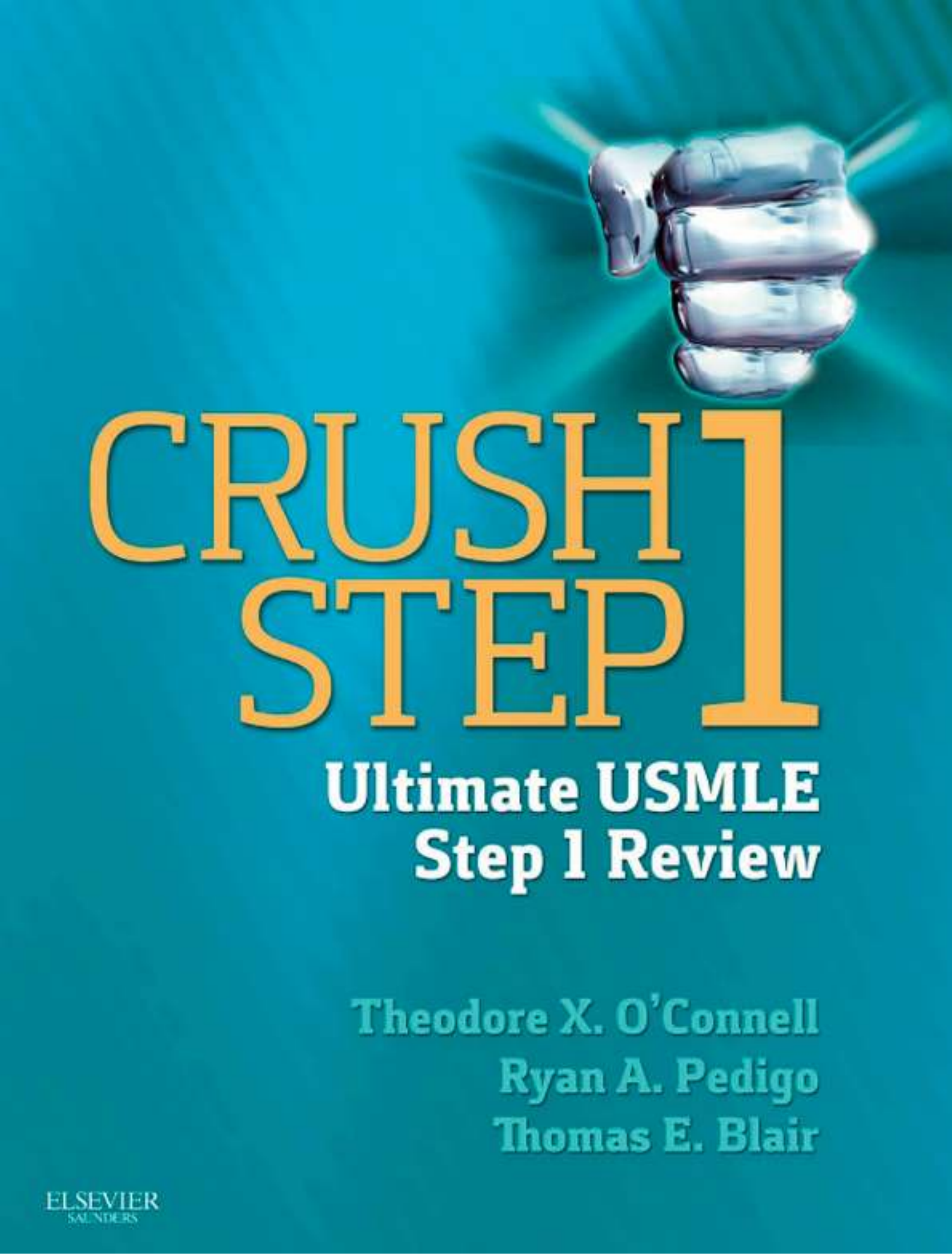
