

Working Procedures for the IARC Handbooks of Tobacco Control

Starting in 2006, the series of International Agency for Research on Cancer (IARC) *Handbooks of Cancer Prevention* added tobacco control as a new area of prevention for their reviews. When appropriate, in addition to cancer, other health outcomes preventable by avoiding tobacco use may be included for evaluation in a *Handbook*.

The text that follows is organised in two principal parts. The first addresses the general scope, objectives and structure of the *Handbooks of Tobacco Control*. The second describes the scientific procedures for evaluating cancer-preventing agents or interventions.

The Working Procedures described herein are largely taken from the *Handbooks of Cancer Prevention* devoted to Chemoprevention and Screening, and from the IARC Monograph Preamble (updated in January 2006).

The term “exposure” appears repeatedly in these procedures, borrowed from the IARC *Monographs* devoted to the evaluation of carcinogenicity. Epidemiological studies conducted to assess the association between exposure to a given hazard and disease outcome are based on the meaning of the term “exposure” implying increased risk to an

undesired health effect. However, in this series of *Handbooks* dedicated to the evaluation of the preventive effects of compounds, biological or pharmaceutical products, behaviours, programmes and interventions, the traditional meaning of the term “exposure” is unfitting. Therefore in several instances the term “intervention”, which lacks a hazardous connotation, is preferred. Examples of interventions with expected benefits in the area of tobacco control are smoking cessation, banning of smoking in public places and taxation on cigarettes. The evaluation of their health effects may be the focus of future *Handbooks*.

Part one: General Principles

General Scope

The prevention and control of cancer are the strategic objectives of the International Agency for Research on Cancer. Cancer prevention may be achieved at the individual level by avoiding cancer-causing agents and at the population level by adopting programmes, legislation and regulations to reduce exposure to cancer-causing agents.

The *Handbooks of Tobacco Control* will evaluate the available

evidence on the role of chemical compounds, biological and pharmaceutical products, behaviours, programmes and interventions in reducing tobacco use and decreasing tobacco-associated morbidity and mortality. The aim of the *Handbook* series is to provide the scientific community, policy-makers and governing bodies of IARC member states as well as of other countries with evidence-based assessments of these interventions at the individual and population levels, with the ultimate goal of assisting in the global implementation of tobacco control provisions within national and international programmes aimed at reducing tobacco-related morbidity and mortality.

Objectives

The objective of the programme is to prepare, and to publish in the form of *Handbooks*, critical reviews and consensus evaluations of evidence on the preventive effect or risk reduction resulting from interventions focusing on tobacco control, with the help of an internationally formed Working Group of experts. The *Handbooks* may also indicate where additional research efforts are needed, specifically when data immediately relevant to an

evaluation are not available. The evaluations in the *Handbooks* are scientific and qualitative judgments of the peer-reviewed published data, conducted during a week-long meeting of peer review and discussions by the Working Group.

Topic for the Handbook

The topic to be evaluated in a *Handbook* is selected approximately twelve months prior to the meeting by the head of the Lifestyle, Environment and Cancer Group after consultation with IARC scientists involved in tobacco research. A *Handbook* may cover a single topic or a group of related topics in the area of Tobacco Control.

Meeting Participants

Soon after the topic of a *Handbook* is chosen, international scientists with relevant expertise are identified by IARC staff, in consultation with other experts. IARC uses literature searches to identify most experts. Each participant serves as an independent scientist and not as a representative of any organisation, government or industry.

Five categories of participants can be present at *Handbook* meetings: Working Group Members, Invited Specialists, Representatives of national and international health agencies, Observers and the IARC Secretariat. Participants in the first two groups generally have pub-

lished significant research related to the topic being reviewed or in tobacco control in particular. Consideration is also given to demographic diversity and balance of area of expertise. All participants are listed, with their addresses and principal affiliations, at the beginning of each *Handbook* volume.

