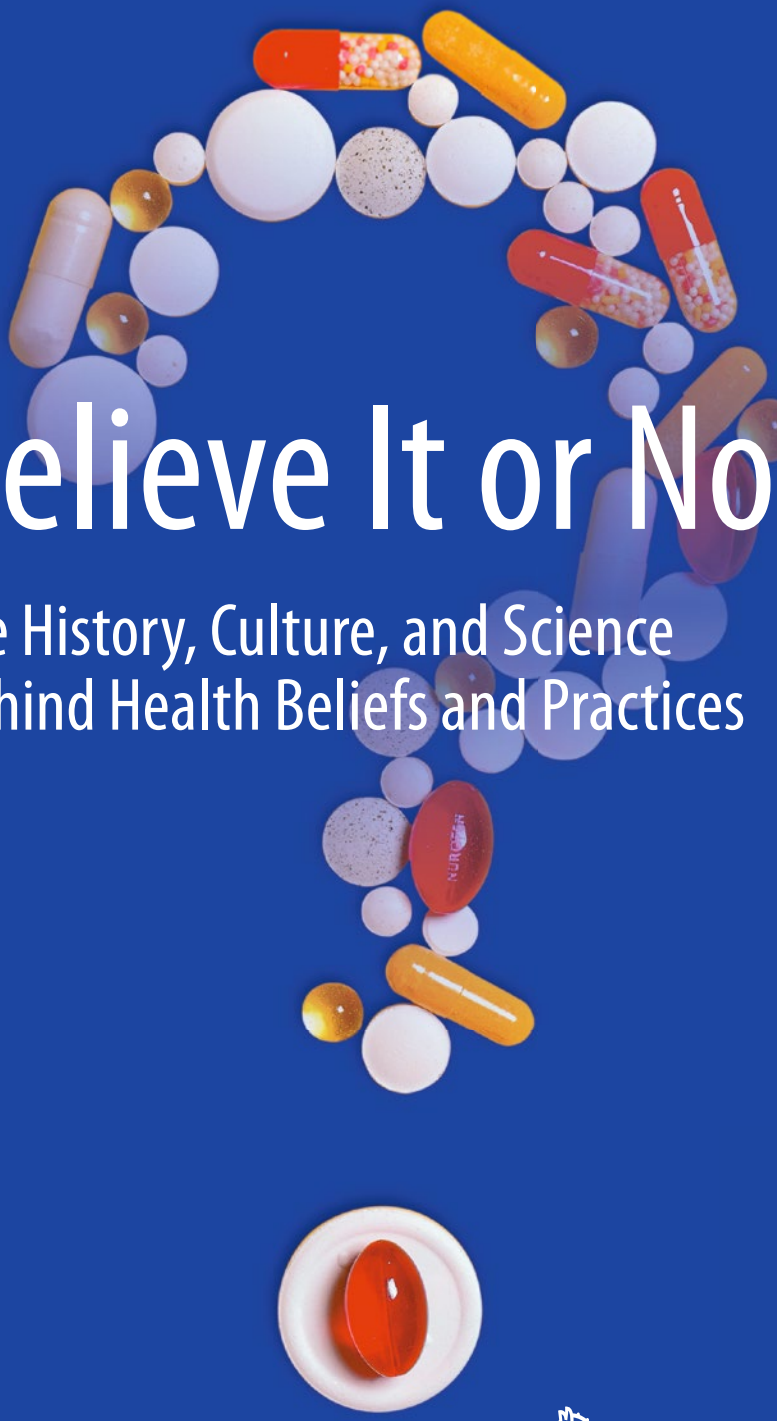
Michael S. Kramer



Believe It or Not

The History, Culture, and Science Behind Health Beliefs and Practices

Springer

Believe It or Not

Michael S. Kramer

Believe It or Not

The History, Culture, and Science Behind Health Beliefs and Practices

Springer

Michael S. Kramer
Faculty of Medicine and Health Sciences
McGill University
Montreal, QC, Canada

ISBN 978-3-031-46021-0 ISBN 978-3-031-46022-7 (eBook)
https://doi.org/10.1007/978-3-031-46022-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

Preface

As a physician, epidemiologist, and devout skeptic, I have been an avid collector of "medical myths" during a professional career spanning more than four decades. I always wanted to research these myths but did not have the time to do them justice while carrying out the tasks for which I earned my salary: research, teaching, clinical work, and administrative responsibilities. The COVID epidemic and my recent retirement from McGill University gave me the opportunity to pursue the project.

