International Agency for Research on Cancer

World Health Organization

IARC Handbooks of Cancer Prevention

Report of the Advisory Group to Recommend an Update to the Preamble to the *IARC Handbooks*

Lyon, France
December 2019

Contents

The Advisory Group to Recommend an Update to the Preamble to the *IARC Handbooks* was convened at a time when the *IARC Handbooks* programme is redefining its scope and objectives, as well as when approaches to information gathering and evidence assessment and integration are increasingly challenged.

Substantial revisions to the Preamble were recommended, to reflect changes in (i) the scope and objectives; (ii) the general principles and procedures, namely the transparency of procedures, the policy for disclosure of conflicts of interests, the evolution of systematic review methods, and the role of peer review; (iii) the types of study reviewed; (iv) the use of data from studies in experimental animals and mechanistic data in the overall evaluation of interventions for primary prevention; and (v) the balance of benefits and harms.

INTRODUCTION

Procedures of the Advisory Group meeting

On 12–14 February 2019, IARC convened an Advisory Group to Recommend an Update to the Preamble to the *IARC Handbooks* on primary and secondary prevention. This meeting was announced on 11 May 2018 on the *IARC Handbooks* website (http://handbooks.iarc.fr), along with a Call for Experts. Shortly after the announcement, IARC solicited comments through a Call for Public Comments on the *IARC Handbooks* website and by engaging with Working Group members for recent volumes of the *Handbooks*. On the basis of the comments received, the IARC Secretariat revised the available documents and submitted revised versions to the Advisory Group, together with the comments. Advisory Group members were invited to comments on those documents during a consultation period. At the end of the consultation period, the IARC Secretariat reviewed all comments from the Advisory Group and produced a revised version of the documents for discussion during the Advisory Group meeting. During the three-day meeting, Advisory Group members reviewed and revised the documents, first in Subgroups according to their fields of expertise and then in plenary sessions to achieve consensus on the various topics under discussion. After the meeting, the IARC Secretariat reviewed the documents on the basis of the comments

collected during the meeting. The post-meeting documents went through two rounds of review by the Advisory Group members before final approval.

Advisory Group members were selected on the basis of their expertise in the field (systematic review and normative process; epidemiology of primary prevention; toxicology; and screening), while also considering gender balance and demographic diversity and balance. The meeting was attended by 21 Advisory Group members, 6 Representatives of national and international health agencies, and 1 Observer. IARC and WHO staff with pertinent expertise, including from the *IARC Handbooks* programme and the WHO Guidelines Review Committee, also participated in the meeting.

Outcomes of the Advisory Group meeting

The outcomes of the Advisory Group meeting were the revised Preambles for primary prevention and for screening, and the present report. Annex 1 provides details about the Advisory Group and other meeting participants, and Annex 2 includes the public comments, which are also available directly on the *IARC Handbooks* website (http://handbooks.iarc.fr/docs/HB-WP_Public_comments.pdf).

The recommendations reflect the consensus of the Advisory Group on the topics that were discussed during the three-day meeting; these include specific considerations about whether a particular change should be made to the Preamble, and the rationale for the recommendations.

GENERAL ISSUES

The Advisory Group made recommendations to the *IARC Handbooks* programme on several overarching issues:

- The Working Procedures were renamed the Preamble to emphasize that this document is the product of an international Advisory Group of external expert scientists, as opposed to Working Procedures that were updated in-house.
- The Advisory Group endorsed the description in the Preamble of the broad range of topics reviewed within the *IARC Handbooks* programme.
- The Advisory Group strongly recommended that the programme reflect on its scope and identify a niche, by identifying topics that extend beyond what is done in other agencies.

- The Advisory Group recommended that the *IARC Handbooks* proactively seek collaborations with other agencies, to avoid duplication of documents and missed opportunities to share resources.
- The Advisory Group strongly recommended the development of an analytical framework in the preparation of each volume of the *Handbooks*. Such a framework will make it possible to establish the scope and rationale of each meeting, the focus of the topic, and the inclusion and exclusion criteria for the studies (detailed below).
- The Advisory Group strongly recommended that the programme limit topics for evaluations to those in which change in exposure, i.e. the actual intervention, has been studied, and to consider studies that compare outcomes in exposed and unexposed people only exceptionally.
- The Advisory Group recognized that the *IARC Handbooks* should not embark on evaluations of tertiary prevention at this stage in the development of the programme.
- The first part of the Preamble, General Principles and Procedures, has been expanded to define the key terms underpinning the programme and those used in the Preamble.

Transparency of procedures

The evaluations provided by the *IARC Handbooks* have broad implications. Therefore, it is critical that the processes used to develop the *Handbooks* be fully transparent and rigorous. IARC's continued attention to scientific rigour and full transparency is essential to ensure the credibility of the conclusions.

The Advisory Group noted that many recommendations for updates were aimed at increasing transparency about the processes used by the *IARC Handbooks*. Although IARC Working Groups have always conducted comprehensive reviews of evidence, the many advances in systematic review methods provide a basis for enhancing transparency through more specific guidance to Working Group members. The *IARC Handbooks* programme has embraced these systematic review methods, incorporating them into its procedures. The Advisory Group recommended updating the Preamble by specifying these procedures, for example by detailing the steps of the systematic review process and where the use of expert judgement is required. The Advisory Group also recognized that there is a balance between transparency and specifying methods in the Preamble too rigidly. The Preamble is designed to accommodate flexibility as

scientific methods evolve. The Instructions for Authors, which are updated more frequently than the Preamble is, describe how these methods are operationalized. The Advisory Group's recommendations for updates to the Preamble also confirmed previous commitments to transparency, including disclosure and publication of conflicts of interest and engagement with the public throughout the process.

Conflicts of interest

The Advisory Group recommended that the updates to the Preamble maintain and strengthen IARC's procedures for protecting evaluations from conflicts of interest. The Advisory Group supported IARC's current policy for disclosure of conflicts of interest and noted that publicly posting the names of potential Working Group members approximately 2 months before a meeting provides an opportunity for undisclosed conflicts of interest to be brought to IARC's attention. The Advisory Group reaffirmed the requirement that all Working Group members be free of conflicts of interest, and recommended that the roles of each category of participant be more explicit and transparent. Furthermore, the Advisory Group strengthened and extended rules for Working Group members by requiring that they refrain from consulting and other activities for financial gain (such as serving as an expert witness) that are related to the topic under review, and refrain from using insider information from the meeting until the final volume of the *Handbooks* is published.

The Advisory Group affirmed IARC's commitment to transparency with respect to data sources, but acknowledged that modern systematic review methods require searches for data or studies that are not published in journals. For the different streams of evidence, IARC conducts comprehensive and transparent searches of bibliographic databases. Identified material is included only if there is sufficient information to permit a scientific evaluation of the quality of the methods and results of the studies. The Advisory Group recommended explicitly clarifying the search criteria for pertinent unpublished studies for certain types of agents (e.g. pharmaceuticals), because research has shown that regulatory agencies may have access to relevant data that are not in the scientific literature. IARC provides the opportunity for regulatory authorities and regulated parties to make pertinent unpublished studies publicly available by the date specified in the Call for Data. Consideration of such studies by the Working Group is dependent on

the public availability of sufficient information to permit an independent evaluation of (i) whether there has been selective reporting (e.g. on outcomes, or from a larger set of conducted studies), (ii) study quality (e.g. design, methodology, and reporting of results), and (iii) study results. This update to the Preamble to clarify the search criteria will ensure that all useful data are identified, and will maintain transparency about which data are used in *IARC Handbooks* evaluations.

Selection of topics and future priorities

The Advisory Group recommended including more detailed information on the process of selecting topics, the criteria and procedures for updating existing evaluations, the screening for new evidence, and the decision to conduct an evaluation.

The Advisory Group recognized that for topics that are vast and/or have been reviewed recently, a review of a narrow, specific aspect of the topic may be carried out; for example, after the review of "absence of excess body fatness", a review related specifically to reduction of body fatness may be carried out. Whatever limitations are placed, they need to be described with sufficient specificity.

