Focus on *Understanding* Statistics

A Clinician's Guide to Statistics in Mental Health. Second Edition. By S. Nassir Ghaemi; Cambridge University Press; Cambridge, UK; 2023; ISBN 9781108814966; pp. 173; $ 34.99 (Paperback)

Richard Balon[1]

Statistics, as a discipline, does not exist in a vacuum. It is a reflection of our views on science, and thus how it is understood and how it is used depends on what we mean by "science." (p. 107)

Most clinicians and educators, and perhaps even some scientists, do not truly understand statistics, or their understanding of it is superficial and does not go much beyond *t*-tests and *p*-values. There are various reasons for this lack of understanding. Most physicians are not highly scientifically oriented, and as the author of this volume, S. Nassir Ghaemi, points out, "are wary of mathematics, and, therefore rather allergic to statistics" (p. x). As Ghaemi adds, the situation is not helped by the fact that most statisticians are comfortable with mathematics, and "when they teach statistics they do so with many mathematical equations" (p. x). Clinicians and educators are "turned off by the highly mathematical nature of statistical texts" (p. x). Statistics is not typically taught contextually by using examples from clinical and educational research to explain its usefulness in specific studies. I can attest to this, as I have sat in many classes and presentations where the explanation and context were missing. Ghaemi, a clinician and researcher who has not been formally educated in statistics beyond a master's degree in public health (and a master's in philosophy), decided to address these issues and write a book of statistics for mental health clinicians. As Ghaemi says, "This is a book of statistics with no or little requirement for prior experience with it, and with few mathematical equations" (p. x). This is certainly a novel and useful approach to open the realm of statistics to those who do not possess math and statistical skills.

✉ Richard Balon
  rbalon@wayne.edu

1   Wayne State University, Detroit, MI, USA

Ghaemi presents his explanation and views of statistics in 20 brief chapters and one appendix, beginning with simple concepts and gradually progressing to more complicated topics. Ghaemi begins with a discussion of the fact that the data never speak for themselves, followed by discussions of levels of evidence; bias; randomization; using clinical trials to improve clinical experience; uses and misuses of *p*-values; the importance of effect size; understanding placebo effects; understanding confidence intervals; observational studies; the alchemy of meta-analysis; Bayesian statistics; and causation. He moves on to describe the philosophy of statistics; a defense and criticism of evidence-based medicine; social and economic factors such as peer review and funding; an explanation of when clinical trials are misinterpreted (using CATIE, STAR*D, and STEP-BD studies); suggestions on how to analyze a study; false positive maintenance clinical trials in psychiatry; and understanding regression.

Throughout the book, Ghaemi repeatedly points to the fact that scientific knowledge is not absolute and all science involves uncertainty, which statistics help us understand through interpreting facts and uncertainty. Ghaemi draws from the writings of many statisticians, scientists, and philosophers. The book's chapters are short and do not deal with the mechanisms of specific statistical tests but, rather, focus on understanding and applying them.

I will summarize some of the important principles and observations from this volume. Ghaemi points out that the most important principle of modern science and efforts to address bias is randomization. "Randomization is the best means of addressing confounding bias, and blinding the best means for addressing measurement bias" (p. 7). The discussion of specific levels of evidence, the highest being double-blind, placebo-controlled trials and the lowest being case series, case reports, and expert opinions, notes that each level has its strengths and weaknesses, and thus, no single level is completely useful or

completely useless. In the discussion of randomized controlled trials (RTCs), Ghaemi argues that "Not all RTCs are created equal" (p. 26), RTCs have many limitations, and "The basic rationale for clinical trials is to remove confounding bias and thereby to clarify what is true and what is false in clinical experience" (p. 27). He also notes that *p*-values are frequently inflated, and thus, statistical correction (Bonferroni) may be required. Ghaemi notes that looking into statistical analysis of subgroups in RTCs may be misleading and that we should turn to descriptive statistical methods instead. He cautions about the generalizability of RTCs as "for any major psychiatric condition, about 10% of patients with the relevant diagnosis will qualify for and agree to participate in available RTCs" (p. 42), and thus, it is difficult to generalize results to the remaining 90%. The text of the chapter on clinical trials also provides a good discourse on a famous meta-analysis that claimed antidepressants to be ineffective treatments (they usually do not work in mild depression [placebo may do better], but they have "clinically meaningful benefit over placebo in severe depression" [p. 28]).

