Preface

Amy Berrington de González, David B. Richardson, and Mary K. Schubauer-Berigan

Purpose of the book

Observational epidemiology is used to identify the causes of cancer and other chronic diseases, to determine the effectiveness of interventions, and to understand reasons for differences in disease rates over time or across locations. For more than 50 years, the *IARC Monographs on the Identification of Carcinogenic Hazards to Humans* have led reviews of observational epidemiology, and other evidence, to identify preventable causes of human cancer. In this review process, expert groups in the *IARC Monographs* programme and similar programmes must judge whether a causal interpretation is supported, including whether chance, bias, and confounding can be reasonably ruled out.

Even with the best study design and analysis, it is nearly impossible to eliminate all sources of systematic bias in observational epidemiology; residual confounding, information bias, and selection bias will often remain, despite the researchers' best efforts. For cancer hazard identification, the primary concern when assessing these systematic biases is whether the direction and magnitude of the bias in the central estimate of association could change the interpretation of the result.

The primary purpose of this book is to summarize the wide range of practical methods that can be used by a reader or reviewer of a publication to assess the potential impact of confounding, information bias (including differential and non-differential exposure and outcome misclassification), or selection bias on the results of an epidemiological study. The methods we present can be implemented with information from the publications or external sources, and do not need the original study data. They include indirect approaches, for example negative control outcomes or exposures and proxies, and other approaches, such as sensitivity analyses.

The original volumes in this IARC Statistical Methods in Cancer Research series, by Breslow and Day, summarized the methods available at the time for the design and analysis of case–control studies and cohort studies. Since these works were published, in the 1980s, there have been important developments in both direct and indirect methods for identifying and quantifying biases, and related advances in causal inference. These methods are scattered across the epidemiological and statistical literature or embedded within more technical textbooks. Our goal here is to draw them together and, to quote Breslow and Day, "to place these new tools in the hands of the practising statistician or epidemiologist" (Breslow and Day, 1980, p. 7). To do this, we present them in a way that is accessible to epidemiologists and other research workers who do not have extensive statistical training, as well as to statisticians who do not have epidemiological training. We then illustrate the methods with practical examples, taken from cancer epidemiology, that recur throughout the chapters. We draw on four agents that have previously been evaluated for carcinogenicity in the *IARC Monographs* programme: radiofrequency electromagnetic field (non-ionizing) radiation, consumption of red meat, night shift work, and consumption of opium. These were chosen because they have features that illustrate the range of concepts being explored throughout the book. We provide links to online code or spreadsheets developed by the coauthors or provided by Fox et al. (2021).

Another purpose of the book is to outline the process for integrating these bias assessments into the evidence synthesis. In systematic reviews, such as those undertaken

by the *IARC Monographs* programme, biases are typically first evaluated within an individual study, and then the integration is performed. A growing range of tools is available for the appraisal of biases in systematic reviews (e.g. Grading of Recommendations, Assessment, Development, and Evaluations [GRADE], Risk of Bias in Non-randomized Studies of Interventions [ROBINS-I] or of Exposure [ROBINS-E]). Many take an algorithmic or checklist approach, which emphasizes the presence or absence of bias without regard for its direction or magnitude, and then exclude studies deemed to have the potential for (or risk of) bias (Steenland et al., 2020). A serious limitation of these tools is that they can purge many potentially informative studies, including studies that could help assess biases. On the opposite end of the evidence synthesis spectrum is the goal of reviewing and synthesizing all informative epidemiological studies. The process we outline uses the wide array of methods described in the book to retain all informative studies. This approach is consistent with the review methods described in the Preamble to the *IARC Monographs* (IARC, 2019 and Chapter 1), which calls for Working Groups to integrate studies into evidence synthesis on the basis of their quality and informativeness but recommends against the use of checklists to assess biases and sources of error.

Despite the many developments in the field of bias assessment, in many epidemiological study papers we still find the ubiquitous limitations section that acknowledges the possibility of residual confounding, measurement error, recall bias, or other deficiencies but does not attempt to assess their potential impact on findings. We hope that this book will encourage authors to apply a wider range of direct and indirect bias assessments in their primary research publications. We also refer the reader who is interested in more involved methods, including multidimensional and probabilistic bias analyses, to Fox et al. (2021). Broader adoption of these analyses will enhance the quality of the original papers and further improve the interpretation of the evidence in subsequent reviews.