Before starting, of course, I needed to find out what others had already written. Every new research project should start with a review of the published literature, and the same is true for a new non-fiction book. I read a dozen or so books on medical myths; most of the books and authors were unknown to me.

What I found convinced me that I was on the right track. For example, several existing books argue that common colds couldn't possibly be caused by exposure to cold air, because colds are caused by respiratory viruses. None considered the possibility of ***joint causation***, that is, that colds might be caused by combined exposure to the cold virus ***and*** cold air. In fact, very few diseases or other health outcomes have a single cause: a factor that is both necessary and sufficient to induce the outcome. Single necessary and sufficient causes exist for a few genetic diseases (like Huntington disease, a rare form of brain degeneration that affects young and middle-aged adults) and for rare infectious agents like Ebola virus. But most genetic and infectious diseases involve an interplay between the genetic defect or infectious agent and other factors involving both the host (the person affected) and the environment. This is even more obviously true for cancer, heart disease, diabetes, and other chronic diseases. It is no longer possible to doubt that cigarette smoking is a cause of lung cancer, while recognizing that most smokers, even heavy smokers, do *not* develop lung cancer, and that occasional cases of lung cancer occur among non-smokers.

The word *myth* can have two quite different meanings: (1) a shared tradition or story, like the Oedipus myth in Greek mythology; and (2) a belief that can be falsified, such as the earth is flat. Most previous books have focused on "busting" the second type of myth: explaining why the myths are untrue, often in a clear and accessible style. I wanted to write something different by exploring the possibility that some longstanding health beliefs might actually be true.

But I am also interested in the first meaning of myth, especially the history and culture surrounding it. When and where did the health belief or practice originate?

What cultural or religious factors led to its origin, its spread to other geographical areas, and its persistence over time? How does the belief or practice vary among countries, and within countries according to age, education, ethnicity, and urban vs rural location? As I discovered in researching the health beliefs and practices I review in this book, some have been around for thousands of years, others only a few decades. For many of them, the historical and cultural influences are as fascinating as the scientific evidence favoring or undermining them.

Finally, previous "myth-busting" books have not attempted rigorous and systematic evaluations of the scientific evidence for and against each of the beliefs and practices they discuss. Some cite published studies, but it is unclear whether the authors came to their conclusions as a ***result*** of the published evidence, or if they selectively sought and cited studies that supported a position to which they had already arrived!

What is required for a scientifically rigorous assessment of any health belief or practice is a *systematic review*, which I define and explain in Chap. 2. To my knowledge, this book is the first to attempt such an assessment. When systematically reviewing the published evidence, it is important to consider the methodological strengths and weaknesses of the studies reviewed. These strengths and weaknesses often have little to do with how recently the study was published, the journal in which it appeared, or the university at which the principal author works.

Carrying out an original systematic review on any one of the health beliefs and practices I discuss would require many months, or even years, of my time and effort. Not only would such an undertaking ignore the excellent work of others who have carried out systematic reviews of the topics I include in the book, but the time required to do that for all the topics would exceed my remaining life expectancy! I will therefore rely heavily on recently published systematic reviews whenever possible. When I am unable to find any, I will acknowledge that fact and cite and assess the best individual studies I can find on the topic, say why I selected them, and explain how they justify the inferences I make.

I can assure you that by the time you read this book, some of its content will be out of date. In the year and a half between the time I began (April 2021) and completed (August 2023) the book, I felt obliged to update my literature search. Although the changes were not earth-shattering, some of the text was substantially revised in the light of new references.

The transience of knowledge should not discourage you, however; it lies at the very heart of science.