1. The *Working Group* is responsible for the critical reviews and evaluations that are developed during the meeting. The tasks of the Working Group 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 critically evaluate the results of epidemiological, clinical, and other type of studies; (v) to prepare recommendations for research and for public health action; and (vi) if the topic being reviewed so permits, to make an overall evaluation of the evidence of a protective effect or reduced risk associated with the exposure or intervention focus of the evaluation. Working Group members are selected based on knowledge and experience pertinent to the topic evaluated and absence of real or apparent conflicts of interest.

2. *Invited Specialists* are experts who also have critical knowledge and experience but have a real or apparent conflict of interest. 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 the intervention being evaluated. Invited Specialists do not serve as meeting chair or subgroup chair, or participate in the evaluations.

3. *Representatives* of national and international health agencies may attend meetings because their agencies are interested in the topic of a *Handbook*. Representatives do not serve as meeting chair or subgroup chair, draft any part of a *Handbook*, or participate in the evaluations.

4. *Observers* with relevant scientific credentials may be admitted to a meeting by IARC in limited numbers. Priority 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 serve as sources of first-hand information from the meeting to their sponsoring organisations. Observers also can play a valuable role in ensuring that all published

information and scientific perspectives are considered. Observers will not serve as chair or subgroup chair, draft any part of a *Handbook*, or participate in the evaluations. At the meeting, the chair and subgroup chairs may grant Observers the opportunity to speak, generally after they have observed a discussion.

5. The *IARC Secretariat* consists of scientists who have relevant expertise and who are designated by the Agency to attend a meeting. 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.

The WHO Declaration of Interest form is sent to each prospective participant at the first contact, with the preliminary letter presenting the *Handbook* meeting. Before an official 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 topic of the meeting. IARC assesses the declared interests to determine whether there is a conflict that warrants some limitation on participation. Working Group Members are selected based on the absence of real or apparent conflicts of interest. If a real or apparent

conflict of interest is identified, then the expert is asked to attend as an Invited Specialist. The declarations are updated and reviewed again at the opening of the meeting, approximately 8 months later. Interests related to the subject of the meeting are disclosed to the meeting participants and in the published volume (Cogliano *et al.*, 2004).

Data for the Handbooks

The *Handbooks* review all pertinent studies on the intervention to be evaluated. Only those data considered relevant to evaluate the evidence are included and summarized. 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.

With regard to reports of basic scientific research, epidemiological studies and clinical trials, only studies that have been published or accepted for publication in the openly available scientific literature are reviewed. In certain instances, government agency reports that have undergone peer review and are widely available can be considered. Exceptions may be made ad hoc to include unpublished reports that are in their final form and publicly available, if their inclusion is considered pertinent to making an evaluation. Abstracts from scientific meetings and other reports that do not provide sufficient detail upon which to base an assessment of their quality are generally not 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 identified by the Working Group are clearly outlined in square brackets (ie, []). The reasons for not giving further consideration to an individual study are also indicated in square brackets. Important aspects of a study, directly impinging on its interpretation, are brought to the attention of the reader. 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. These analyses and their results are outlined in square brackets or in italics in the *Handbook*.

Working Procedures

(a) Literature to be reviewed

After the topic of the *Handbook* is chosen, pertinent studies are identified by IARC from recognized sources of information such as PubMed and made available to Working Group members and Invited Specialists to prepare the working papers for the meeting. Meeting participants are invited to supplement the IARC literature searches with their own searches. Studies cited in the working papers are available at the time of the meeting.

(b) Chair of the Meeting

The chair of the *Handbook* meeting is identified among leading international experts soon after the topic of a *Handbook* is chosen. The chair will help develop an outline for the *Handbook* early on, participate on conference calls with Working Group members and Invited Specialists in preparing for the meeting, provide early feedback on working papers and chair the meeting.

