemiological studies can apply only to the type(s) of cancer studied, to the dose levels reported, and to the intervals between first exposure and disease onset observed in these studies. Experience with human cancer indicates that the period from first exposure to the development of clinical cancer is sometimes longer than 20 years; latent periods substantially shorter than 30 years cannot provide evidence for lack of carcinogenicity.

3. Studies of cancer in experimental animals

All known human carcinogens that have been studied adequately for carcinogenicity in experimental animals have produced positive results in one or more animal species

(Wilbourn *et al.*, 1986; Tomatis *et al.*, 1989). For several agents (e.g. aflatoxins, diethylstilbestrol, solar radiation, vinyl chloride), carcinogenicity in experimental animals was established or highly suspected before epidemiological studies confirmed their carcinogenicity in humans (Vainio *et al.*, 1995). Although this association 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) also present a carcinogenic hazard to humans. Accordingly, in the absence of additional scientific information, these agents are considered to pose a carcinogenic hazard to humans. Examples of additional scientific information are data that demonstrate that a given agent causes cancer in animals through a species-specific mechanism that does not operate in humans or data that demonstrate that the mechanism in experimental animals also operates in humans (see Part B, Section 6).

Consideration is given to all available long-term studies of cancer in experimental animals with the agent under review (see Part A, Section 4). In all experimental settings, the nature and extent of impurities or contaminants present in the agent being evaluated are given when available. Animal species, strain (including genetic background where applicable), sex, numbers per group, age at start of treatment, route of exposure, dose levels, duration of exposure, survival and information on tumours (incidence, latency, severity or multiplicity of neoplasms or preneoplastic lesions) are reported. Those studies in experimental animals 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, 2002).

Other studies considered may include: experiments in which the agent was administered in the presence of factors that modify carcinogenic effects (e.g. initiation–promotion studies, co-carcinogenicity studies and studies in genetically modified animals); studies in which the end-point was not cancer but a defined precancerous lesion; experiments on the carcinogenicity of known metabolites and derivatives; and studies of cancer in non-laboratory animals (e.g. livestock and companion animals) exposed to the agent.

For studies of mixtures, consideration is given to the possibility that changes in the physicochemical properties of the individual substances may occur during collection, storage, extraction, concentration and delivery. Another consideration is that chemical and toxicological interactions of components in a mixture may alter dose–response relationships. The relevance to human exposure of the test mixture administered in the animal experiment is also assessed. This may involve consideration of the following aspects of the mixture tested: (i) physical and chemical characteristics, (ii) identified constituents that may indicate the presence of a class of substances and (iii) the results of genetic toxicity and related tests.

The relevance of results obtained with an agent that is analogous (e.g. similar in structure or of a similar virus genus) to that being evaluated is also considered. Such results may provide biological and mechanistic information that is relevant to the

understanding of the process of carcinogenesis in humans and may strengthen the biological plausibility that the agent being evaluated is carcinogenic to humans (see Part B, Section 2f).

(*a*) *Qualitative aspects*

An assessment of carcinogenicity involves several considerations of qualitative importance, including (i) the experimental conditions under which the test was performed, including route, schedule and duration of exposure, species, strain (including genetic background where applicable), sex, age and duration of follow-up; (ii) the consistency of the results, for example, across species and target organ(s); (iii) the spectrum of neoplastic response, from preneoplastic lesions and benign tumours to malignant neoplasms; and (iv) the possible role of modifying factors.

Considerations of importance in the interpretation and evaluation of a particular study include: (i) how clearly the agent was defined and, in the case of mixtures, how adequately the sample characterization was reported; (ii) whether the dose was monitored adequately, particularly in inhalation experiments; (iii) whether the doses, duration of treatment and route of exposure were appropriate; (iv) whether the survival of treated animals was similar to that of controls; (v) whether there were adequate numbers of animals per group; (vi) whether both male and female animals were used; (vii) whether animals were allocated randomly to groups; (viii) whether the duration of observation was adequate; and (ix) whether the data were reported and analysed adequately.

When benign tumours (a) occur together with and originate from the same cell type as malignant tumours in an organ or tissue in a particular study and (b) appear to represent a stage in the progression to malignancy, they are usually combined in the assessment of tumour incidence (Huff *et al.*, 1989). The occurrence of lesions presumed to be preneoplastic may in certain instances aid in assessing the biological plausibility of any neoplastic response observed. If an agent induces only benign neoplasms that appear to be end-points that do not readily undergo transition to malignancy, the agent should nevertheless be suspected of being carcinogenic and requires further investigation.