How to Read the Book

The first chapter of this book reviews the principles of scientific inference of cause and effect and provides the conceptual basis for evaluating individual research studies. The second chapter moves beyond individual studies to consider the entire body of published research on the topics reviewed: how to sum up the combined evidence based on systematic reviews, where available, of the best research. These first two

chapters are dense and should not be quickly skimmed, except by readers with formal training in epidemiologic and clinical research methods. I have done my best to minimize the technical jargon and to "translate" complicated epidemiologic and statistical concepts in an understandable way for a well-educated general audience.

The subsequent chapters of the book deal with specific health beliefs and practices and can be read in any order, much like an encyclopedia. The topic chapters are divided into sections related to infection, skin and eye conditions, foods and beverages, and pregnancy and childhood. You should feel free to skip around within and between sections, however, to focus on those health beliefs and practices in which you are most interested. Each topic chapter begins with an Introduction that traces the historical and geographical origins of the belief or practice, including the cultural and religious factors that have favored its adoption and persistence. The second section of each topic chapter summarizes the information available on the current prevalence of the beliefs and practices that are the focus of the chapter, that is, ***who believes***. In this section, I prioritize systematic reviews or individual studies of variations in the prevalence of the belief or practice across and within countries and according to age, education, ethnicity, and other factors. In the third section of each topic chapter, I provide a detailed evidence review of the published studies supporting or undermining the belief or practice: in other words, ***why*** you should believe it or not. Readers who read this section may find it helpful to refer back to Chaps. 1 and 2. For those who do not want or need in-depth information on the published studies, I summarize the key evidence points in a textbox placed just before the detailed review. I conclude each topic chapter with a brief summary.

Montreal, QC, Canada Michael S. Kramer

Acknowledgments

I received encouragement and valuable advice from many persons in planning, organizing, writing, and editing this book. These persons include colleagues at McGill University, as well as family and friends. I would particularly like to cite a few of them:

My wife Claire not only supported the need for the book, but also read every word. Although untrained in science, she is an unabashed but constructive critic. She made many insightful suggestions about how to "translate" difficult scientific concepts into language that is accessible to educated lay readers.

My daughter Elise is a practicing optometrist who excels at expressing complex medical concepts and terms that her patients and the general public can understand. She too read every word of the book and suggested major improvements in content and style. She also contributed to the content and interpretation of the chapters on myopia and eye strain.

Contents

About the Author

Michael S. Kramer completed most of his early schooling in Miami, Florida. He left Miami to pursue his undergraduate studies at the University of Chicago, then moved to Yale, where he completed medical school, a residency in pediatrics, and a research fellowship in clinical epidemiology. Following his education and professional training, he moved north to accept a faculty position at the McGill University Faculty of Medicine in Montreal, Canada, where he spent his entire academic career of 42 years and is now Professor Emeritus. He practiced clinical pediatrics for nearly 25 years, but most of his time has been devoted to research and teaching.

Dr Kramer has published over 500 scientific articles and has won numerous national and international awards for his research. He served as a member of expert committees of the World Health Organization (WHO), the U.S. Institute of Medicine, and the Council of Canadian Academies. He helped establish the Canadian Perinatal Surveillance System in 1995 and from 2003 to 2011 was Scientific Director of the Institute of Human Development, Child and Youth Health at the Canadian Institutes of Health Research, Canada's national health research funding agency. In 2011, he was elected to Fellowship in the Royal Society of Canada. Dr Kramer's systematic review of the scientific evidence on the optimal duration of exclusive breastfeeding led directly to new infant feeding recommendations by WHO in 2001. His research on preterm birth helped draw attention to labor induction and elective cesarean delivery as drivers of the rise in preterm birth from the 1980s to the early 2000s. That research contributed to obstetric guidelines to restrict provider-initiated early delivery, which helped reverse the trend. Dr Kramer was recently cited as among the most impactful 0.01% of the world's researchers across all scientific fields. His book written for the general public, entitled *Beyond Parenting Advice: How Science Should Guide Your Decisions on Pregnancy and Child-Rearing*, was published by Springer Nature in late 2021.

Dr Kramer is married and has three children and six grandchildren. He plays violin and is an avid chamber musician. He also enjoys a variety of outdoor activities, including cycling, hiking, tennis, and skiing.