The Advisory Group discussed the pertinence of the *IARC Handbooks* programme to interventions that have an effect on an intermediate outcome (and not directly to cancer). It was recognized that many studies assessing cancer prevention investigate interventions that have an effect on an intermediate outcome, with the understanding that such an effect would, in turn, lead to a decrease in cancer incidence and/or mortality. The Advisory Group requested that before embarking on the evaluation of an intervention that has an effect on an intermediate outcome, there must be solid evidence for the link (preventive or risk factor) between the intermediate outcome and cancer. The Advisory Group agreed that, in well-defined circumstances, a volume of the *Handbooks* could first evaluate the effect of the intermediate outcome on cancer, followed by an evaluation of the effects of one or several interventions on the intermediate outcome.

The two latter elements are addressed through the construction of an analytical framework at the very early stages of preparation of a volume of the *Handbooks*.

Development of an analytical framework in preparation of a volume of the *Handbooks*

A new evaluation scheme has been developed for the classification of an intervention of primary prevention, to accommodate the framework of either a direct or an indirect pathway from the intervention to cancer prevention.

To this end, two scenarios have been developed. Scenario 1 describes the direct pathway, with only one step from the intervention to cancer prevention. Scenario 2 describes the indirect pathway, with one step from the intervention to an intermediate outcome (Step 1) and one step from the intermediate outcome to cancer (Step 2). In Scenario 2, there will be two independent sets of data. Evidence from Step 2 will likely come from other authoritative sources that have clearly established a causal association between an intermediate outcome and cancer – either a positive causal association between a risk factor and increased cancer incidence or a protective causal association between a preventive factor and reduced cancer incidence.

These scenarios will make it possible to evaluate interventions that do not directly have an impact on cancer, but that have an effect on a factor (a risk factor or a preventive factor) that itself has a known effect on cancer – for instance, in evaluating the impact of tax policies in reducing the prevalence of tobacco smoking, which is a known risk factor for many cancer types. If this link is only suspected but has not yet been established, then the *IARC Handbooks* may undertake the review of the different streams of evidence that report on such an association, but this will not lead to an overall evaluation.

Role of peer review in the *IARC Handbooks*

Scientific rigour, transparency, and peer review are critical to the validity of the *IARC Handbooks* evaluations. The process consists of several cycles of peer review: (i) before the meeting, with the review of each Working Group member's first draft by several Working Group scientists and by scientists in the *IARC Handbooks* programme, and subsequent revision into a revised draft; and (ii) at the meeting, with the review of the meeting drafts by the entire Subgroup, and the review of the plenary drafts by the entire Working Group. The final report is considered to be the product of the entire Working Group, and sections are not ascribed to individual authors.

The Advisory Group emphasized that it is critical to communicate the several steps and level of rigour of the peer-review process in the Preamble, because the peer review is more thorough than that of a manuscript submitted to a scientific journal.

Instructions for Authors

The Preamble refers to the Instructions for Authors, which are provided to Working Group members to guide them in developing the drafts before the Working Group meeting. The Instructions for Authors constitute the documentation for implementing the principles set out in the Preamble. Therefore, they need to be dynamic and to be updated frequently as methodologies evolve. Their importance is amplified by the increased demand for transparency. Consequently, the Advisory Group recommended that attention be given to modification of the Instructions for Authors in line with the adopted Preamble.

Communication and dissemination of the outcome

The *IARC Handbooks* identify interventions and strategies that can prevent cancer. Because such identification is a first step in cancer prevention, the *IARC Handbooks* evaluations are an important international activity that provides information for decision-making to improve public health worldwide. The *IARC Handbooks* are considered authoritative and are used by many stakeholders, such as national health agencies, research scientists, industry, and the general public. These stakeholders use this information in different ways, such as to identify research gaps, to estimate the proportion of cancers that are preventable through a specific intervention, and to develop guidelines and recommendations to limit exposure to a potential carcinogen or to increase preventive measures.

Some evaluations, such as for cervical cancer prevention, affect large populations, whereas others may pose risks only to particular subgroups. IARC has substantial experience in communicating its findings, and also recognizes the challenges of reaching across the diversity of backgrounds of those interested in the *IARC Handbooks* programme and its findings.

The Advisory Group suggested that IARC continue to develop approaches to disseminate the findings of the *IARC Handbooks* beyond the Special Reports in the *New England Journal of Medicine* and posts on

the IARC and *Handbooks* websites. Frequently asked questions with answers should be developed for those topics that are expected to reach a wide audience.

The Advisory Group recommended that the *IARC Handbooks* programme make communication a high priority and an integral part of the programme.

SCIENTIFIC REVIEW

Experimental and observational studies in humans

Epidemiological studies now include the collection of biological samples and the measurement of various biomarkers, and therefore they offer enhanced assessment of exposures and outcomes, and the opportunity to bridge from laboratory findings on mechanisms to human populations. Epidemiological cohort studies are also evolving, to incorporate biobanks and large populations.

The type and quality of the exposure assessment can have an important impact when interpreting epidemiological findings; therefore, these are important considerations when reviewing studies. The Advisory Group noted that exposure assessment is particularly complicated for outcomes with long latency periods, such as cancer, for which detailed information on past exposures is often missing and exposure intensity and timing must be estimated. The Advisory Group recommended that, in the assessment of study quality, Working Groups add an explicit consideration of the quality of the exposure assessment in each study.

Study quality and informativeness

In describing the quality considerations to be evaluated by Working Groups in their review of epidemiological studies, the Advisory Group recommended adding an explicit consideration of the *informativeness* of each study. The informativeness of a study is its ability to show a true association, if there is one, between the agent and cancer, and the lack of an association, if no association exists. Cooper et al. (2016) used the term *sensitivity* to mean substantially the same aspect of a study. *Informativeness* means not only the absence of bias or confounding, and accuracy in effect estimates, but also that the results provide relevant information on the intervention–cancer association. An informative study is one that is likely to detect an association if one actually exists. Considerations include: (i) having a study

population of sufficient size to obtain precise estimates of effect; (ii) sufficient elapsed time from intervention to measurement of outcome for an effect, if present, to be observable; (iii) the presence of an adequate exposure contrast (intensity, frequency, and/or duration); (iv) biologically relevant definitions of exposure (Smith and Kriebel, 2010); and (v) relevant and well-defined time windows for exposure and outcome.

Studies of cancer in experimental animals for primary prevention

As explained above, the analytical framework recommended by the Advisory Group may articulate both direct pathways (i.e. the intervention has a direct effect on cancer outcomes) and indirect pathways (i.e. the intervention has an effect on an intermediate outcome that has an established causal or preventive association with cancer incidence). In this framework, studies in experimental animals may contribute to the direct evidence that the intervention prevents cancer (Scenario 1) or to the evidence that the intermediate outcome prevents or causes cancer (Step 2 of Scenario 2), while evidence for Step 1 comes from studies in humans only. In this framework, studies in experimental animals can inform the cancer-preventive effect of an intervention or an intermediate outcome in humans, in particular when evidence from epidemiological studies is not compelling or is lacking.

The Advisory Group debated about the suitability of animal models for human safety assessment of primary cancer prevention. One important aspect of the question is the high heterogeneity of human tumours compared with the relative homogeneity of tumours in animal models. The Advisory Group recommended that the criteria for evaluating studies in experimental animals be very stringent, given the limitations in extrapolating from animals to humans, to remain on the conservative side and to avoid the hasty implementation of an intervention on the basis of *sufficient evidence* in experimental animals leading to deleterious effects in humans. Nevertheless, the Advisory Group noted that there have been many recent developments in the design and conduct of cancer-preventive studies in experimental animals, which must be considered.

The types of animal models that mimic human cancers that were considered by the Advisory Group to be relevant now include transplantable systems (xenografts and organoid). In addition, the criteria for

consideration by the Working Group to assess the informativeness of individual studies have been expanded on the basis of recent developments in the field, to ensure the highest relevance to humans of the included studies (Lewis et al., 2017). Concomitantly, less emphasis is given to information related to chemical compounds, because of the broader spectrum of intervention and strategies considered for primary prevention in recent volumes.