(*b*) *Quantitative aspects*

The probability that tumours will occur may depend on the species, sex, strain, genetic background and age of the animal, and on the dose, route, timing and duration of the exposure. Evidence of an increased incidence of neoplasms with increasing levels 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, depending on the particular agent under study and the target organ. Mechanisms such as induction of DNA damage or inhibition of repair, altered cell division and cell death rates and changes in intercellular communication are important determinants of dose–response relationships for some carcinogens. Since many chemicals require metabolic activation before being

converted to their reactive intermediates, both metabolic and toxicokinetic aspects are important in determining the dose–response pattern. Saturation of steps such as absorption, activation, inactivation and elimination may produce non-linearity in the dose–response relationship (Hoel *et al.*, 1983; Gart *et al.*, 1986), as could saturation of processes such as DNA repair. The dose–response relationship can also be affected by differences in survival among the treatment groups.

(*c*) *Statistical analyses*

Factors considered include the adequacy of the information given for each treatment group: (i) number of animals studied and number examined histologically, (ii) number of animals with a given tumour type and (iii) length of survival. 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 or not 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. 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 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-Haenzel 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).

Formal statistical methods have been developed to incorporate historical control data into the analysis of data from a given experiment. These methods assign an appropriate weight to historical and concurrent controls on the basis of the extent of between-study and within-study variability: less weight is given to historical controls when they show a high degree of variability, and greater weight when they show little variability. 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, particularly when historical controls show high between-study variability and are, thus, of little relevance to the current experiment. In analysing results for uncommon tumours, however, 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, gender

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).

Although meta-analyses and combined analyses are conducted less frequently for animal experiments than for epidemiological studies due to differences in animal strains, they can be useful aids in interpreting animal data when the experimental protocols are sufficiently similar.

4. Mechanistic and other relevant data

Mechanistic and other relevant data may provide evidence of carcinogenicity and also help in assessing the relevance and importance of findings of cancer in animals and in humans. The nature of the mechanistic and other relevant data depends on the biological activity of the agent being considered. The Working Group considers representative studies to give a concise description of the relevant data and issues that they consider to be important; thus, not every available study is cited. Relevant topics may include toxicokinetics, mechanisms of carcinogenesis, susceptible individuals, populations and life-stages, other relevant data and other adverse effects. When data on biomarkers are informative about the mechanisms of carcinogenesis, they are included in this section.

These topics are not mutually exclusive; thus, the same studies may be discussed in more than one subsection. For example, a mutation in a gene that codes for an enzyme that metabolizes the agent under study could be discussed in the subsections on toxicokinetics, mechanisms and individual susceptibility if it also exists as an inherited polymorphism.

(*a*) *Toxicokinetic data*

Toxicokinetics refers to the absorption, distribution, metabolism and elimination of agents in humans, experimental animals and, where relevant, cellular systems. Examples of kinetic factors that may affect dose–response relationships include uptake, deposition, biopersistence and half-life in tissues, protein binding, metabolic activation and detoxification. Studies that indicate the metabolic fate of the agent in humans and in experimental animals are summarized briefly, and comparisons of data from humans and animals are made when possible. Comparative information on the relationship between exposure and the dose that reaches the target site may be important for the extrapolation of hazards between species and in clarifying the role of in-vitro findings.

(*b*) *Data on mechanisms of carcinogenesis*

To provide focus, the Working Group attempts to identify the possible mechanisms by which the agent may increase the risk of cancer. For each possible mechanism, a representative selection of key data from humans and experimental systems is summarized. Attention is given to gaps in the data and to data that suggests that more than one

mechanism may be operating. The relevance of the mechanism to humans is discussed, in particular, when mechanistic data are derived from experimental model systems. Changes in the affected organs, tissues or cells can be divided into three non-exclusive levels as described below.

(i) *Changes in physiology*

Physiological changes refer to exposure-related modifications to the physiology and/or response of cells, tissues and organs. Examples of potentially adverse physiological changes include mitogenesis, compensatory cell division, escape from apoptosis and/or senescence, presence of inflammation, hyperplasia, metaplasia and/or preneoplasia, angiogenesis, alterations in cellular adhesion, changes in steroidal hormones and changes in immune surveillance.