Part I

How to Evaluate Scientific Evidence

How Science Helps Decide What to Believe

1

"The most costly of all follies is to believe passionately in the palpably not true. It is the chief occupation of mankind."

H. L. Mencken

What Is Science?

Before explaining what science is, I will start by discussing what it is not. Science is not technology. Yes, developing new technologies requires scientific training and knowledge. Conversely, many scientific advances benefit from, and may even require, technologic innovation. Technology is a tool that enables good science—not an end in itself, but a means to an end. The Large Hadron Collider (the giant nuclear accelerator located near Geneva, Switzerland) creates high-speed collisions of subatomic particles. But it is scientific hypotheses that lead to the design of specific experiments using the collider, and analysis of the data from those experiments, that lead to new knowledge about the fundamentals of matter.

If you ask school-age children or most adults without formal scientific education to define science, they are likely to mention white coats, laboratory glassware, or high-tech machines. They rarely invoke the testing of hypotheses through carefully designed and conducted experiments or other studies.

If science is not technology, neither is it unquestioned and untested belief in the truth of a proposition. The so-called natural remedies are derived from natural sources and are therefore ***believed*** to be safe. Because of their long history, popularity, and apparent safety, natural remedies can be sold in pharmacies and grocery stores at any price the market will bear. But you are probably unaware that the companies manufacturing natural remedies are not required to demonstrate that they are

M. S. Kramer, *Believe It or Not*, https://doi.org/10.1007/978-3-031-46022-7_1

effective, that is, that they actually work. People who buy these products do so out of faith: the ***belief*** that the products are effective. Because the manufacturers are not legally required to demonstrate efficacy, they don't even try. They have nothing to gain from science and everything to lose.

In contrast, drugs and vaccines cannot be legally marketed in most countries unless they have been approved by national health agencies on the basis of rigorous scientific studies that demonstrate both safety and effectiveness. These rigorous studies are called randomized controlled trials, or RCTs, and I will have much more to say about them later in this chapter. National health agencies do allow the sale of some drugs without evidence of efficacy from RCTs. Such drugs can be purchased "over the counter" without a prescription and were "grandfathered" in after long periods of prior use without major safety concerns. Medicines for the common cold are examples of such drugs.

Beliefs are often based on anecdote. For example, some people are unshakably convinced that their colds are *always* caused by exposure to cold air. Every time they come down with a sneeze and cough, they reflect back on the previous few days (or hours) and recall, "Oh, yeah, I went out on Monday when my hair was still wet" or "My office was freezing cold yesterday." The same reasoning is applied to prevention ("I haven't had a single cold since I started taking vitamin C tablets") and successful recovery ("Every time I have a bad cold, my doctor prescribes antibiotics, and my cold gets better within a few days"). I have devoted chapters to these beliefs about colds in the next section of the book.

These examples demonstrate a very strong cognitive bias: *"post hoc ergo propter hoc"* (after this, therefore because of this), which also known as the post hoc fallacy. But just as the rooster's morning crow doesn't cause the sun to rise, a correct temporal sequence (or, more likely, biased recollection) of events is weak evidence of causality. For example, *any* treatment taken for a cold will appear to be beneficial when it is taken at the peak of symptoms, since down is the only direction possible after a peak! Anecdotes tend to become reinforced by similar episodes that recur, or are selectively recalled, another type of cognitive bias called *confirmation bias*. Eventually, these reinforced beliefs become established in society at large as "folk wisdom."

What about the role of serendipity, a beneficial chance occurrence? Serendipity has enjoyed a rich history in science. But as Louis Pasteur famously said, chance favors the prepared mind. One often-cited medical example of serendipity is Alexander Fleming's discovery of the antibacterial properties of *Penicillium*, a common bread mold that had contaminated one of Fleming's bacteria-containing culture dishes that he had mistakenly left open. Fleming noticed a clear halo (where bacterial growth had been inhibited) surrounding the mold. The serendipitous discovery of penicillin, which is produced by the mold, ushered in the modern era of antibiotic treatment of infections. But observations like Fleming's are not in themselves scientific. They generate hypotheses when, in Pasteur's words, the mind is suitably

prepared. Those hypotheses then lead to experiments and other studies to test the hypotheses—that is, science. When scientific tests convincingly support a hypothesis, it is said to be confirmed ("proven").