With respect to the evaluations, the Advisory Group recognized that, similarly to studies in humans, there is no formulaic answer to the number or type of studies and of outcomes necessary to reach an evaluation of *sufficient* or *limited* evidence, because of the large diversity of study designs and types of end-points. The main updates are: (i) the end-points of cancer-related survival and latency now also contribute to the body of evidence; (ii) effects on cancer progression are included; (iii) an evaluation of *sufficient evidence* can be reached with "an appropriate combination of benign and malignant neoplasms"; and (iv) the criteria for reaching *sufficient evidence* and for *evidence suggesting lack of cancer prevention in experimental animals* have been strengthened.

The Advisory Group re-emphasized that the evaluation is the result of expert judgement based on a rigorous peer review of the literature and an assessment of the body of evidence.

Mechanistic evidence for primary prevention

Similarly to studies in experimental animals, mechanistic data can contribute to the overall evaluation either with the direct evidence that the intervention prevents cancer (Scenario 1) or with the evidence that the intermediate outcome prevents or causes cancer (Step 2 of Scenario 2), while evidence for Step 1, that the intervention prevents the intermediate outcome, comes from studies in humans only. In this framework, mechanistic data can inform the cancer-preventive effect of an intervention or an intermediate outcome in humans.

The spectrum and contribution of mechanistic data to the overall evaluation have been completely reconsidered by the Advisory Group.

The Advisory Group recognized that the wide array of possible contributions by mechanistic studies means that outcomes and end-points will vary widely, depending on the types of intervention and the

specific types of cancer examined in each volume of the *Handbooks*. Therefore, the Advisory Group recommended that IARC should rely on the Working Group's expertise for each volume to determine, in consultation with the IARC Secretariat, the areas of research that would be pertinent for that volume.

Possible mechanisms of action have been redefined and sorted into a limited number of major mechanisms, while stressing that the list is not exhaustive. A strong emphasis has been placed on studies in humans, including studies with cancer-relevant biomarkers. Data from experimental models may also be incorporated but are given less weight. Importantly, information on the harms of the preventive intervention may also come from mechanistic studies, and this is reviewed in the section on mechanistic evidence.

The categories of *strong*, *limited*, and *inadequate evidence* will apply to the entire body of evidence from mechanistic data and not to each mechanism separately, and the criteria for each level of evidence have been qualified and parallel those for studies in experimental animals.

Efficacy and effectiveness studies of screening methods

The concepts of efficacy versus effectiveness, and the types of studies evaluated, are now clearly defined in the Preamble. Although studies on efficacy and on effectiveness are reviewed separately, an evaluation of the entire body of evidence is performed.

Assessing the effectiveness of new technologies

The Advisory Group noted that in *IARC Handbooks* Volume 17: Colorectal Cancer Screening, several of the screening methods used did not have evidence of an impact on cancer-specific mortality coming from randomized controlled trials. Referring to the approach proposed for the evaluation of new colorectal cancer screening tests, the Advisory Group indicated that it may not be necessary to require a full randomized controlled trial of comparative efficacy (using cancer-specific mortality as an end-point) of the new method with the old technology, if the new method is based on the same principles as the old technology and targets lesions with the same biology as the old technology, and if the old technology has been shown to reduce cancer incidence or mortality. In such instances, the assessment of a new screening method requires the demonstration of its impact on relevant intermediate screening outcomes, validated in

studies with an experimental design in populations at average risk, supported by programmatic population evaluation in the screening context, addressing not just efficacy but also acceptability and measurement of effectiveness on an intention-to-screen basis.

The availability of evidence on the comparative effectiveness of a new technology with a proven comparator may also be used as a criterion to determine whether a screening method will be evaluated. An example of such an approach was the decision to include in *IARC Handbooks* Volume 17 computed tomography colonography, which is not routinely used in either opportunistic screening or organized screening programmes for colorectal cancer. The updated Preamble now clearly states such criteria. This change is important, because an increasing number of new technologies will become available for which it will not be ethical to run traditional randomized controlled trials.

Comparative effectiveness

The Advisory Group also discussed the possibility of conducting a formal evaluation of the comparative effectiveness of two methods if both have been proven to be effective. This scenario is likely to become increasingly frequent as the number of methods or technologies that have been proven to be effective for the screening of a given type of cancer will increase. For instance, there are already a large range of methods for screening for colorectal cancer and for cervical cancer, several of which have been proven to be effective. Therefore, it will be pertinent, even critical, for the *IARC Handbooks* programme to provide the level of evidence for the comparative effectiveness of two methods that have been proven to be effective. This is now included in the Preamble.

Benefit–harm ratio and cost–effectiveness analyses

The Advisory Group insisted that more consideration be given to the description of harms associated with the intervention, because the intervention is a preventive measure. No intervention can be implemented without a thorough review of the harms and some evidence of a positive benefit–harm ratio. However, the Advisory Group also recognized that it is not within the scope of the *IARC Handbooks*, and is not logistically possible because of resources and expertise, to conduct comprehensive analyses of the benefit–harm ratio, and of cost–effectiveness. Indeed, such analyses take into account setting-specific

information, which IARC has neither the competence nor the resources to do. It is up to each setting to apply the outcome of the *Handbooks* evaluation to the specific conditions and implement strategies or develop recommendations on the basis of the local conditions. The *IARC Handbooks* perform a qualitative evaluation of the cancer-preventive effect of an intervention. The Advisory Group recommended that IARC review a representative set of the most-informative, best-conducted studies on the benefit–harm balance and of cost–effectiveness analyses and summarize the results in the volume of the *Handbooks*. Such information will not feed into the evaluation but is summarized along with the evaluation to aid users in their consideration of action they may take in light of the evaluation.

EVALUATION

Importance of expert judgement

The application of informed judgement by experts is an integral and critical component of the *IARC Handbooks* evaluation process. In particular, evidence from studies of cancer in humans usually derives from observational designs that do not follow the structure of a randomized controlled trial. When evidence from observational studies is evaluated, the idiosyncratic nature of each study needs to be considered, and this invariably involves judgement. Reliance on standardized checklists and formulas would be counterproductive to a thoughtful evaluation of a study's strengths and limitations. The Preamble provides a framework for evaluating the strength of the evidence, and within a Working Group, experts will exercise their judgement as needed. To achieve transparency in its evaluation, the Working Group should lay out clear reasoning for its decisions, describe the role of expert judgement in those decisions, and explain the basis for that judgement.

Maintaining the distinction between Group B1 and Group B2 in primary prevention

In developing the evaluation scheme, the Advisory Group reflected on the validity of maintaining two Group B categories, i.e. Group B1 and Group B2. Such a distinction seemed obsolete if one considers that only *sufficient evidence* in humans will allow a Group A evaluation. However, the removal of the two categories of Group B would not make it possible to distinguish between a situation in which the level of evidence for experimental animals or for mechanistic data is *sufficient* and a situation in which the level of

evidence for those evidence streams is *less than sufficient*. Therefore, both Group B1 and Group B2 were eventually retained in the overall evaluation scheme.

To parallel the classifications of primary prevention, a classification into four groups was also established for the evaluation of screening. The categories define the strength of evidence that screening for a specific cancer with a given method can decrease the incidence of that cancer or mortality related to that cancer.

References

Cooper GS, Lunn RM, Ågerstrand M, Glenn BS, Kraft AD, Luke AM, et al. (2016). Study sensitivity: evaluating the ability to detect effects in systematic reviews of chemical exposures. Environ Int. 92–93:605–10. https://doi.org/10.1016/j.envint.2016.03.017 PMID:27156196

Lewis SJ, Gardner M, Higgins J, Holly JMP, Gaunt TR, Perks CM, et al. (2017). Developing the WCRF International/University of Bristol methodology for identifying and carrying out systematic reviews of mechanisms of exposure-cancer associations. Cancer Epidemiol Biomarkers Prev. 26(11):1667–75. https://doi.org/10.1158/1055-9965.EPI-17-0232 PMID:28978562

Smith T, Kriebel D (2010). A biologic approach to environmental assessment and epidemiology. New York (NY), USA: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195141566.001.0001

ANNEX 1. LIST OF PARTICIPANTS

Advisory Group Members and Invited Specialists serve in their individual capacities as scientists and not as representatives of their government or any organization with which they are affiliated. Affiliations are provided for identification purposes only.