(c) Working papers

Working papers are due about 6 to 8 months after original contact of invited experts. The first version of the working papers is compiled and formatted by IARC staff about two months prior to the meeting, or as soon as they are received, and made available ahead of time through IARC's Internet to all Working Group members, Invited Specialists and the IARC Secretariat. Reception of working papers ahead of the established deadline is encouraged, as it allows review of their content, facilitating identification of information gaps early enough. When possible or when deemed necessary, some working papers may be discussed early on among experts to expedite the review process to be accomplished during the meeting. A conference call will be scheduled after reception of all working papers and prior to the meeting, with the aim of identifying areas deserving additional work by experts before the meeting.

Acknowledgement of significant contributions to the chapters by colleagues of the invited experts, either at their home institution or elsewhere, can be included in the *Handbook* under an acknowledgement paragraph to be shown following the listing of the meeting participants.

(d) Meeting

The Working Group members meet at IARC for seven to eight days to discuss and finalize the texts of the *Handbook* and to formulate the evaluations. The Working Group members and Invited Specialists are grouped into sub-groups according to their area of expertise. Sub-groups meet during the first three to four days to review in detail the first versions of their working papers, develop a joint subgroup draft, and write summaries. Scheduling of plenary and sub-group time may change from one *Handbook* meeting to another. During the last few days the participants meet in plenary session to review the subgroup working papers, summaries and to develop the consensus evaluations.

(e) Post-Meeting

After the meeting, the draft of the *Handbook* composed during the meeting is verified (by consulting the original literature), edited and prepared for publication by IARC staff. The aim is to publish *Handbooks* within twelve months of the meeting. If applicable, summaries reporting the results of

the evaluation may be available on the IARC website (<http://www.iarc.fr>) soon after the meeting, and a short report may be published in the international literature.

Part two: Scientific Review of the Evidence and Evaluation

1. Scientific Review

The results of the studies reviewed will constitute the evidence forming the foundation of the evaluation. The validity of these studies should be examined critically to determine the weight of the studies contributing to the assessment. This will entail judging the appropriateness of study design, data collection (including adequate description of the intervention and follow-up), data analysis, and ultimately deciding if chance, bias, confounding or lack of statistical power may account for the observed results. The experts will ascertain how the limitations of the studies affect the results and conclusions reported. The criteria that follow apply to epidemiological and clinical studies and therefore may not be as relevant to studies where other quality criteria would be indicated—for example, those assessing the impact of economic policies.

(a) 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 operation of factors in the study design or execution that lead erroneously to a stronger or weaker association than in fact exists between the exposure/intervention being evaluated and the outcome. Confounding is a form of bias that occurs when the association with the 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 taken into account in an individual study, the *Handbook* considers a number of aspects of design and analysis as described in the report of the study.

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

Second, the authors should have taken into account—in the study design and analysis—other variables that can influence the risk of disease or impact of an intervention, and that may have been related to the intervention 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 the disease may or may not be more appropriate than those with national rates. Internal comparisons of disease frequency among individuals at different levels of the intervention are also desirable in cohort studies, since they minimize the potential for confounding related to difference in risk factors between an external reference group and the study population.

Third, the authors should have reported the basic data on which the conclusions are founded, even if sophisticated statistical analyses were employed. 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).

Aspects that are particularly important in evaluating experimental studies are: the selection of participants, the nature and adequacy of the randomisation procedure, evidence that randomisation achieved an adequate balance between groups, the exclusion criteria used before and after randomisation, compliance with the intervention in the intervention group, and 'contamination' with the intervention in the control group. Other considerations are the means by which the end-point was determined and validated, the length and completeness of follow-up of the groups, and the adequacy of the analysis. 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 exposure ceased, will be reviewed and summarized when available.

Independent population-based studies of the same exposure or intervention may lead to ambiguous results. Combined analyses of data from multiple studies may be a means of resolving 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).

The advantages of combined analyses include increased precision due to increased sample size as well as 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, data collection procedures, measurement methods and effects of unmeasured covariates that may differ between studies.