Scientific Inference

Not all scientific inferences are cause-and-effect. Some studies have a predictive purpose, such as quantifying the probability of having a fetus affected by Down's syndrome (a birth defect also called trisomy 21, a third copy of the 21st chromosome), based on measurements of various hormones and proteins in the blood during the second trimester of pregnancy. The number of study women, the methods used to recruit them, and their age and other factors will affect the accuracy of the prediction. But no cause-and-effect relationship is inferred. The hormones and proteins measured are not causes of Down's syndrome, but rather, biological markers that help predict its occurrence and thereby help the clinician decide whether or not to recommend a more expensive test based on fetal DNA in the mother's blood or a riskier test like amniocentesis (obtaining and analyzing a sample of amniotic fluid to examine the fetus's chromosomes).

Other scientific inquiries have a descriptive goal. Some population health studies, for example, describe geographic differences or temporal trends in occurrence of health events. Are certain types of cancer more common in some states or provinces than in others? Have a country's hysterectomy rates risen or fallen over time? As mentioned in the Preface to this book, a major section of each of the book's topic chapters is devoted to the current and past prevalence of the beliefs and practices that are the focus of the chapter. That section is primarily based on descriptive studies. No cause-and-effect relationships are inferred from descriptive studies, but they may lead to new causal hypotheses about ***why*** the observed geographic or temporal differences have occurred. Those hypotheses can then be tested in subsequent studies.

It may surprise you to learn that most scientific questions bearing on health beliefs involve causes and consequences. Does eating oily food and chocolate cause acne? Will vaccination increase the risk of infections caused by bacteria or viruses not related to those vaccinated against? Will going outdoors in the winter with wet hair increase your risk of catching a cold?

As shown in Fig. 1.1, such questions have two essential ingredients: a hypothesized cause and a hypothesized effect. In health research, we call these the *exposure* and *outcome*, respectively. The hypothesis is that the exposure causes a change in the outcome. The process of causal inference is thus: formulate a hypothesis about an exposure and its effect on outcome, design a study to test that hypothesis, analyze and interpret the data that result from the study, and infer the validity of—that is, confirm or refute—the hypothesis.

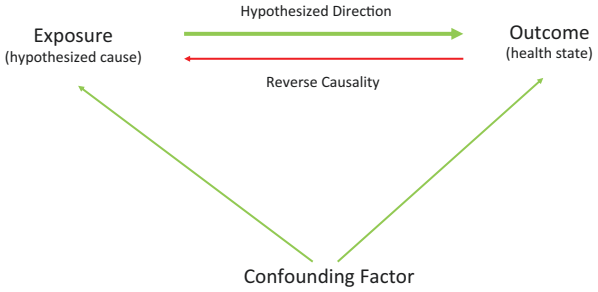


Fig. 1.1 The essentials of causal inference. The study exposure is the hypothesized cause of the outcome, and the outcome is the health state on which an effect of exposure is hypothesized. Arrows point from causes to effects. The direction of an arrow also denotes temporal sequence; the tail occurs earlier in time than the head. Green arrows denote known or hypothesized causal directions, while the red arrow from outcome to exposure denotes reverse causality: the study outcome precedes and causes the exposure. A confounding factor is an underlying (antecedent) cause of both the exposure and outcome and biases the apparent effect of exposure on outcome. It needs to be adjusted ("controlled") for to remove the bias

Experiments Versus Observational Studies

It is important to distinguish two broad types of studies bearing on human health. The first type is called an *experiment*. An experiment means that the researcher actively intervenes to change the exposure and then observes the outcome in the study participants. In health research, the intervention is often a treatment intended to improve the study participant's health, either by preventing an illness or lessening its impact—sometimes even curing it. The outcome is the health state: an illness or some measure of discomfort or disability due to the illness. A *controlled experiment* is a study in which two treatments are compared, or an active treatment is compared to an inactive placebo. The *"control"* part is key to the comparison. It provides another group of participants in whom the outcome (disease or no disease, average blood pressure, cure or no cure) can be compared to the outcome observed in the active treatment group.