Members

Annie S. Anderson, University of Dundee, United Kingdom
Bruce Armstrong, The University of Western Australia, Australia[1] (Overall Chair)
Frédéric Bost, Institut national de la santé et de la recherche médicale, France
Luisa Camacho, United States Food and Drug Administration, USA
Karen Canfell, Cancer Council New South Wales, Australia[2]
Andrew Chan, Massachusetts General Hospital, USA[3] (Subgroup Chair, Experimental studies of primary prevention (animals and mechanistic))
Steven Clinton, The Ohio State University, USA[4] (remote participation)
Sue Curry, The University of Iowa, USA (Subgroup Chair, General principles)
Markus Follmann, German Cancer Society, Germany
David Forman, University of Leeds, United Kingdom
Christine Friedenreich, Alberta Health Services, Canada
Susan Gapstur, American Cancer Society, USA
David Hunter, University of Oxford, United Kingdom
Stephen Hursting, University of North Carolina at Chapel Hill, USA
Iris Lansdorp-Vogelaar, Erasmus Medical Center, The Netherlands
Michael Leitzmann, University of Regensburg, Germany (Subgroup Chair, Epidemiology of primary prevention)
Sarah Lewis, University of Bristol, United Kingdom
Pamela Marcus, National Cancer Institute, USA
Arn Migowski, National Cancer Institute, Brazil
Linda Rabeneck, Cancer Care Ontario, Canada (Subgroup Chair, Screening)
Rengaswamy Sankaranarayanan, Research Triangle Institute International, India
Nancy Santesso, McMaster University, Canada

Invited Specialist

Carlo Senore, Reference Center for Epidemiology and Cancer Prevention (CPO) in Piemonte, Italy[5]

[1] Bruce Armstrong reports providing expert opinion to Maurice Blackburn Lawyers.
[2] Karen Canfell reports holding intellectual property rights on cancer natural history modelling.
[3] Andrew Chan reports having received personal consultancy fees from Bayer Pharma AG.
[4] Steven Clinton reports that his unit at The Ohio State University benefits from research funding from the National Cattlemen's Beef Association and the Ohio Soybean Council.
[5] Carlo Senore reports that his Unit at the Reference Center for Epidemiology and Cancer Prevention (CPO) received support from Medtronics and EndoChoice.

Representatives of national and international health agencies

Chisato Hamashima, National Cancer Center, Japan
Solveig Hofvind, Cancer Registry of Norway, Norway
JaeKwan Jun, National Cancer Center, Republic of Korea
Siti Zuhrini Kahan, Ministry of Health, Brunei Darussalam
Sok King Ong, Ministry of Health, Brunei Darussalam
Nadia Vilahur Chiaraviglio, European Commission, Italy

Observer

Martin Wiseman, World Cancer Research Fund International, United Kingdom

IARC/WHO Secretariat

Maribel Almonte, Prevention and Implementation Group (Group Head)
Andrea Altieri, Nutritional Methodology and Biostatistics Group *(Rapporteur)*
Iacopo Baussano, Infections and Cancer Epidemiology Group
Lamia Benbrahim-Tallaa, *IARC Monographs* Group (unable to attend)
Andre Carvalho, Screening Group
Ian Cree, WHO Classification of Tumours Group (Acting Head, Section of Evidence Synthesis and Classification)
Laure Dossus, Biomarkers Group
Pietro Ferrari, Nutritional Methodology and Biostatistics Group (Group Head)
Jennifer Girschik, *IARC Monographs* Group
Yann Grosse, *IARC Monographs* Group
Iciar Indave, WHO Classification of Tumours Group *(Rapporteur)*
Béatrice Lauby-Secretan, *IARC Handbooks* Group (Group Head) *(Responsible Officer)*
Karen Müller, Communications Group (Editor)
Susan Norris, Department of Knowledge Translation (HQ/HIS/IER/REK), WHO[6]
Jin Young Park, Prevention and Implementation Group
Mary Schubauer-Berigan, *IARC Monographs* Group *(Rapporteur)*
Kurt Straif, Consultant for the Section of Evidence Synthesis and Classification *(Rapporteur)*

NOTE REGARDING CONFLICTS OF INTERESTS: Each participant first received a preliminary invitation with the request to complete and sign the IARC/WHO Declaration of Interests form, which covers employment and consulting activities, individual and institutional research support, and other financial interests.

Official invitations were extended after careful assessment of any declared interests that might constitute a real or perceived conflict of interest. Pertinent and significant conflicts are disclosed here. Information about other potential conflicts that are not disclosed may be sent to the Head of the *IARC Handbooks* Group at ihb@iarc.fr.

Participants identified as Invited Specialists did not serve as Meeting Chair or Subgroup Chair. The Declarations were updated and reviewed again at the opening of the meeting.

[6] Susan Norris reports holding stocks in several private companies.

NOTE REGARDING OBSERVERS: Each Observer agreed to respect the Guidelines for Observers at *IARC Handbooks* meetings. Observers did not serve as Meeting Chair or Subgroup Chair, or draft or revise any part of the updated Preambles or the Advisory Group Report. They also agreed not to contact participants before or after the meeting, not to lobby them at any time, not to send them written materials, and not to offer them meals or other favours. IARC asked and reminded Advisory Group Members to report any contact or attempt to influence that they may have encountered, either before or during the meeting.

Posted on 11 December 2018, updated on 21 June 2019

Annex 2

Table of contents

1-	Dr David Thomas	Fred Hutchinson Cancer Research Center (retired)
2-	Dr Robert Smith	American Cancer Society
3-	Dr Gary Stoner	Ohio State University
4-	Dr Ludovic Reveiz	Pan American Health Organization
5-	Dr Paul Pinsky	United States National Cancer Institute

Public Comments Form

To Propose an Update to the Working Procedures of the *IARC Handbooks of Cancer Prevention*

Please send completed comment forms to hbworkingproc@iarc.fr.

1. Name and affiliation of commenter

Your name	David B Thomas
Your principal affiliation	Fred Hutchinson Cancer Research Center (retired)
If another party suggested that you submit this nomination, please identify	Click here to enter text.
WHO Declaration of Interests form (to sign and submit via hbworkingproc@iarc.fr)	No conflict of interest

2. Proposed update to the Working Procedures of the *IARC Handbooks of Cancer Prevention*

(for reference, see the current Working Procedures for primary prevention and for screening, full text available as PDF at: http://handbooks.iarc.fr/docs/Handbooks-Working_Procedures-Primary_Prevention.pdf

W O R K I N G P R O C E D U R E S - handbooks.iarc.fr

handbooks.iarc.fr

3 The Working Procedures describe theobjectives and scope of IARC Handbooks of 4 Cancer Prevention programme, the scientific principles and procedures used to develop a 5 Handbook, the types of evidence considered, and scientific criteria that guide the 6 evaluations. The Working Procedures should be consulted when reading aHandbook or 7 summary of evaluations made bythe IARC Handbooks.

and http://handbooks.iarc.fr/docs/Handbooks-Working_Procedures-Screening.pdf, respectively)

Location of text to be updated: Document	As requested, here are my comments on working procedures for IARC Handbooks for Cancer Prevention, Screening. A.4. Meeting Participants (1) In describing the criteria for selecting participants for working groups, it is stated that "consideration is also given to demographic and gender diversity and balance of scientific

	findings and views". I understand and agree that gender and demographic diversity is desirable. However, these should be a secondary consideration after it has been ascertained that the person being considered has the necessary expertise and scientific qualifications. I assume that this is IARC policy, but I have seen instances where this policy may not have been rigorously followed. More importantly, I am concerned by the statement that "scientific …views" are considered. Ideally scientists should be totally objective in their assessment of scientific information and should not have pre-formed views. One should beware of persons who advocate for one position or another. It may indicate that they have conflicts of interests or biases for other reasons that influence their judgment. I suggest that this statement be reconsidered. B.4. Efficacy and Effectiveness I strongly disagree with the distinction made between studies of efficacy and studies of effectiveness. As written, it is implied that randomized trials are the only kinds of studies that assess efficacy, and that all observational studies are studies of effectiveness. Consider first randomized trials. In a trial of screening, individuals are randomized to a screening group or a control group. To avoid biased results, analyses are based on intent to screen. That is, endpoints (mortality or incidence of advanced disease) are compared in those randomized to screening or not, regardless of whether those who were randomized to screening actually were screened (or received all of the screenings in studies of more than one round of screening), and regardless of whether some of those in the control group received some screening. (The results of randomized trials can also be influenced by other factors, such as loss to follow-up.) The ratio of the rates of the endpoint in the two groups (e.g. the mortality rate in the screened group divided by the mortality rate in the control group) is the measure of the success of the intervention. As with all studies, it is a measure of the benefit of the screening under the circumstances under which the study was performed. Is this a measure of efficacy or

effectiveness? The answer is "both."