(ii) *Functional changes at the cellular level*

Functional changes refer to exposure-related alterations in the signalling pathways used by cells to manage critical processes that are related to increased risk for cancer. Examples of functional changes include modified activities of enzymes involved in the metabolism of xenobiotics, alterations in the expression of key genes that regulate DNA repair, alterations in cyclin-dependent kinases that govern cell cycle progression, changes in the patterns of post-translational modifications of proteins, changes in regulatory factors that alter apoptotic rates, changes in the secretion of factors related to the stimulation of DNA replication and transcription and changes in gap–junction-mediated intercellular communication.

(iii) *Changes at the molecular level*

Molecular changes refer to exposure-related changes in key cellular structures at the molecular level, including, in particular, genotoxicity. Examples of molecular changes include formation of DNA adducts and DNA strand breaks, mutations in genes, chromosomal aberrations, aneuploidy and changes in DNA methylation patterns. Greater emphasis is given to irreversible effects.

The use of mechanistic data in the identification of a carcinogenic hazard is specific to the mechanism being addressed and is not readily described for every possible level and mechanism discussed above.

Genotoxicity data are discussed here to illustrate the key issues involved in the evaluation of mechanistic data.

Tests for genetic and related effects are described in view of the relevance of gene mutation and chromosomal aberration/aneuploidy to carcinogenesis (Vainio *et al*., 1992; McGregor *et al*., 1999). The adequacy of the reporting of sample characterization is considered and, when necessary, commented upon; with regard to complex mixtures, such comments are similar to those described for animal carcinogenicity tests. The available data are interpreted critically according to the end-points detected, which may include DNA damage, gene mutation,

sister chromatid exchange, micronucleus formation, chromosomal aberrations and aneuploidy. The concentrations employed are given, and mention is made of whether the use of an exogenous metabolic system *in vitro* affected the test result. These data are listed in tabular form by phylogenetic classification.

Positive results in tests using prokaryotes, lower eukaryotes, insects, plants and cultured mammalian cells suggest that genetic and related effects could occur in mammals. Results from such tests may also give information on the types of genetic effect produced and on the involvement of metabolic activation. Some end-points described are clearly genetic in nature (e.g. gene mutations), while others are associated with genetic effects (e.g. unscheduled DNA synthesis). In-vitro tests for tumour promotion, cell transformation and gap–junction intercellular communication may be sensitive to changes that are not necessarily the result of genetic alterations but that may have specific relevance to the process of carcinogenesis. Critical appraisals of these tests have been published (Montesano *et al*., 1986; McGregor *et al*., 1999).

Genetic or other activity manifest in humans and experimental mammals is regarded to be of greater relevance than that in other organisms. The demonstration that an agent can induce gene and chromosomal mutations in mammals *in vivo* indicates that it may have carcinogenic activity. Negative results in tests for mutagenicity in selected tissues from animals treated *in vivo* provide less weight, partly because they do not exclude the possibility of an effect in tissues other than those examined. Moreover, negative results in short-term tests with genetic end-points cannot be considered to provide evidence that rules out the carcinogenicity of agents that act through other mechanisms (e.g. receptor-mediated effects, cellular toxicity with regenerative cell division, peroxisome proliferation) (Vainio *et al*., 1992). Factors that may give misleading results in short-term tests have been discussed in detail elsewhere (Montesano *et al*., 1986; McGregor *et al*., 1999).

When there is evidence that an agent acts by a specific mechanism that does not involve genotoxicity (e.g. hormonal dysregulation, immune suppression, and formation of calculi and other deposits that cause chronic irritation), that evidence is presented and reviewed critically in the context of rigorous criteria for the operation of that mechanism in carcinogenesis (e.g. Capen *et al.*, 1999).

For biological agents such as viruses, bacteria and parasites, other data relevant to carcinogenicity may include descriptions of the pathology of infection, integration and expression of viruses, and genetic alterations seen in human tumours. Other observations that might comprise cellular and tissue responses to infection, immune response and the presence of tumour markers are also considered.

For physical agents that are forms of radiation, other data relevant to carcinogenicity may include descriptions of damaging effects at the physiological, cellular and molecular level, as for chemical agents, and descriptions of how these effects occur. ‘Physical agents’ may also be considered to comprise foreign bodies, such as surgical implants of

various kinds, and poorly soluble fibres, dusts and particles of various sizes, the pathogenic effects of which are a result of their physical presence in tissues or body cavities. Other relevant data for such materials may include characterization of cellular, tissue and physiological reactions to these materials and descriptions of pathological conditions other than neoplasia with which they may be associated.

(*c*) *Other data relevant to mechanisms*

A description is provided of any structure–activity relationships that may be relevant to an evaluation of the carcinogenicity of an agent, the toxicological implications of the physical and chemical properties, and any other data relevant to the evaluation that are not included elsewhere.