The controlled experiment is analogous to a laboratory study in experimental animals. One group of animals receives the active treatment, the other group receives an inactive placebo or another active treatment. Two main differences distinguish animal and human experiments: a scientific one and an ethical one.

The scientific difference is that the animals who receive both treatments are usually genetically identical mice, rats, fruit flies, etc. Humans, thankfully, are not genetically identical, unless they are monozygotic (from a single fertilized egg) twins, triplets, etc. The question then becomes: How can a researcher ensure that the two groups of human participants receiving the two different treatments are identical in all respects *other than* receipt of the active versus the control treatment?

The answer is randomization. Letting the flip of a coin or a computer-generated random sequence of numbers determine which participants receive which treatment does not guarantee that each participant is equivalent to every other participant in

the two study groups. Instead, it guarantees *exchangeability*. Exchangeability means that the two groups are virtually identical *on average* and would have been equally similar had those receiving the active and control treatments been switched—in other words, had they received the opposite treatment. This type of human experiment is called a *randomized controlled trial*, or RCT. As mentioned earlier in this chapter, the RCT design is required for licensing new drugs. The RCT is the "gold standard" for making causal scientific inferences, not only in drug studies but in all human health research.

The ethical difference between human experiments (RCTs) and animal experiments goes beyond the legal and moral necessity to obtain human participants' informed consent. That necessity applies to all human studies, not just RCTs. But it is unethical to administer interventions that are known or suspected to be harmful to human beings, even if they consent to those interventions. We cannot randomize children to be exposed to lead versus a placebo or to physical punishment versus "time out" approaches to discipline.

Instead, studies of the effects of hypothesized harmful exposures must be non-experimental by design. We call these *observational studies*. Of course, RCTs also require observations; all participants must be observed to see if and when then they develop the outcome hypothesized to be caused or prevented by the active intervention. But in observational studies, the researcher does not intervene. He or she merely observes both the exposure (treatment) and the outcome and then compares the outcomes in groups of exposed and unexposed participants. Observational studies are also used to investigate exposures that are not known to be harmful, including common health behaviors and treatments chosen by the participants or their care givers. The key feature of observational studies that distinguishes them from RCTs is the lack of exchangeability of exposed and unexposed participants that randomized treatment allocation provides. Table 1.1 compares and contrasts the main features of experimental and observational studies.

Table 1.1 Comparison of experimental and observational human health studies

Experimental studies	Observational studies
Researcher intervenes to change exposure	No intervention by researcher
Interventions often referred to as treatments	Exposure is observed, not assigned, by researcher
Treatment is hypothesized to prevent or ameliorate an illness	Exposure may be hypothesized as harmful or beneficial, may be a treatment
Two or more treatments usually compared:	Two or more exposures usually compared:
• Experimental (new) treatment	• Main exposure of interest
• One or more control treatments	• Non-exposure or control exposure
• Most rigorous design randomly allocates treatment to participants (randomized controlled trial, RCT)	
Health outcome is observed	Health outcome is observed

As shown in Fig. 1.1, the inference that exposure causes a change in outcome critically depends on knowing the temporal sequence of exposure and outcome. Whether a study is experimental or observational, it is essential that participants have not yet developed the outcome at the time they are exposed. An outcome that precedes the exposure cannot have been caused by that exposure.

Bias and Precision

In the context of causal inference in human health studies, bias refers to an observed association between exposure and outcome that differs *systematically* (that is, not merely by chance) from the true causal effect of exposure on outcome. In other words, the researcher is likely to observe an association in the absence of a true effect, fail to observe an association in spite of a true effect, or observe an association that is stronger or weaker than the true effect. I will focus on the two most important sources bias: confounding and reverse causality. Both are illustrated in Fig. 1.1.