Now consider the observational studies. First, it must be recognized that a screening modality must be efficacious to be effective. Therefore, any study with results which are considered valid and which are interpreted as showing effectiveness is, by definition, evidence for efficacy. Furthermore, some observational studies, at least theoretically, could come closer to measuring the true magnitude of the beneficial effect attributable to a screening modality (i.e. its inherent efficacy) than some controlled trials because they aim to measure the influence of the screening modality in those actually screened, rather than those randomized to screening as in the trials. For example, in a case control study of screening for breast cancer, women who present with advanced breast cancer would be compared with a sample of women without breast cancer who are selected from the same population from which the cases came, and the past history of screening would be ascertained for all of the women. The odds ratio (as an estimate of the relative risk) of the risk of advanced disease in screened vs. unscreened women would then be calculated. This ratio is closer to an estimate of the true impact of the screening on the risk of advanced disease than would be the results of a trial with some women in the screened group not being screened and some women in the control group being screened. This is not to say that case-control studies are better than trials. They are not. But they do provide evidence for efficacy. A similar argument could be made for observational cohort studies.

In summary, the studies of efficacy and effectiveness should not be categorized by method. The results of all relevant trials and observational studies should be considered as providing evidence for assessing whether a screening modality is efficacious. Also, the closer that a study comes to measuring the influence of screening among those actually screened the more likely it is to be measuring the potential magnitude of the beneficial effect of the screening modality.

	6. Evaluation. If my argument above is accepted, then the evaluation should be made separately for efficacy and effectiveness. Efficacy is a function of the screening modality itself. I would use the categories of sufficient, limited, inadequate, and evidence against for efficacy only. The criteria for each of the categories seem reasonable to me, with one exception. In the criteria for evidence against, I would eliminate the last criterion, i.e. "There is evidence showing that harms overweight (sic) benefits from the specific intervention." A screening modality may be efficacious and still not be useful because of its harms. I would then separately discuss effectiveness, which is a measure of how well the screening modality works in actual practice. I don't see the value of evaluating effectiveness in the same way as evaluating efficacy, because the effectiveness is so dependent on the local circumstances under which a screening program is conducted. A screening modality can be efficacious and be effective in some situations and not effective in others. Instead, I suggest that the specific factors that are thought to enhance and reduce effectiveness be identified and the evidence for their influence on effectiveness be evaluated. This is where such things as acceptability, compliance, and benefits vs. harms are considered. The likely impact of these factors should be discussed qualitatively. Quantifying their impact is probably not practical or useful because the magnitude of the effect is so dependent on local circumstances. It is also probably not practical to evaluate each factor using the same four categories used in the evaluation of efficacy. I hope that you find my comments useful. Thank you for asking for my input. Sincerely,

	David B. Thomas

Public Comments Form

To Propose an Update to the Working Procedures of the *IARC Handbooks of Cancer Prevention*

Please send completed comment forms to hbworkingproc@iarc.fr.

1. Name and affiliation of commenter

Your name	Robert A. Smith, PhD
Your principal affiliation	American Cancer Society
If another party suggested that you submit this nomination, please identify	Click here to enter text.
WHO Declaration of Interests form (to sign and submit via hbworkingproc@iarc.fr)	No interests to declare; Disclosure Form on file

2. Proposed update to the Working Procedures of the *IARC Handbooks of Cancer Prevention*

(for reference, see the current Working Procedures for primary prevention and for screening, full text available as PDF at: http://handbooks.iarc.fr/docs/Handbooks-Working_Procedures-Primary_Prevention.pdf and http://handbooks.iarc.fr/docs/Handbooks-Working_Procedures-Screening.pdf, respectively)

Comment on IARC Handbooks of Cancer Prevention—Screening—Working Procedures

Page 1—Good introduction

Page 2—lines 3 to 7. I think I would express this differently, and I offer this suggestion to put another suggestion into context. First, the text is a little bit pessimistic. If a cancer screening test is effective and generally has met Wilson and Jungner criteria, then offering it to the public can be endorsed. Benefits will have been judged to outweigh harms over the life course. So, I would suggest the following edits:

Screening requires a commitment among healthy individuals to a series of repeated interactions with health-care providers. It is important for the screening program to reinforce the value of screening and to implement best practices to minimize inconvenience, costs, and the potential for screening fatigue. Effective screening requires an ongoing commitment between the public and health-care providers, and a commitment to use public health resources efficiently.

Page 2, line 25: Suggest "demonstrated," rather than "proven"

Page 2, line 26: Suggest In the past, screening with a given procedure….

Note, there are two circumstances here in which a test would move forward without evidence from an RCT: the first is the example of cervical cancer. The promotion of the Pap test was based on the observation that invasive disease was proceeded by the steady progression of cellular changes. Many years later, there is considerable data suggesting that it is a good screening test, and realistically it was too late to do a RCT (which had not yet been applied to address questions in prevention and early detection research), so careful evaluation of national observational data pushed it over the finish line. A happy ending, but everyone agrees that this is not ideal. The second situation, as we discussed at length during the preparation of the CRC Handbook, is the introduction of alternative screening tests after the efficacy of early detection has been demonstrated earlier with a different test. In this case, "earlier is better" has already been demonstrated.

Page 2, line 32: I think we could be more assertive here…"should be avoided" seems too mild. I know this is all narrative and just a lead up to the methodology, but…..

However, uncontrolled interventions in which individuals are exposed to unknown risks to achieve a benefit that has not yet been proven cannot be supported. To avoid the uptake of "wild type" screening, promising early detection interventions should be promptly evaluated in experimental settings to determine if they are efficacious.

Page 3, lines 1-8: For lines 1-2, I'd suggest being more specific about what we mean by different populations, and perhaps provide some examples. I don't think you're referring to subpopulations, such as the different levels of performance by age

Comment [A1]: Thus far, it always requires repeated interactions. I think you could drop "usually."

Deleted: usually

Deleted: between 3 "healthy" individuals and

Deleted: , which can be inconvenient and 4 costly. Furthermore, effective

Deleted: 5

Deleted: and has inherent public health 6 costs.

Deleted: 30

Deleted: and benefits should be 31 avoided.

within a population, i.e., cervical cancer screening being less effective in older vs. younger women if the same sampling tool is used, or the challenge of greater breast density in younger women in breast cancer screening. Lines 2-5, here again, I suggest that the potential for harms and excess costs should be those that remain after rigorous efforts focused on quality assurance. In the U.S., for example, there is a great deal of complaining about harms, but little effort to reduce avoidable hams. Line 5-8, suggest you be more specific about pertinent health services, i.e., diagnosis and treatment services, and finally, line 8, it often is the case that the what is accomplished by a RCT is the avoidance of bias, but as technology and experience improve, the performance in the community setting exceeds (effectiveness) what was achieved in the trial. This certainly is the case with breast. I think the sentence is fine as is, but you could say "achieve or exceed the outcomes…..