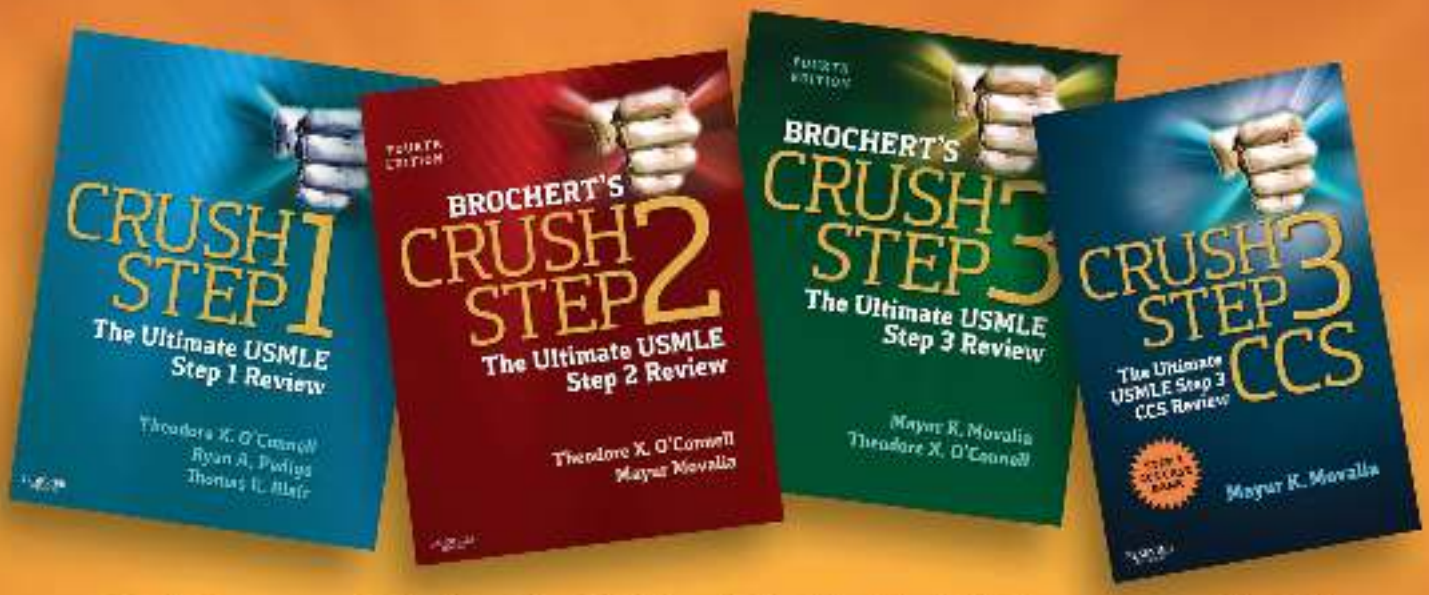