This IARC Scientific Publication was supported by a 4-day workshop held in October 2022 in Lyon, France, attended by the coauthors. Before the workshop, participants developed initial literature reviews and outlined draft chapters. At the workshop, participants discussed the methods, developed worked examples, and finalized the organization of the material. The draft chapters were reviewed internally and by a group of external peer reviewers.

Definitions of biases in observational cancer epidemiology

Brief descriptions of bias in measures of association are presented next and are then further elaborated within the relevant chapters. Three major sources of systematic bias are recognized in estimates of a measure of association: confounding, information bias, and selection bias. Because the focus here is on hazard identification, rather than risk assessment, it is critical to evaluate the direction of the bias in relation to the direction of the observed effect. Therefore, we have used the terminology of bias towards or away from the null (no effect) to describe the direction wherever possible. In some special circumstances, we may deviate from this, particularly if the direction of the bias is (nearly) always upwards (positive) or downwards (negative), such as for the healthy worker effect.

Confounding

Confounding is bias that arises when the exposure and the outcome of interest share a common cause (Hernán and Robins, 2006). Confounding is a routine concern in observational epidemiology, because of the lack of random assignment to exposure that would ensure that extraneous factors (e.g. other causes of cancer) are randomly distributed among those with different exposure values. To be a confounder, a factor must be related to both exposure (the agent of interest) and outcome (the cancer of interest). Confounding can lead to spurious associations (away from the null, also termed positive confounding) or mask true associations (towards the null, also termed negative confounding). Confounding can be controlled for or minimized in the design or analysis phase of an observational study. This often requires the identification and specification of confounding factors (or confounders) that well represent the source of the potential confounding. For example, confounding by tobacco smoking may be specified in various ways, such as number of years of tobacco smoking, intensity of tobacco smoking, or time since quitting (or any combination thereof). Importantly, the quality of the specification of the confounder can influence the extent to which confounding

is controlled. Control for poorly specified confounders may result in incompletely controlled (or residual) confounding. Conversely, adjusting for a confounder that is on the causal pathway between an exposure and a cancer would have the effect of removing some of the effect of that exposure on the cancer and would give an inaccurate assessment of the true total causal effect.

Information bias

Information bias results from mismeasurement or misclassification of exposure or outcome. The extent of exposure mismeasurement or misclassification can be non-differential or differential with respect to outcome status (e.g. those with cancer can have equally accurate, more accurate, or less accurate exposure measurement or classification, compared with those without cancer). This mismeasurement can be systematic (e.g. always higher than the true value) or random (e.g. sometimes higher and sometimes lower than the true value). An example of differential and systematic exposure measurement error is the recall bias that may be observed in case–control studies, in which participants in the case group may be more likely to recall an exposure than participants in the control group, and the control group would therefore have systematically underestimated exposures. Likewise, outcome misclassification can be non-differential or differential with respect to exposure status, although the latter is less common in most observational epidemiology studies.

Selection bias

Selection bias can occur when entry into or retention in a study is related to both exposure and outcome, for example in a cohort study when exposed individuals are systematically more (or less) likely to be found to be diagnosed with cancer compared with unexposed individuals, or when dropout from a cohort is related to both exposure and outcome status. In a case–control study, selection bias can occur when people with cancer (case participants) are more (or less) likely to agree to take part in a study if they have had an exposure that they think might be related to cancer. Importantly, selection bias requires that study inclusion is related to both exposure and outcome. Study inclusion that is related to only one of these factors does not necessarily lead to selection bias. For example, if the source population giving rise to the study population is more highly exposed than the target (e.g. general) population but inclusion is unrelated to cancer outcomes, then the study might suffer from non-representativeness of the target population, but the estimate will not, in expectation, be biased for the source population. This is often the case with occupationally exposed (source) populations, who may have higher exposures than the general (target) population but whose mechanisms of follow-up would not be different from those of the general population. More information about these concepts is available in Richiardi et al. (2013).