Another important discussion in this book focuses on the importance of effect size, as "Effect sizes tells you *how much* change was seen in a parameter" (p. 52), while *p*-value will just tell you whether the change was significant (think statistically vs. clinically significant). This is especially important to consider in the studies examining whether antidepressants work. I liked the writing on the placebo effect that complicates the results of both medication and psychotherapy studies: "A drug effect is never just a drug effect. It is added to or subtracted from by the clinician-patient relationship" (p. 64) and even highly manualized and empirical therapies "have a placebo effect based on the interpersonal relationship between the therapist and the client" (p. 64). The chapter on confidence intervals helped me understand this concept much better. "Confidence intervals can be defined as the range of plausible values for the effect size" (p. 71) or "that it is the likelihood that the real value for the variable would be captured in 95% of trials" (p. 71). I appreciated the chapter on meta-analysis (called "statistical alchemy" by epidemiologist Alvan Feinstein), which emphasizes that meta-analyses are frequently not very useful and that" Meta-analysis is never more valid than an equally large single RTC" (p. 83). Ghaemi writes succinctly, "Meta-analysis is a tool, not an answer" (p. 83).

In many parts of the book, Ghaemi turns to the philosophy of science. It starts in the chapter on causation, where Ghaemi writes that "Causation is essentially a philosophical, not a statistical problem" (p. 97), and that "The numbers can never give the complete answer, because they are never definitive" (p. 99) and "Statistics, by nature, are never absolute: they are about measuring the probability of error; they can never remove error" (p. 99). Here, he also reviews Bradford Hill's criteria of causation, which included strength of the association; consistency of the association; specificity of the association; temporality; biological gradient; plausibility ("clinical observation trumps biology, not vice versa" [p. 101]); coherence; experiment; and analogy. Ghaemi also emphasizes the importance of replication. The chapter on the philosophy of statistics poses the question of whether to verify or refute, discussing the work of Sir Karl Popper. Equally interesting is the chapter on defense and criticism of evidence-based medicine. The text emphasizes the importance of clinical observation, which is aptly illustrated by the statement that "The Cochrane database completely ignores all observational studies, and thus it would not include any "evidence" that penicillin is effective" (p. 118).

The chapter on social and economic factors includes a discussion of peer review ("akin to Winston Churchill's definition of democracy: It is the worst system imaginable, except for all others" [p. 121]) and related issues such as who reviews the reviewers, mediocrity being rewarded through peer review, the "almighty" impact factor, and its distorting effect; the intangibles of coauthorship and self-censorship on the part of coauthors; ghost authorship; who owns the data; unpublished negative studies; and funding of research. These are not exactly statistical issues but are related to our understanding of research and published science. Important also is the chapter discussing the well-known psychopharmacological studies: CATIE, STAR*D, and STEP-BD. It concludes that although these studies "were conducted outside the realm of the pharmaceutical industry, by academic centers and with governmental funding, in each case the researchers seemed strongly biased toward the most positive spin they could give to the drugs being studied" (p. 138). The critique is good but could have been more detailed. The last two chapters present some advice on how to analyze a study (one wishes that there was more on this topic and clearer directions were given), and a discussion of false positive maintenance trials in psychiatry, using, for example, analysis of maintenance studies of lithium and lamotrigine.

This is an interesting, provocative, and unusual statistical book. I liked the frequent use of examples of studies from the psychiatric literature (and the famous study on the harms of tobacco). The focus on understanding statistics is important and useful. It does not include every statistical test, but do clinicians need them all? Those interested in learning more about statistics could probably combine this book with some classic text. I believe that educators may find this volume especially useful for the required teaching of research literacy.

Declarations