CRUSH
STEP 1
Ultimate USMLE
Step 1 Review
Theodore X. O'Connell
Ryan A. Pedigo
Thomas E. Blair
ELSEVIER
SAUNDERS

Don't just **PASS** your **USMLEs**—

CRUSH THEM!



Crush Step 1
ISBN: 978-1-4557-3621-6
2014

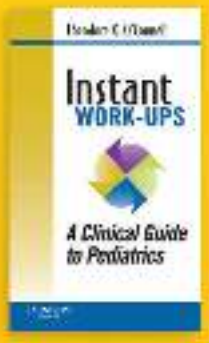
Brochert's Crush Step 2, 4th Edition
ISBN: 978-1-4557-0831-1
2013

Brochert's Crush Step 3, 4th Edition
ISBN: 978-1-4557-0830-4
2013

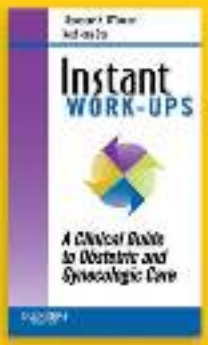
Crush Step 3 CCS
ISBN: 978-1-4557-2874-4
2013

More **GREAT RESOURCES** from Dr. O'Connell!

USMLE Step 2 Secrets, 4th Edition
ISBN: 978-0-323-28814-2
2015



Instant Work-ups: A Clinical Guide to Pediatrics
ISBN: 978-1-4160-5462-7
2010



Instant Work-ups: A Clinical Guide to Obstetric and Gynecologic Care
ISBN: 978-1-4160-5461-2
2009

Instant Work-ups: A Clinical Guide to Medicine
ISBN: 978-1-4160-5460-5
2009

Shop online today at **us.elsevierhealth.com**, or call **1-800-545-2522**!

ELSEVIER

CRUSH STEP 1

THE ULTIMATE USMLE STEP 1 REVIEW

This page intentionally left blank

CRUSH STEP 1

THE ULTIMATE USMLE STEP 1 REVIEW

Theodore X. O'Connell, M.D.
Program Director, Family Medicine Residency Program
Kaiser Permanente Napa-Solano, California
Assistant Clinical Professor, Department of Family Medicine
UCLA David Geffen School of Medicine, Los Angeles, California

Ryan A. Pedigo, M.D.
Resident Physician Emergency Medicine
Harbor-UCLA Medical Center, Los Angeles, California

Thomas E. Blair, M.D.
Resident Physician Emergency Medicine
Harbor-UCLA Medical Center, Los Angeles, California

ELSEVIER

ELSEVIER
SAUNDERS

1600 John F. Kennedy Blvd.
Ste 1800
Philadelphia, PA 19103-2899

Crush Step 1 The Ultimate USMLE Step 1 Review ISBN: **978-1-4557-5621-6**

Library of Congress Cataloging-in-Publication Data
O'Connell, Theodore X., author.
Crush step 1 : the ultimate USMLE step 1 review / Theodore X. O'Connell, Thomas Blair, Ryan Pedigo.
p. ; cm.
Includes index.
ISBN 978-1-4557-5621-6 (pbk. : alk. paper)
I. Blair, Thomas (Emergency physician), author. II. Pedigo, Ryan, author. III. Title.
[DNLM: 1. Clinical Medicine--Examination Questions. WB 18.2]
RC58
616.0076–dc23 2013033016

Senior Content Strategist: James Merritt
Content Development Specialist: Julia Roberts
Publishing Services Manager: Hemamalini Rajendrababu
Project Managers: Divya KrishnaKumar
Designer: Steven Stave

Printed in China
Last digit is the print number: 9 8 7 6 5 4 3 2 1

FACULTY REVIEW BOARD

This page intentionally left blank

STUDENT REVIEW BOARD

The authors and publisher express sincere thanks to these students and residents, who provided many useful comments and helpful suggestions for improving the content and for ensuring the accuracy of this book.

Joshua Elder, M.D.
Resident Physician
Department of Emergency Medicine
Stanford University Medical Center/Kaiser Permanente
Stanford, California

Jennifer Neeper, M.D.
Resident Physician
Department of Obstetrics and Gynecology
University of California, Irvine Medical Center
Irvine, California

Jake Ortiz, M.D.
Resident Physician
Department of Internal Medicine
University of California, Davis Medical Center
Davis, California

Steven Storage
M.D. Candidate, Class of 2013
UCLA David Geffen School of Medicine
Los Angeles, California

This page intentionally left blank

CONTRIBUTORS

Will Babbitt
M.D. Candidate, Class of 2013
University of California, San Francisco School of Medicine
San Francisco, California

John Baird
M.D. Candidate, Class of 2013
University of Arkansas College of Medicine
Little Rock, Arkansas

Brenton Bauer, M.D.
Resident Physician
Department of Internal Medicine
UCLA Ronald Reagan Medical Center
Los Angeles, California

Thomas E. Blair, M.D.
Resident Physician
Emergency Medicine Resident Physician
Harbor-UCLA Medical Center
Los Angeles, California