Confounding occurs when a third factor (neither the exposure nor the outcome) biases the association between the study exposure and outcome. The bias occurs because, as shown by the arrows in Fig. 1.1, the confounding third factor is an underlying (antecedent) cause of both the exposure and the outcome. For example, let's say we knew nothing about the fact that cigarette smoking causes lung cancer. A clever researcher carries out an observational study of 100 cases of lung cancer and 100 controls without lung cancer; this is called a case-control study. The researcher carefully interviews and examines the 100 cases and 100 controls. Of the 100 cases, 30 are found to have yellow fingers on their dominant hand, whereas only 3 of the 100 controls have this finding. It would be incorrect to infer that yellow fingers cause lung cancer, because (as we now know) both the yellow fingers and the lung cancer are caused by smoking cigarettes. This bias can be reduced or eliminated by measuring and adjusting for the confounding factor through one of several statistical techniques. For example, if we analyze smokers and non-smokers separately, we will find none of the non-smoking cases or controls to have yellow fingers but a similarly high proportion of smokers with yellow fingers both among cases and controls.

The second important source of bias is *reverse causality*. It is illustrated by the red arrow in Fig. 1.1. This bias occurs when the outcome actually precedes and causes the exposure, rather than the reverse. It is particularly likely to occur in what are called cross-sectional studies, because exposure and outcome are ascertained at the same moment (cross-section) of time. For example, many of the studies investigating whether a large number of hours per day spent in front of a television or computer screen causes obesity are based on a cross-sectional design in which particpants are weighed and measured and are interviewed about how many hours per day they spend watching television or using a computer. If those measurements and interviews occur around the same time, we have no way

of knowing if a positive association reflects the causal effect of screen time on causing obesity or the causal effect of obesity on screen time. Either direction is biologically plausible. The only way to be sure of inferring the correct direction is to design a *longitudinal* (*prospective*) study in which eligible participants have a normal body weight at baseline when the hypothesized cause, prolonged screen time, is also measured. The participants are then followed up over time, and the proportion of *new* cases of obesity is compared in those with and without prolonged screen time at baseline.

Confounding and reverse causality biases are much more likely in observational studies than in randomized trials (RCTs). Because "association does not prove causation," it is sometimes claimed that causal inference *requires* a randomized trial. But bias can occur even in randomized trials. For example, confounding can occur if the treatment received is not well concealed from participants or care givers (we say they are not "blinded") and leads to other co-interventions that affect the trial outcome.

Conversely, well-designed observational studies that consistently show strong associations with a dose–response relation (for example, higher risks of the outcome in participants with higher levels of exposure), as well as confirmation in repeated studies in different settings, often provide sufficient evidence of causation to take action. That cigarette smoking causes lung cancer can no longer be debated, despite the efforts of tobacco companies to undermine the "merely observational" evidence base. The reduced lung cancer risk in ex-smokers, the fall in lung cancer incidence in countries that have succeeded in reducing their smoking rates, and the rise in incidence in other countries with increased smoking provide strong evidence of causality despite the "observational" design of the studies demonstrating the association. Similar arguments can be made for prone sleeping position as a cause for the sudden infant death syndrome, or SIDS.

Precision is different from bias. Like bias, insufficient precision can lead to an error in the estimate of an exposure-outcome association, whether that estimate comes from an observational study or an RCT. Unlike bias, however, precision is the degree of uncertainty about the magnitude of association due to chance variation. Imprecision, or low precision, leads to an estimate that is not systematically too high or too low, but one that shows a wide range of statistical uncertainty around the observed estimate. It is usually due to a small sample size and often prevents detection of a true association or effect. The observed association or effect is called *statistically non-significant*. In other words, it may be entirely attributable to chance. For example, if our above-mentioned study of lung cancer and yellow fingers had included only ten cases and ten controls, we might have observed three cases and no controls with yellow fingers. That would be a statistically non-significant result, because the sample size of only 20 total participants might well yield a difference of this magnitude (three out of ten vs. zero out of ten) solely by chance even if yellow fingers had no true association with lung cancer. This "false-negative" finding has nothing to do with confounding by cigarette smoking. It is merely a consequence of an insufficient sample size, that is, imprecision.