High-output data, such as those derived from gene expression microarrays, and high-throughput data, such as those that result from testing hundreds of agents for a single end-point, pose a unique problem for the use of mechanistic data in the evaluation of a carcinogenic hazard. In the case of high-output data, there is the possibility to overinterpret changes in individual end-points (e.g. changes in expression in one gene) without considering the consistency of that finding in the broader context of the other end-points (e.g. other genes with linked transcriptional control). High-output data can be used in assessing mechanisms, but all end-points measured in a single experiment need to be considered in the proper context. For high-throughput data, where the number of observations far exceeds the number of end-points measured, their utility for identifying common mechanisms across multiple agents is enhanced. These data can be used to identify mechanisms that not only seem plausible, but also have a consistent pattern of carcinogenic response across entire classes of related compounds.

(*d*) *Susceptibility data*

Individuals, populations and life-stages may have greater or lesser susceptibility to an agent, based on toxicokinetics, mechanisms of carcinogenesis and other factors. Examples of host and genetic factors that affect individual susceptibility include sex, genetic polymorphisms of genes involved in the metabolism of the agent under evaluation, differences in metabolic capacity due to life-stage or the presence of disease, differences in DNA repair capacity, competition for or alteration of metabolic capacity by medications or other chemical exposures, pre-existing hormonal imbalance that is exacerbated by a chemical exposure, a suppressed immune system, periods of higher-than-usual tissue growth or regeneration and genetic polymorphisms that lead to differences in behaviour (e.g. addiction). Such data can substantially increase the strength of the evidence from epidemiological data and enhance the linkage of in-vivo and in-vitro laboratory studies to humans.

(*e*) *Data on other adverse effects*

Data on acute, subchronic and chronic adverse effects relevant to the cancer evaluation are summarized. Adverse effects that confirm distribution and biological effects at the sites of tumour development, or alterations in physiology that could lead to tumour development, are emphasized. Effects on reproduction, embryonic and fetal survival and development are summarized briefly. The adequacy of epidemiological studies of reproductive outcome and genetic and related effects in humans is judged by the same criteria as those applied to epidemiological studies of cancer, but fewer details are given.

5. Summary

This section is a summary of data presented in the preceding sections. Summaries can be found on the *Monographs* programme website (http://monographs.iarc.fr).

(*a*) *Exposure data*

Data are summarized, as appropriate, on the basis of elements such as production, use, occurrence and exposure levels in the workplace and environment and measurements in human tissues and body fluids. Quantitative data and time trends are given to compare exposures in different occupations and environmental settings. Exposure to biological agents is described in terms of transmission, prevalence and persistence of infection.

(*b*) *Cancer in humans*

Results of epidemiological studies pertinent to an assessment of human carcinogenicity are summarized. When relevant, case reports and correlation studies are also summarized. The target organ(s) or tissue(s) in which an increase in cancer was observed is identified. Dose–response and other quantitative data may be summarized when available.

(*c*) *Cancer in experimental animals*

Data relevant to an evaluation of carcinogenicity in animals are summarized. For each animal species, study design and route of administration, it is stated whether an increased incidence, reduced latency, or increased severity or multiplicity of neoplasms or preneoplastic lesions were observed, and the tumour sites are indicated. If the agent produced tumours after prenatal exposure or in single-dose experiments, this is also mentioned. Negative findings, inverse relationships, dose–response and other quantitative data are also summarized.

(*d*) *Mechanistic and other relevant data*

Data relevant to the toxicokinetics (absorption, distribution, metabolism, elimination) and the possible mechanism(s) of carcinogenesis (e.g. genetic toxicity, epigenetic effects) are summarized. In addition, information on susceptible individuals, populations and life-

stages is summarized. This section also reports on other toxic effects, including reproductive and developmental effects, as well as additional relevant data that are considered to be important.

6. Evaluation and rationale

Evaluations of the strength of the evidence for carcinogenicity arising from human and experimental animal data are made, using standard terms. The strength of the mechanistic evidence is also characterized.

It is recognized that the criteria for these evaluations, described below, cannot encompass all of the factors that may be relevant to an evaluation of carcinogenicity. In considering all of the relevant scientific data, the Working Group may assign the agent to a higher or lower category than a strict interpretation of these criteria would indicate.

These categories refer only to the strength of the evidence that an exposure is carcinogenic and not to the extent of its carcinogenic activity (potency). A classification may change as new information becomes available.