Meeting participants seems OK—I'm increasingly concerned about the way conflicts of interest are being approached, specifically that attention to avoiding COIs has created a kind of a "guilty until proven innocent" environment. In journal disclosures, all relationships are described as "your conflicts of interest." I fully agree that there needs to be full disclosure of all interests, but these should not be described as conflict of interests until they are determined to be real or possible conflicts, including the appearance of a COI ("possible COI" or "the appearance of a COI"s preferable to apparent, since aren't "real" and "apparent" the same?). It also should be possible for a person to have an interest or relationship, or there could be a relationship with the individual's institution and not that individual directly, that could be determined to pose no conflict at all. Shall we just call these interests, and then place emphasis on whether they are determined to be either real or constitute the appearance of a potential COI that would cause some doubt about the credibility of the work? Just a thought.

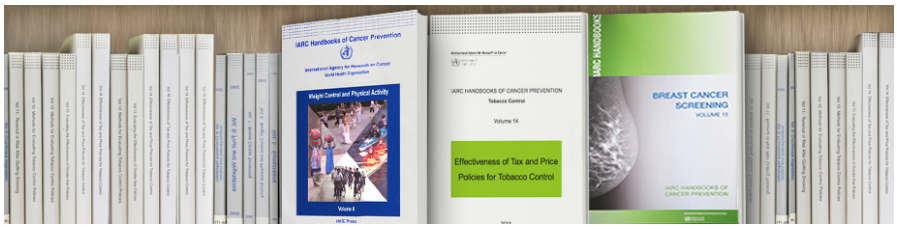
Review and Evaluation Process

Page 6, line 5—I think everyone would agree that the pace of work when the Working Group meets in Lyon is intense and demanding, and towards the end of the 8-day period, that pace intensifies during the critical review and voting period. The two experiences I had (Breast and Colorectal) would have benefited from at least 1 extra day, and perhaps 2. Less stress during one of the most critical periods of the Handbook R&D.

Page 11, line 5—Perhaps it would be informative to note that estimating overdiagnosis, as an adverse outcome of screening, is extremely difficult, since it depends on measuring excess incidence in the context of screening, comparing an exposed group with an unexposed group. Usually, the data are not so cooperative. The possibility that there is overdiagnosis is an important consideration and the evidence can be described, but thus far, we don't have good estimates for the rate of overdiagnosis for any cancer.

Last comment—I expected to see some discussion on the application of new technology in the presence of previously proven technology….will that be something that is the focus of the upgrade in the methodology?

International Agency for Research on Cancer
World Health Organization
IARC Handbooks of Cancer Prevention



WORKING PROCEDURES

A. GENERAL PRINCIPLES AND PROCEDURES

Back to Table of Contents

1. Background

The global burden of cancer is high and continues to increase: the annual number of new cases was estimated at 14.1 million in 2012 and is expected to reach 22.2 million by 2030 (Ferlay et al., 2015). With current trends in demographics and exposure, the cancer burden has been shifting from high-resource countries to low- and medium-resource countries.

1. Seems low

Prevention of cancer is one of the key objectives of the International Agency for Research on Cancer (IARC). Cancer prevention can be achieved by primary prevention - aimed at preventing the occurrence of cancer - or by secondary prevention - aimed at diagnosing cancer sufficiently early to reduce related mortality and suffering.

2. Chemoprevention as another mode of secondary prevention?

Screening and early clinical diagnosis are the principal instruments of secondary prevention of cancer and a fundamental component of any cancer control strategy. Screening may enable detection of cancer sufficiently early that cure and resulting reduction in mortality and suffering from the disease are realistic possibilities given suitable treatment. Screening for some cancers, such as cervical or colorectal cancer, may also detect precancerous lesions, effective treatment of which can prevent occurrence of cancer.

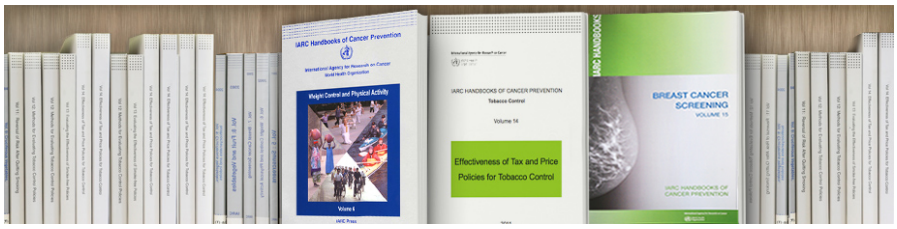
When screening is planned as part of a cancer control programme, only procedures proved to be effective (see below) should be proposed to the general population. Screening usually requires repeated interactions between "healthy" individuals and health-care providers, which can be inconvenient and costly. Furthermore, effective screening requires an ongoing commitment between the public and health-care providers and has inherent public health costs.

Updated 14 November 2017

IARC, 150 Cours Albert Thomas, 69372 Lyon CEDEX 08, France - Tel: +33 (0)4 72 73 84 85

International Agency for Research on Cancer
World Health Organization

IARC Handbooks of Cancer Prevention



WORKING PROCEDURES

B. SCIENTIFIC REVIEW AND EVALUATION

Back to Table of Contents

4. Efficacy and effectiveness of a screening procedure

For the evaluation of both efficacy and effectiveness, the Working Group considers the following general principles in making judgements about the available studies:

- Relevance of the study;
- Appropriateness of the study design and analysis to the question being asked;
- Adequacy and completeness of the presentation of the results;
- Degree to which chance, bias, and confounding may have affected the results.

4.1 Efficacy

In this section, evidence from randomized controlled trial (RCT) studies is reviewed. All aspects of study design and analysis are critically discussed. Indicators of the efficacy of the procedure in terms of mortality or incidence, as well as other relevant indicators, such as the detectable phases of the natural history of the disease, are presented.

Aspects that are particularly important in evaluating RCTs are: the selection of participants, the nature and adequacy of the randomization procedure, evidence that randomization achieved an adequate balance between the groups, exclusion criteria used before and after randomization, compliance with the intervention in the screened group, and â€œcontaminationâ€ of the control. Other considerations include the means by which the outcome (preneoplastic lesions or cancer) was determined and validated (either by screening or by other means of detection of the disease), the length and completeness of follow-up of the groups, and the adequacy of the analysis.

3.

When RCTs are lacking, efficacy cannot be directly evaluated, but only indirectly inferred from observational studies (see below).

4.2 Effectiveness of population-based screening

The impact of the screening procedure when implemented in defined populations is examined in this section.

In this section, mostly observational studies are reviewed, conducted in

o The screening procedure is consistently associated with no reduction in mortality from or incidence of invasive cancer, and chance, bias, and confounding can be ruled out with reasonable confidence.

o There is evidence showing that harms overweight benefits from the specific intervention.

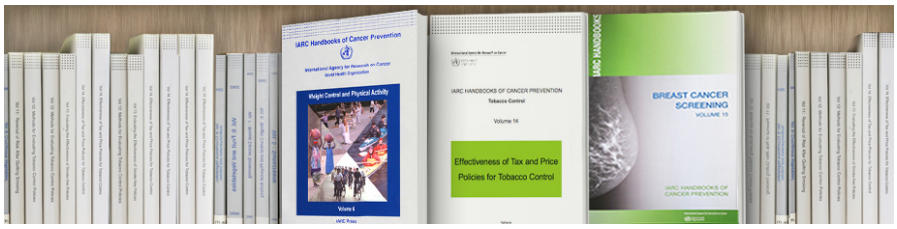
4. ?

Updated 14 November 2017

IARC, 150 Cours Albert Thomas, 69372 Lyon CEDEX 08, France - Tel: +33 (0)4 72 73 84 85

International Agency for Research on Cancer
World Health Organization

IARC Handbooks of Cancer Prevention



WORKING PROCEDURES

A. GENERAL PRINCIPLES AND PROCEDURES

Back to Table of Contents

1. Background

Prevention of cancer is one of the key objectives of the International Agency for Research on Cancer (IARC). The aim of the *IARC Handbooks of Cancer Prevention* series is to review and evaluate scientific information on interventions that may reduce the incidence of or mortality from cancer. As a result of The *Handbooks* evaluations, national and international health agencies have been able, on scientific grounds, to take measures to develop interventions or recommendations that will reduce the risk of developing cancer.