Manuel Celedon, M.D.
Resident Physician
Department of Emergency Medicine
Harbor-UCLA Medical Center
Los Angeles, California

Masood Memarzadeh, M.D.
Resident Physician
Department of Anesthesia and Perioperative Care
University of California
San Francisco, California

Theodore X. O'Connell, M.D.
Program Director, Family Medicine Residency Program
Kaiser Permanente Napa-Solano, California
Assistant Clinical Professor, Department of Family Medicine
UCLA David Geffen School of Medicine, Los Angeles, California

Gwen Owens
M.D./Ph.D. Candidate
UCLA David Geffen School of Medicine
California Institute of Technology
Los Angeles, California

Ryan A. Pedigo, M.D.
Resident Physician
Emergency Medicine Resident Physician
Harbor-UCLA Medical Center
Los Angeles, California

Tiffany Prugpichailers, M.D.
Resident Physician
Department of Pediatrics
UCLA Mattel Children's Hospital
Los Angeles, California

Tina Roosta
M.D. Candidate, Class of 2013
UCLA David Geffen School of Medicine
Los Angeles, California

Lauren Sanchez
M.D. Candidate, Class of 2013
UCLA David Geffen School of Medicine
Los Angeles, California

Manpreet Singh, M.D.
Resident Physician
Department of Emergency Medicine
Harbor-UCLA Medical Center
Los Angeles, California

INTRODUCTION

Crush Step 1 was conceptualized, designed, and created by medical students, then edited by experts in their fields, with the goal of being the best resource on the market for the USMLE Step 1. The idea for Crush Step 1 came from the common frustration from medical students that either review books are subject specific and too detailed for the USMLE Step 1 or they are the opposite: too short and require memorization of facts rather than understanding of concepts. Crush Step 1 aims to allow students to gain a deeper understanding of concepts that are tested in the first 2 years of medical school and on the USMLE Step 1 exam. Most important, Crush Step 1 is a book that students can use on their first day of medical school. When students understand the concepts of how anatomy, physiology, pathology, and pharmacology interact with one another, medicine becomes less about memorization and more about truly learning how the body works—and doesn't work.

Writing Crush Step 1 was a complex task that required careful analysis and planning. First, medical students from multiple medical schools were questioned regarding their satisfaction with current review books, what they did and did not like about other review books, and what characteristics their ideal review book would have. The response was resounding and almost unanimous: medical students want a text that allows the actual understanding of subjects, includes review questions, and contains full-color figures. This was the framework for what Crush Step 1 set out to accomplish. After many of the chapters were written, they were "beta-tested" to make sure that students liked the format and understood the concepts. Each chapter underwent multiple rounds of editing and refining until we were happy with the end result. Finally, we added high-yield boxes for rapid review and questions at the end of key sections to facilitate review of critical concepts and ensure comprehension. At the end, faculty editors lent their expertise on key chapters and helped guide the refinement of each chapter.

We sincerely hope you enjoy Crush Step 1. Do not hesitate to contact us to help us in our journey to make the best review book available–the review book that we wish we would have had. This is your book, and we want to make sure that everything is clear, precise, and serves your needs. Although we have used numerous editors and have tried our best to ensure that all the content is accurate, it is possible that errors occur within the text. Please help us with our goal to make this text the best resource for medical students available by submitting any comments, suggestions, or corrections to CrushStepOne@gmail.com. We welcome your feedback and corrections and wish you the best for the USMLE Step 1 and beyond.

Thanks again, and best of luck to you on your journey toward becoming outstanding physicians.