An evaluation of the degree of evidence is limited to the materials tested, as defined physically, chemically or biologically. When the agents evaluated are considered by the Working Group to be sufficiently closely related, they may be grouped together for the purpose of a single evaluation of the degree of evidence.

(*a*) *Carcinogenicity in humans*

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

***Sufficient evidence of carcinogenicity*:** The Working Group considers that a causal relationship has been established between exposure to the agent and human cancer. That is, a positive relationship has been observed between the exposure and cancer in studies in which chance, bias and confounding could be ruled out with reasonable confidence. A statement that there is *sufficient evidence* is followed by a separate sentence that identifies the target organ(s) or tissue(s) where an increased risk of cancer was observed in humans. Identification of a specific target organ or tissue does not preclude the possibility that the agent may cause cancer at other sites.

***Limited evidence of carcinogenicity*:** A positive association has been observed between exposure to the agent and cancer for which a causal interpretation is considered by the Working Group to be credible, but chance, bias or confounding could not be ruled out with reasonable confidence.

***Inadequate evidence of carcinogenicity*:** The available studies are of insufficient quality, consistency or statistical power to permit a conclusion regarding the presence or absence of a causal association between exposure and cancer, or no data on cancer in humans are available.

***Evidence suggesting lack of carcinogenicity*:** There are several adequate 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 any studied cancer at any observed level of exposure. The results from these studies alone or combined should have narrow confidence intervals with an upper limit close to the null value (e.g. a relative risk of 1.0). Bias and confounding should be ruled out with reasonable confidence, and the studies should have an adequate length of follow-up. A conclusion of *evidence suggesting lack of carcinogenicity* is inevitably limited to the cancer sites, 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.

In some instances, the above categories may be used to classify the degree of evidence related to carcinogenicity in specific organs or tissues.

When the available epidemiological studies pertain to a mixture, process, occupation or industry, the Working Group seeks to identify the specific agent considered most likely to be responsible for any excess risk. The evaluation is focused as narrowly as the available data on exposure and other aspects permit.

(*b*) *Carcinogenicity in experimental animals*

Carcinogenicity in experimental animals can be evaluated using conventional bioassays, bioassays that employ genetically modified animals, and other in-vivo bioassays that focus on one or more of the critical stages of carcinogenesis. In the absence of data from conventional long-term bioassays or from assays with neoplasia as the end-point, consistently positive results in several models that address several stages in the multistage process of carcinogenesis should be considered in evaluating the degree of evidence of carcinogenicity in experimental animals.

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

***Sufficient evidence of carcinogenicity*:** The Working Group considers that a causal relationship has been established between the agent and 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 or under different protocols. An increased incidence of tumours in both sexes of a single species in a well-conducted study, ideally conducted under Good Laboratory Practices, can also provide *sufficient evidence*.

A single study in one species and sex might 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 strong 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, e.g. (a) the evidence of carcinogenicity is restricted to a single experiment; (b) there are unresolved questions regarding the adequacy of the design, conduct or interpretation of the studies; (c) the agent increases the incidence only of benign neoplasms or lesions of uncertain neoplastic potential; or (d) the evidence of carcinogenicity is restricted to studies that demonstrate only promoting activity in a narrow range of tissues or organs.

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

***Evidence suggesting lack of carcinogenicity*:** Adequate studies involving at least two species are available which show that, within the limits of the tests used, the agent is not carcinogenic. A conclusion of *evidence suggesting lack of carcinogenicity* is inevitably limited to the species, tumour sites, age at exposure, and conditions and levels of exposure studied.

(*c*) *Mechanistic and other relevant data*

Mechanistic and other evidence judged to be relevant to an evaluation of carcinogenicity and of sufficient importance to affect the overall evaluation is highlighted. This may include data on preneoplastic lesions, tumour pathology, genetic and related effects, structure–activity relationships, metabolism and toxicokinetics, physicochemical parameters and analogous biological agents.

The strength of the evidence that any carcinogenic effect observed is due to a particular mechanism is evaluated, using terms such as 'weak', 'moderate' or 'strong'. The Working Group then assesses whether that particular mechanism is likely to be operative in humans. The strongest indications that a particular mechanism operates in humans derive from data on humans or biological specimens obtained from exposed humans. The data may be considered to be especially relevant if they show that the agent in question has caused changes in exposed humans that are on the causal pathway to carcinogenesis. Such data may, however, never become available, because it is at least conceivable that certain compounds may be kept from human use solely on the basis of evidence of their toxicity and/or carcinogenicity in experimental systems.