Meta-analyses may be conducted by the Working Group during the course of preparing a *Handbook* and are identified as original calculations by placement of the results in square brackets or in italics. These may be de-novo analyses or updates of previously conducted analyses that incorporate the results from new studies. Whenever possible, however, such analyses are preferably conducted prior to the *Handbook* meeting. Publication of the results of such meta-analyses prior to or concurrently with the *Handbook* meeting is encouraged for purposes of peer review. The same criteria for data quality that would be applied to individual studies must be applied to combined analyses, and such analyses must take into account heterogeneity between studies.

(b) Criteria for causality

After the quality of each study has been summarized and assessed, a judgement is made concerning the strength of evidence that the exposure or intervention in question reduces the risk of disease or is protective for humans. Hill (1965) lists areas for evaluating the strength of epidemiological associations used in the review of human data when assessing carcinogenesis. These criteria, in many instances, will apply to the assessment included in a *Handbook*

- Consistency of observed associations across studies and populations;
- Magnitude of the reported association;
- Temporal relationship between exposure/intervention and change in disease;
- Exposure-response biologic gradient;
- Biological plausibility;
- Coherence of results across other lines of evidence; and
- Analogy present in related exposures and their effects on health.

If the results are inconsistent among investigations, possible reasons (such as differences in level of exposure/intervention) are sought, and results of studies judged to be of high quality are given more weight than those of studies judged to be methodologically less sound.

When several studies show little or no indication of an

association between an intervention and cancer prevention, the judgement may be made that, in the aggregate, they show evidence of lack of effect. The possibility that bias, confounding or misclassification of exposure or outcome that could explain the observed results should be considered and excluded with reasonable certainty.

2. Summary of the data reviewed (evidence)

This section summarizes the results of the evidence presented in the preceding sections in a *Handbook* in a concise manner. Traditionally, this section does not include citation of literature as do preceding sections presenting and discussing the evidence covered in a *Handbook*.

3. Evaluation of the evidence

An evaluation of the strength of the evidence for disease prevention or reduction in morbidity and mortality is made using standard terms. It is conceivable that not every exposure/intervention reviewed in a *Handbook* of tobacco control will permit a formal evaluation of the evidence, as traditionally done in other *Handbooks* of Cancer Prevention and in the Monographs. In evaluating the strength of the evidence, a topic may allow a more formal evaluation (i.e. assigning causality or a protective effect in the prevention of cancer).

If assignment of causality is pertinent and possible, the possible outcomes of an evaluation can include:

Sufficient evidence of a reduction in risk:

The Working Group considers that a causal relationship has been established between the intervention under consideration and a reduction in morbidity and mortality. That is, a relationship has been observed between the exposure/intervention and disease morbidity and mortality in studies in which chance, bias and confounding could be ruled out with reasonable confidence. A statement that there is *sufficient evidence* should be followed by a separate sentence that identifies the types of cancer and other diseases where a decreased morbidity and mortality was observed in humans.

Limited evidence of a reduction in risk:

An association has been observed between the exposure/interven-

tion under consideration and a reduction in disease morbidity and mortality 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 a reduction in risk:

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 the exposure/intervention and a reduced morbidity and mortality. Alternatively, this category is used when no data are available.

Evidence suggesting lack of effect:

There are several adequate studies that are mutually consistent in not showing an association between the exposure/intervention and disease morbidity and mortality. A conclusion of evidence suggesting *lack of risk reduction* is inevitably

limited to the disease sites, conditions and levels of control, and length of observation covered by the available studies.

4. Overall evaluation

The overall evaluation, usually in the form of a narrative, will include a summary of the body of evidence considered as a whole and summary statements made about the strength of the evidence for a health protective or preventive effect, or adverse effects, as appropriate.

5. Recommendations

After reviewing the data and deliberating on them, the Working Group may formulate recommendations, where applicable, for further research and public health action.