Sincerely,
The Crush Step 1 Team

This page intentionally left blank

CONTENTS

1 BIOSTATISTICS

Thomas E. Blair

MEAN, MEDIAN, AND MODE

- **Sample value set:** 1, 1, 2, 4, 5, 7, 7, 25, where $n = 8$
- **Mean:** The average of a sample. It is calculated by adding all values, then dividing by the number of values (n). In this example $(1 + 1 + 2 + 4 + 5 + 7 + 7 + 25)/8 = 6.5$. The mean is sensitive to extreme values.
- **Median:** The middle value of a sample. It is equivalent to the 50th percentile such that half the sample values are above and half are below. It is identified by arranging the values in ascending order, then finding the middle-most number. If n is odd, the median is the $[(n + 1)/2]$th largest observation. If n is even, the median is the average of the $(n/2)$th and the $(n/2 + 1)$th largest observations. In this example, there is an even number of values, so the median is the average of the two middle-most numbers; that is, for 1, 1, 2, 4, 5, 7, 7, 25, the median $= (4 + 5)/2 = 4.5$. An advantage of the median is that it is insensitive to extreme values. You may notice that in this sample the mean is greater than the median. This indicates that the distribution has a positive skew (see later discussion).
- **Mode:** The most frequently occurring value in a sample. In this example, both 1 and 7 are modes because they both appear twice. Therefore, this data set can be said to be bimodal.
- **Standard deviation (SD):** A measure of the spread and variability of a data set, calculated as the square root of the variance. It represents the average deviation from the mean. The closer the values remain to the mean, the smaller the SD (Fig. 1-1). The concept, not the mathematics, may be tested on Step 1.
  - *Example:* Normal body temperature will have a small SD because an individual's anterior and posterior hypothalamus maintains temperature homeostasis within a very limited range. Blood sugars, on the other hand, will have a larger SD because glycemic loads change throughout the day.

Mean: average value
Median: middle value
Mode: most frequent value

Standard deviation represents the average deviation from the mean.

DEFINITIONS

- **Incidence:** The number of **new** cases of a disease in a population over a specific period of time (longitudinal).
- **Prevalence:** The total number of people in a population affected by a condition at one point in time (cross-sectional).
- **Duration** relates incidence to prevalence.
  - *Example:* Upper respiratory infections (URIs) have a high incidence every year during winter months, but a low prevalence because most URIs resolve quickly. Diabetes has a relatively low incidence but high prevalence because a patient who has diabetes has it for life.
- **Normal distribution:** Also known as a gaussian distribution or bell-shaped curve. A probability function in which values are symmetrically distributed around a central value, and the **mean, median, and mode are equal**. In a normal distribution, 1 SD accounts for 68% of all values, 2 SDs account for 95% of all values, and 3 SDs account for 99.7% of all values—the **68-95-99 rule** (Fig. 1-2). The area under the curve is 1 (100%).
  - *Example:* The intelligence quotient (IQ) test is constructed to follow a normal distribution with a mean of 100 and SD of 15. That means that 95% of the population (2 SDs) will have an IQ between 70 and 130. Of clinical import, mental retardation is defined as an IQ of <70.

In a normal distribution, mean = median = mode.

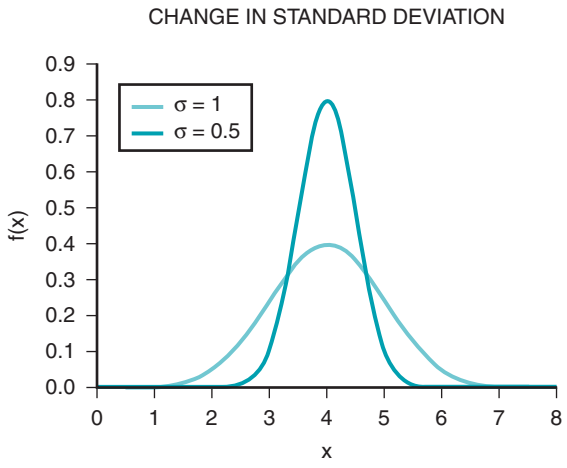


FIGURE 1-1 Data sets with the same mean, but different standard deviations. The wider curve has a larger standard deviation than the taller curve.