The conclusion that a mechanism operates in experimental animals is strengthened by findings of consistent results in different experimental systems, by the demonstration of biological plausibility and by coherence of the overall database. Strong support can be obtained from studies that challenge the hypothesized mechanism experimentally, by demonstrating that the suppression of key mechanistic processes leads to the suppression of tumour development. The Working Group considers whether multiple mechanisms might contribute to tumour development, whether different mechanisms might operate in different dose ranges, whether separate mechanisms might operate in humans and experimental animals and whether a unique mechanism might operate in a susceptible group.

The possible contribution of alternative mechanisms must be considered before concluding that tumours observed in experimental animals are not relevant to humans. An uneven level of experimental support for different mechanisms may reflect that disproportionate resources have been focused on investigating a favoured mechanism.

For complex exposures, including occupational and industrial exposures, the chemical composition and the potential contribution of carcinogens known to be present are considered by the Working Group in its overall evaluation of human carcinogenicity. The Working Group also determines the extent to which the materials tested in experimental systems are related to those to which humans are exposed.

(*d*) *Overall evaluation*

Finally, the body of evidence is considered as a whole, in order to reach an overall evaluation of the carcinogenicity of the agent to humans.

An evaluation may be made for a group of agents that have been evaluated by the Working Group. In addition, when supporting data indicate that other related agents, for which there is no direct evidence of their capacity to induce cancer in humans or in animals, may also be carcinogenic, a statement describing the rationale for this conclusion is added to the evaluation narrative; an additional evaluation may be made for this broader group of agents if the strength of the evidence warrants it.

The agent is described according to the wording of one of the following categories, and the designated group is given. The categorization of an agent is a matter of scientific judgement that reflects the strength of the evidence derived from studies in humans and in experimental animals and from mechanistic and other relevant data.

Group 1: The agent is *carcinogenic to humans*.

This category is used when there is *sufficient evidence of carcinogenicity* in humans. Exceptionally, an agent may be placed in this category when evidence of carcinogenicity in humans is less than *sufficient* but there is *sufficient evidence of carcinogenicity* in experimental animals and strong evidence in exposed humans that the agent acts through a relevant mechanism of carcinogenicity.

Group 2.

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

Group 2A: The agent is *probably carcinogenic to humans*.

This category is used when there is *limited evidence of carcinogenicity* in humans and *sufficient evidence of carcinogenicity* in experimental animals. In some cases, an agent may be classified in this category when there is *inadequate evidence of carcinogenicity* in humans and *sufficient evidence of carcinogenicity* in experimental animals and strong evidence that the carcinogenesis is mediated by a mechanism that also operates in humans. Exceptionally, an agent may be classified in this category solely on the basis of *limited evidence of carcinogenicity* in humans. An agent may be assigned to this category if it clearly belongs, based on mechanistic considerations, to a class of agents for which one or more members have been classified in Group 1 or Group 2A.

Group 2B: The agent is *possibly carcinogenic to humans*.

This category is used for agents for which there is *limited evidence of carcinogenicity* in humans and less than *sufficient evidence of carcinogenicity* in experimental animals. It may also be used when there is *inadequate evidence of carcinogenicity* in humans but there is *sufficient evidence of carcinogenicity* in experimental animals. In some instances, an agent for which there is *inadequate evidence of carcinogenicity* in humans and less than *sufficient evidence of carcinogenicity* in experimental animals together with supporting evidence from mechanistic and other relevant data may be placed in this group. An agent may be classified in this category solely on the basis of strong evidence from mechanistic and other relevant data.

Group 3: The agent is *not classifiable as to its carcinogenicity to humans*.

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

Exceptionally, agents for which the evidence of carcinogenicity is *inadequate* in humans but *sufficient* in experimental animals may be placed in this category when there is strong evidence that the mechanism of carcinogenicity in experimental animals does not operate in humans.

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

An evaluation in Group 3 is not a determination of non-carcinogenicity or overall safety. It often means that further research is needed, especially when exposures are widespread or the cancer data are consistent with differing interpretations.

Group 4: The agent is *probably not carcinogenic to humans*.

This category is used for agents for which there is *evidence suggesting lack of carcinogenicity* in humans and in experimental animals. In some instances, agents for which there is *inadequate evidence of carcinogenicity* in humans but *evidence suggesting lack of carcinogenicity* in experimental animals, consistently and strongly

supported by a broad range of mechanistic and other relevant data, may be classified in this group.