The criteria guiding the evaluations were first established in 1995 at inception of the *IARC Handbooks* series, and were revised in subsequent volumes.

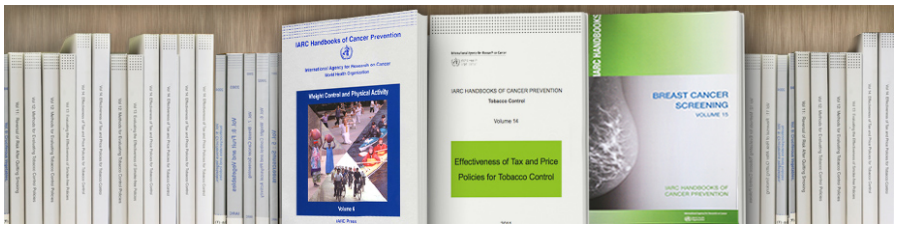
Posted 5 July 2016

5.

IARC, 150 Cours Albert Thomas, 69372 Lyon CEDEX 08, France - Tel: +33 (0)4 72 73 84 85

International Agency for Research on Cancer
World Health Organization

IARC Handbooks of Cancer Prevention



WORKING PROCEDURES

A. GENERAL PRINCIPLES AND PROCEDURES

Back to Table of Contents

2. Objective and scope

The objective of the *IARC Handbooks* programme is the preparation of critical reviews and evaluations of the evidence that a particular intervention can prevent cancer. The evaluations, which are prepared by a Working Group of international experts, are scientific judgements about the available evidence on efficacy, effectiveness, and safety of a wide range of cancer-preventive interventions. No recommendation is given with regard to national or international regulations or legislation, which are the responsibility of individual governments and/or other international authorities. The *IARC Handbooks* may assist national and international authorities in devising programmes of health promotion and cancer prevention, and in making benefit-risk assessments.

In this document, the term "intervention" refers to any chemical, activity, or strategy that is subject to evaluation in a *Handbook*. Cancer-preventive interventions encompass pharmacological, immunological, dietary, and behavioural interventions that may delay, block, or reverse carcinogenic processes, or reduce underlying risk factors.

Preventive interventions can be applied across a continuum of: (1) the general population; (2) subgroups with particular predisposing host or environmental risk factors, including genetic susceptibility to cancer; (3) persons with precancerous lesions; and (4) cancer patients at risk of developing second primary tumours. Use of the same interventions in the treatment of cancer patients to control the growth, metastasis, and recurrence of tumours is considered to be patient management and not prevention, although data from clinical trials of such interventions may be pertinent when reaching anevaluation.

6. Separate these two out

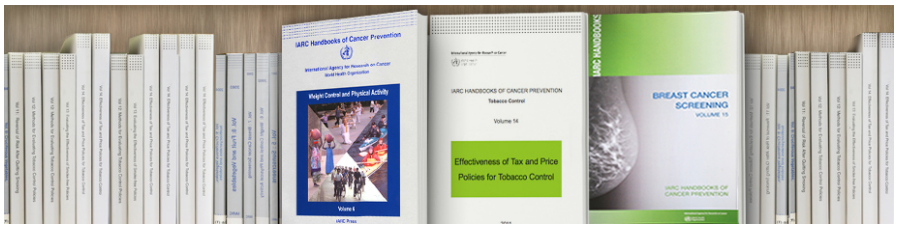
7.

Posted 5 July 2016

IARC, 150 Cours Albert Thomas, 69372 Lyon CEDEX 08, France - Tel: +33 (0)4 72 73 84 85

International Agency for Research on Cancer
World Health Organization

IARC Handbooks of Cancer Prevention



WORKING PROCEDURES

B. SCIENTIFIC REVIEW AND EVALUATION

Back to Table of Contents

2. Studies of cancer prevention in humans

This section includes all pertinent experimental and observational epidemiological studies of cancer prevention in humans, with cancer as an outcome (see Part A, Section 4). Studies of biomarkers as indicators of the intervention are included in Section 4 when they are relevant to an evaluation of the cancer-preventive effect in humans.

(a) Types of study considered

This section focuses on studies that assess the prevention of cancer as an outcome in humans. Relevant evidence is normally provided by experimental studies (for example, randomized clinical trials and community intervention trials), and analytical observational studies, primarily cohort studies and case-control studies. For certain interventions applied at the population level, well-designed ecological studies (studies measuring both outcome and exposure on the aggregate, or population, level) or interrupted time-series studies may also be informative. Cross-sectional studies, descriptive epidemiological studies, case-series and case reports are usually not reviewed. The uncertainties that surround the interpretation of such studies make them inadequate, except in exceptional circumstances, to form the basis for inferring a preventive relationship. However, when considered together with experimental and analytical observational studies, these types of study can sometimes contribute to the decision of the Working Group as to whether or not a causal relationship exists. 8. ?

Intervention studies are experimental in design - that is, the use of, or exposure to, the intervention is assigned by the investigator. Experimental studies can provide the strongest and most direct evidence of a protective or preventive effect; however, the use of such studies is limited for practical and ethical reasons and the subjects are often drawn from select groups that may not represent the population at large.

In exceptional cases, epidemiological studies on advanced preneoplastic lesions and other end-points thought to be relevant to cancer are also reviewed in this section. The results of such studies may strengthen inferences drawn from other studies.

(b) Quality of studies considered

In considering whether a particular study should contribute to the evaluation of an intervention, the Working Group considers the following

suppression of effects that are on the pathway to cancer. The mechanistic evidence can be strengthened by findings of consistent results in different experimental designs, by the demonstration of biological plausibility, and by coherence of the overall database.

The Working Group considers whether multiple mechanisms might contribute to cancer prevention, whether different mechanisms might operate in different dose ranges or at different sites, or whether separate mechanisms might operate in a susceptible group.

For complex interventions, such as food categories, the chemical composition and the potential contribution of different nutrients known to be present may be considered by the Working Group in its overall evaluation of cancer prevention.

(d) Overall evaluation

Finally, the body of evidence is considered as a whole, and summary statements are made that encompass the effects of the intervention with regard to cancer-preventive effects in humans. The overall evaluation is described according to the wording of one of the following standard categories. The categorization of an intervention 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.

(i) The intervention prevents cancer (Group A)

This category is used for interventions for which there is *sufficient evidence* of a cancer-preventive effect in humans.
The sites on which the evidence in humans is based are given.

(ii) The intervention probably prevents cancer (Group B1)

This category is used for interventions for which there is *limited evidence* of a cancer-preventive effect in humans and *sufficient evidence* in animals. An intervention may also be classified in this category when there is *limited evidence* in humans, less than *sufficient evidence* in experimental animals, and strong supporting evidence from mechanistic and other relevant data that the mechanism(s) of prevention also operates in humans.

The sites on which the evidence in humans is based are given.

(iii) The intervention possibly prevents cancer (Group B2)

This category is used for interventions for which there is *inadequate evidence* in humans, and *sufficient evidence* in experimental animals. An intervention may also be classified in this category when there is *inadequate evidence* in humans, *limited evidence* in experimental animals, and strong supporting evidence from mechanistic and other relevant data that the mechanism(s) of prevention also operates in humans.

(iv)The intervention is unclassifiable as to its cancer-preventive effects (Group C)

This category is used for interventions for which the evidence is *inadequate* in humans and less than *sufficient* in experimental animals. Interventions that do not fall into any other group are also placed in this category.

(v) The intervention probably does not prevent cancer (Group D)

This category is used for interventions for which there is *evidence suggesting lack of a cancer-preventive effect* both in humans and in experimental animals.

(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 in humans, studies 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, and an explanation of the reasoning of the Working Group in weighing data and making evaluations. The human populations that were the subject of study should be identified. Additionally, important health concerns identifiedâ€"such as adverse effects, including cancer-causing properties should be clearly addressed. 10.

When there are significant differences in 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.