Other bias descriptors

Other terms are used to discuss bias in epidemiological studies, although such terms often relate to problems of confounding, information bias, or selection bias.

Immortal time bias occurs when study participants (e.g. in a cohort study) cannot experience the outcome during some periods of their follow-up after exposure begins. This is usually related to the establishment of a cohort (and, hence, the start of follow-up) after the start of exposure, as might occur in occupational or pharmaco-epidemiological studies. Because immortal time bias occurs in studies that condition on disease status during some period after exposure begins, it is a form of selection bias.

Reverse causation, for example in which diagnosis with disease at time 0 causes a change in exposure status at time 1, typically refers to a type of information bias that arises when subjects are not classified with respect to baseline exposure status. Protopathic bias, which is related to reverse causation, occurs when prediagnostic symptoms of the outcome under study affect the exposure. For example, in Volume 126 of the *IARC Monographs* (IARC, 2021), protopathic bias and reverse causation were of concern for opium consumption and certain cancers, because consumption of opium (a narcotic, antitussive drug) might have been initiated to reduce early symptoms of cancers of the larynx, lung, or oesophagus.

Organization of the book

We have made some pragmatic decisions about the organization of the material in this book. The detailed index should facilitate the location of specific topics and relevant worked examples.

In Chapter 1, we provide background on the *IARC Monographs* programme and its systematic review and evidence synthesis process, as a key example of issues faced by expert review groups. We briefly discuss other major programmes of cancer hazard identification worldwide and their similarities to and differences from the processes of IARC. We also introduce the concept of study informativeness (the ability for a study to correctly identify a real positive association or a real null association) and discuss the related topic of conflicts of interest and how these could affect the potential for study bias or informativeness.

In Chapter 2, we introduce the concept of directed acyclic graphs and describe how they can be useful tools for identifying sources of bias, particularly confounding and selection bias. In Chapters 3, 4, and 5, we summarize methods that can be applied to assess and quantify the three major sources of bias (confounding, information bias, and selection bias). Chapter 5 also covers the miscellaneous topics of immortal time bias, protopathic bias, reverse causation, and considerations when using biomarkers of exposure. We then describe how to integrate these bias assessments into the evidence synthesis process in Chapter 6, and discuss some approaches for the evaluation of multiple biases. We conclude, in Chapter 7, with some recommendations for reporting results and data elements in original study publications that could facilitate bias assessment for future systematic reviewers.

References

Breslow NE, Day NE (1980). Statistical methods in cancer research, Volume I. The analysis of case–control studies (IARC Scientific Publications No. 32). Lyon, France: International Agency for Research on Cancer. Available from: https://publications.iarc.who.int/175.

Fox MP, MacLehose RF, Lash TL (2021). Applying quantitative bias analysis to epidemiologic data. Statistics for biology and health. 2nd ed. Cham, Switzerland: Springer; doi:10.1007/978-3-030-82673-4

Hernán MA, Robins JM (2006). Estimating causal effects from epidemiological data. J Epidemiol Community Health. 60(7):578–86. doi:10.1136/jech.2004.029496 PMID:16790829

IARC (2019). Preamble to the *IARC Monographs* (amended January 2019). Lyon, France: International Agency for Research on Cancer. Available from: https://monographs.iarc.who.int/iarc-monographs-preamble-preamble-to-the-iarc-monographs/.

IARC (2021). Opium consumption. IARC Monogr Identif Carcinog Hazard Hum. 126:1–253. Available from: https://publications.iarc.who.int/600 PMID:36395294

Richiardi L, Pizzi C, Pearce N (2013). Commentary: Representativeness is usually not necessary and often should be avoided. Int J Epidemiol. 42(4):1018–22. doi:10.1093/ije/dyt103 PMID:24062290

Steenland K, Schubauer-Berigan MK, Vermeulen R, Lunn RM, Straif K, Zahm S, et al. (2020). Risk of bias assessments for evidence syntheses of observational epidemiologic studies of environmental and occupational exposures: strengths and limitations. Environ Health Perspect. 128(9):95002. doi:10.1289/EHP6980 PMID:32924579