(*e*) *Rationale*

The reasoning that the Working Group used to reach its evaluation is presented and discussed. This section integrates the major findings from studies of cancer in humans, studies of cancer in experimental animals, and mechanistic and other relevant data. It includes concise statements of the principal line(s) of argument that emerged, the conclusions of the Working Group on the strength of the evidence for each group of studies, citations to indicate which studies were pivotal to these conclusions, and an explanation of the reasoning of the Working Group in weighing data and making evaluations. When there are significant differences of scientific interpretation among Working Group Members, a brief summary of the alternative interpretations is provided, together with their scientific rationale and an indication of the relative degree of support for each alternative.

References

Bieler, G.S. & Williams, R.L. (1993) Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity. *Biometrics*, **49**(3), 793–801

Breslow, N.E. & Day, N.E. (1980) *Statistical Methods in Cancer Research*, Vol. 1, *The Analysis of Case–Control Studies* (IARC Scientific Publications No. 32), Lyon, IARC

Breslow, N.E. & Day, N.E. (1987) *Statistical Methods in Cancer Research*, Vol. 2, *The Design and Analysis of Cohort Studies* (IARC Scientific Publications No. 82), Lyon, IARC

Buffler, P., Rice, J., Baan, R., Bird, M. & Boffetta, P., eds (2004) *Mechanisms of Carcinogenesis: Contributions of Molecular Epidemiology* (IARC Scientific Publications No. 157), Lyon, IARC

Capen, C.C., Dybing, E., Rice, J.M. & Wilbourn, J.D. (1999) *Species Differences in Thyroid, Kidney and Urinary Bladder Carcinogenesis* (IARC Scientific Publications No. 147), Lyon, IARC

Cogliano, V.J., Baan, R.A., Straif, K., Grosse, Y., Secretan, M.B., El Ghissassi, F. & Kleihues, P. (2004) The science and practice of carcinogen identification and evaluation. *Environmental Health Perspect.*, **112**(13), 1269–1274

Cogliano, V., Baan, R., Straif, K., Grosse, Y., Secretan, B., El Ghissassi, F. & Boyle, P. (2005) Transparency in IARC Monographs. *Lancet Oncol.*, **6**(10), 747

Dunson, D.B., Chen, Z. & Harry, J. (2003) A Bayesian approach for joint modeling of cluster size and subunit-specific outcomes. *Biometrics*, **59**(3), 521–30

Fung, K.Y., Krewski, D. & Smythe, R.T. (1996) A comparison of tests for trend with historical controls in carcinogen bioassay. *Can. J. Statist.*, **24**, 431–454

Gart, J.J., Krewski, D., Lee, P.N., Tarone, R.E. & Wahrendorf, J. (1986) *Statistical Methods in Cancer Research*, Vol. 3, *The Design and Analysis of Long-term Animal Experiments* (IARC Scientific Publications No. 79), Lyon, IARC

Greenland, S. (1998) Meta-analysis. In: Rothman, K.J. & Greenland, S., eds, *Modern Epidemiology*, Philadelphia, Lippincott Williams & Wilkins, pp. 643–673

Greim, H., Gelbke, H.-P., Reuter, U., Thielmann, H.W. & Edler, L. (2003) Evaluation of historical control data in carcinogenicity studies. *Hum. exp. Toxicol.*, **22**, 541–549

Haseman, J.K., Huff, J. & Boorman, G.A. (1984) Use of historical control data in carcinogenicity studies in rodents. *Toxicol. Pathol.*, **12**(2), 126–135

Hill, A.B. (1965) The environment and disease: Association or causation? *Proc. R. Soc. Med.*, **58**, 295–300

Hoel, D.G., Kaplan, N.L. & Anderson, M.W. (1983) Implication of nonlinear kinetics on risk estimation in carcinogenesis. *Science*, **219**, 1032–1037

Huff, J.E., Eustis, S.L. & Haseman, J.K. (1989) Occurrence and relevance of chemically induced benign neoplasms in long-term carcinogenicity studies. *Cancer Metastasis Rev.*, **8**, 1–21

IARC (1977) *IARC Monographs Programme on the Evaluation of the Carcinogenic Risk of Chemicals to Humans*. Preamble (IARC intern. tech. Rep. No. 77/002)

IARC (1978) *Chemicals with Sufficient Evidence of Carcinogenicity in Experimental Animals* – IARC Monographs *Volumes 1–17* (IARC intern. tech. Rep. No. 78/003)