Posted 5 July 2016

Public Comments Form

To Propose an Update to the Working Procedures of the *IARC Handbooks of Cancer Prevention*

Please send completed comment forms to hbworkingproc@iarc.fr.

1. Name and affiliation of commenter

Your name	Ludovic Reveiz
Your principal affiliation	Pan American Health Organization
If another party suggested that you submit this nomination, please identify	Click here to enter text.
WHO Declaration of Interests form (to sign and submit via hbworkingproc@iarc.fr)	PAHO / WHO staff

2. Proposed update to the Working Procedures of the *IARC Handbooks of Cancer Prevention*

(for reference, see the current Working Procedures for primary prevention and for screening, full text available as PDF at: http://handbooks.iarc.fr/docs/Handbooks-Working_Procedures-Primary_Prevention.pdf and http://handbooks.iarc.fr/docs/Handbooks-Working_Procedures-Screening.pdf, respectively)

Location of text to be updated: Document Section Page number Line number	Line 25 to 28 IARC Handbooks of Cancer Prevention Screening WORKING PROCEDURES **5. Review and evaluation process** 5 18 - 29
Current text	"IARC performs literature searches to compile the relevant bibliography in relation to the topic that will be evaluated. Meeting participants are expected to supplement the IARC literature searches with their own searches of published evidence".
Proposed update (revised text)	"IARC will develop the Handbooks of Cancer - Prevention Screening in accordance with standard procedures set out in the World Health Organization handbook for guideline development [ref]. The quality of the evidence supporting the recommendations will be graded using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) [ref] and the Confidence in the Evidence from Reviews of Qualitative Research (CERQual) approaches [ref]. The GRADE Evidence to Decision Framework (EtD) approach incorporates intervention effects, resources, values, equity, feasibility and acceptability criteria among others".
Brief rationale for update (max. 200 words)	By incorporating the WHO, GRADE and CERQual approaches, other sections should be modified in the SCIENTIFIC REVIEW AND EVALUATION section. Main points for suggesting the modification are: 1. WHO/GRC, GRADE and CERQual are explicit, standardize, comprehensive and widely recognized approaches for developing health evidence-informed recommendations 2. The current method will not necessarily summarize or cite the entire body of literature on the intervention being evaluated (Page 7, Line 6-8). There are no clear criteria for deciding this.

	GRADE provides a comprehensive and standardize approach for decision making. 3. More methodological details are needed in the document (how systematic reviews will be conducted, how the quality of a body of evidence on the effectiveness will be judged and rated and which Strength of the Recommendations Grading System will be used). GRADE has developed an explicit approach for diagnostic tests and strategies that can be applied for the development of the handbooks of cancer prevention - screening. 4. Many countries have incorporated WHO/GRC and GRADE approaches for the development of their national guidelines. Having a similar approach could facilitate the adaptation / adoption of the IARC handbooks of cancer prevention – screening by Member States. PAHO is currently developing and international database of GRADE guidelines (BIGG). Although still under construction, we have identified so far more than 700 GRADE guidelines published in the last five years worldwide (only half are currently available in the database). 5. Using this approach could facilitate subsequent updates by IARC or MS
References, if any (max. 5)	1. WHO handbook for guideline development – 2nd ed https://www.who.int/publications/guidelines/guidelines_review_committee/en/ 2. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. Updated October 2013. https://gdt.gradepro.org/app/handbook/handbook.html 3. Pan American Health Organization 2018. Strengthening national evidence-informed guideline programs. A tool for adapting and implementing guidelines in the Americas ISBN: 978-92-75-12016-3 http://iris.paho.org/xmlui/handle/123456789/49145 ©

Comments

Section A2.
Line 56. The PPV of many recommended screening tests is well below 50%. Therefore, the statement
"Screening tests sort out apparently-well people who probably have a disease from those who probably do not".
is not accurate. It is more accurate to say that screening separates out people who are more likely to have the disease from those who are less likely.

Line 95 – Since screening generally does not "prevent" cancer, suggest changing the wording of this sentence
" … the Handbooks base their evaluation of the strength of the evidence that a putative preventive intervention actually can prevent cancer on the quality and results of all relevant research".
Suggest something like "reduce the burden of cancer" or "prevent cancer incidence or mortality".

Section A3.
Line 105-
It is not clear what is meant by
"The intervention is of putative protective value, but the efficacy, effectiveness or safety have not been established formally"
For example, screening mammography and screening for colorectal cancer were both subjects of handbooks and at the time of the handbook, both had had many trials showing efficacy and were recommended by many entities. What does "established formally" mean?

Section B4.
Line 333 - The term "epidemiologic studies" (in the heading) is often thought to refer only to observational studies, as opposed to randomized trials. Cost-effectiveness is also commonly not considered an "epidemiologic study". Suggest another term, for example, "Experimental and observational studies of each screening method".

Section B4.1
In the discussion of experimental studies, or randomized trials, somewhere the distinction between ITT (intent-to-treat or intent-to-screen) and per-protocol estimates of the RR (or other metric) should be discussed. Some consider ITT estimates to be more relevant for judging effectiveness and per-protocol estimates more relevant for judging efficacy, though not all agree. Any per-protocol estimates must be based on methods that avoid selection bias.
Line 407 –

The sentence “Confidence intervals … that could be produced by chance alone” is confusing. For a point estimate, say RR, whose 95% CI does not cross 1.0, the general conclusion would be that the protective effect is not due to chance; therefore, saying the 95% CI is what could be produced by chance alone is not accurate or at least misleading. Could say the “range of plausible values”.

Section B4.1d
Line 523
It states

“At least, however, (i) a cross-sectional comparison of new test’s accuracy with that of a screening test that has been established to prevent cancer death and (ii) a randomised controlled trial to establish whether, in comparison with the old technology, the new technology can reduce risk of interval cancer should be performed”

The first study (cross-sectional comparison) makes sense. However, the 2nd (trial for interval cancers) is problematic. Interval cancers are generally fairly rare and the sample size for such a trial would frequently have to be enormous and thus not practical. Note that the current TMIST trial (comparing 2D mammography with tomosynthesis) does not have as its endpoint the interval cancer rate, but instead a composite endpoint of advanced disease. There should also be some consideration here of how similar mechanistically the new technology is to the established one (e.g., colonoscopy vs sigmoidoscopy), with less direct evidence needed for the new technology the closer that technology is mechanistically to the established one.

Section B4.2
Line 586 – Interval cancers are not a harm of screening, but rather a failure of the efficacy of screening. That is very different. An interval cancer is one diagnosed by symptoms in a screening setting; in the absence of screening, essentially all cancers would be diagnosed that way. The only way an interval cancer (or a false negative) would be a harm is if knowledge of the prior negative screen delayed diagnostic work-up. That is speculative and to my knowledge, there is no real evidence for it. It is certainly not a major harm of screening, in general.
Line 586 – In addition to harms of “adverse consequences of unnecessary treatment of an overdiagnosed cancer”, there is harms of overdiagnosis per se (being labelled as having cancer and psychological harms of being told one has cancer), as well as harms of surveillance, when treatment is delayed.
Somewhere in this section, the concept of ancillary findings should be discussed. These are a potential harm, and also a potential benefit, e.g., with low-dose CT lung cancer screening or CT colonography.

Section B4.3, 4.4.
It should be emphasized that benefits-to-harm ratios and cost-effectiveness are very dependent on disease prevalence in a population to be screened, in that costs and harms are generally relatively independent of disease prevalence (except for overdiagnosis) while benefits are often directly proportional to prevalence. Often such data (on harms and costs) are from studies conducted in developed countries where the

prevalence of the cancer may be substantially higher than in less developed countries. Therefore, extrapolating benefits-to-harms ratios and cost-effectiveness ratios to other settings must be done taking into account disease prevalence. This is in addition to the fact that costs of medical tests and treatments may clearly differ between these settings, as is pointed out in the current text.

Section B6.1.
For the evaluation for Group A of the benefits and harms and whether the benefits outweigh the harms, it should be noted that this depends on the population setting, and specifically on underlying disease prevalence. Benefits may outweigh harms in high prevalence settings (including where the studies were conducted) but not in lower prevalence settings.