IARC (1979) *Criteria to Select Chemicals for* IARC Monographs (IARC intern. tech. Rep. No. 79/003)

IARC (1982) *IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans*, Suppl. 4, *Chemicals, Industrial Processes and Industries Associated with Cancer in Humans* (IARC Monographs, Volumes 1 to 29), Lyon, IARC

IARC (1983) *Approaches to Classifying Chemical Carcinogens According to Mechanism of Action* (IARC intern. tech. Rep. No. 83/001)

IARC (1987) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, Supplement 7, *Overall Evaluations of Carcinogenicity: An Updating of* IARC Monographs *Volumes 1 to 42*, Lyon, IARC

IARC (1988) *Report of an IARC Working Group to Review the Approaches and Processes Used to Evaluate the Carcinogenicity of Mixtures and Groups of Chemicals* (IARC intern. tech. Rep. No. 88/002)

IARC (1991) *A Consensus Report of an IARC Monographs Working Group on the Use of Mechanisms of Carcinogenesis in Risk Identification* (IARC intern. tech. Rep. No. 91/002)

IARC (2004) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, Vol. 84, *Some Drinking-water Disinfectants and Contaminants, including Arsenic*, Lyon, IARC, pp. 39–267

IARC (2005) *Report of the Advisory Group to Recommend Updates to the Preamble to the* IARC Monographs (IARC Int. Rep. No. 05/001)

IARC (2006) *Report of the Advisory Group to Review the Amended Preamble to the* IARC Monographs (IARC Int. Rep. No. 06/001)

McGregor, D.B., Rice, J.M. & Venitt, S., eds (1999) *The Use of Short- and Medium-term Tests for Carcinogens and Data on Genetic Effects in Carcinogenic Hazard Evaluation* (IARC Scientific Publications No. 146), Lyon, IARC

Montesano, R., Bartsch, H., Vainio, H., Wilbourn, J. & Yamasaki, H., eds (1986) *Long-term and Short-term Assays for Carcinogenesis—A Critical Appraisal* (IARC Scientific Publications No. 83), Lyon, IARC

OECD (2002) *Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies* (Series on Testing and Assessment No. 35), Paris, OECD

Peto, R., Pike, M.C., Day, N.E., Gray, R.G., Lee, P.N., Parish, S., Peto, J., Richards, S. & Wahrendorf, J. (1980) Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal experiments. In: *IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans*, Suppl. 2, *Long-term and Short-term Screening Assays for Carcinogens: A Critical Appraisal*, Lyon, IARC, pp. 311–426

Portier, C.J. & Bailer, A.J. (1989) Testing for increased carcinogenicity using a survival-adjusted quantal response test. *Fundam. appl. Toxicol.*, **12**(4), 731–737

Sherman, C.D., Portier, C.J. & Kopp-Schneider, A. (1994) Multistage models of carcinogenesis: An approximation for the size and number distribution of late-stage clones. *Risk Anal.*, **14**(6), 1039–1048

Stewart, B.W. & Kleihues, P., eds (2003) *World Cancer Report*, Lyon, IARC

Tomatis, L., Aitio, A., Wilbourn, J. & Shuker, L. (1989) Human carcinogens so far identified. *Japan. J. Cancer Res.*, **80**, 795–807

Toniolo, P., Boffetta, P., Shuker, D.E.G., Rothman, N., Hulka, B. & Pearce, N., eds (1997) *Application of Biomarkers in Cancer Epidemiology* (IARC Scientific Publications No. 142), Lyon, IARC

Vainio, H., Magee, P., McGregor, D. & McMichael, A., eds (1992) *Mechanisms of Carcinogenesis in Risk Identification* (IARC Scientific Publications No. 116), Lyon, IARC

Vainio, H., Wilbourn, J.D., Sasco, A.J., Partensky, C., Gaudin, N., Heseltine, E. & Eragne, I. (1995) [Identification of human carcinogenic risk in IARC Monographs.] *Bull. Cancer*, **82**, 339–348 (in French)

Vineis, P., Malats, N., Lang, M., d'Errico, A., Caporaso, N., Cuzick, J. & Boffetta, P., eds (1999) *Metabolic Polymorphisms and Susceptibility to Cancer* (IARC Scientific Publications No. 148), Lyon, IARC

Wilbourn, J., Haroun, L., Heseltine, E., Kaldor, J., Partensky, C. & Vainio, H. (1986) Response of experimental animals to human carcinogens: An analysis based upon the *IARC Monographs* Programme. *Carcinogenesis*, **7**, 1853–1863