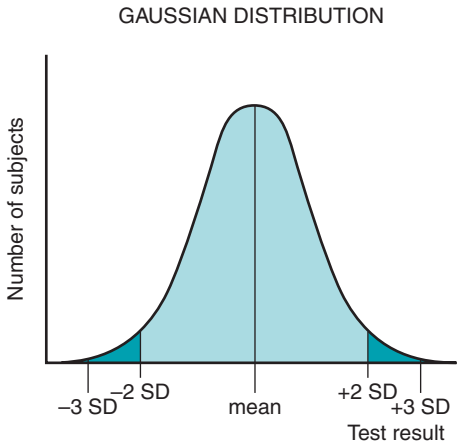


FIGURE 1-2 In a normal distribution, 1 standard deviation (SD) accounts for 68% of all values, 2 SDs account for 95% of all values, and 3 SDs account for 99% of all values. *(From Marshall WJ, Bangert SK.* Clinical Chemistry. *6th ed. Philadelphia: Elsevier; 2008.)*

- **Bimodal distribution:** A distribution with two modes.
  - *Example:* The incidence of Crohn's disease displays a bimodal distribution with the first peak between 15 and 30 years of age and the second peak between 60 and 80 years of age.
- **Negative skew:** An asymmetrical distribution in which a tail on the left indicates that **mean < median < mode**. The tail is due to outliers on the left side of the curve.
  - *Example:* A graphic representation of age at death would show a negative skew, with most people clustered at the right end of the distribution and relatively few dying at a younger age (Fig. 1-3A).
- **Positive skew:** An asymmetrical distribution, in which a tail on the right side indicates that **mean > median > mode**. The tail is due to outliers on the right side of the curve.
  - *Example:* A graphic representation of age at initiation of smoking would display positive skew. Most people would be clustered around their late teens, but a small number of middle-aged and older adults, who initiated smoking later in life, create a positive tail (Fig. 1-3B).

Negative skew: tail on left
Positive skew: tail on right

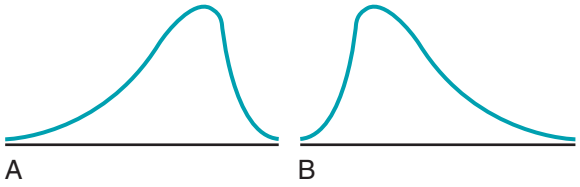


FIGURE 1-3 A, Negative skew. **B,** Positive skew. *(From Jekel JF, Katz DL, Wild JG, Elmore DMG.* Epidemiology, Biostatistics, and Preventive Medicine. *3rd ed. Philadelphia: Elsevier; 2007.)*

Statistics for Diagnostic Tests

- **True positive (Tp):** Disease is present and diagnostic test is positive (a correct result).
- **True negative (Tn):** Disease is absent and diagnostic test is negative (a correct result).
- **False positive (Fp):** Disease is absent and diagnostic test is positive (an incorrect result).
- **False negative (Fn):** Disease is present and diagnostic test is negative (an incorrect result).
- **2 × 2 Table:** Consider drawing a 2 × 2 table whenever faced with a biostatistics question. To consistently put the table in the right order, try using the mnemonic "truth over test results"; that is, the truth is more important than the test. Intuitively, "yes" precedes "no" on both axes (Table 1-1).
- **Prevalence:** This is defined as (total number of people with disease)/(total number of people studied) or, from the 2 × 2 table, (Tp + Fn)/(Tp + Tn + Fp + Fn). Note that prevalence varies by population. For example, Chagas disease has a higher prevalence in South America than it does in the United States.
- **Sensitivity:** Given the disease is present, the probability that the test will be positive. Stated another way, sensitivity is the ability of a test to become positive in the presence of the disease. It is defined as Tp/(Tp + Fn) or Tp/(total number of people with disease). Sensitive tests are useful for screening because there are few false negatives. A highly sensitive test can, therefore, rule out the disease. Consider the mnemonic **SN-N-OUT**—for a test that is **S**e**N**sitive, a **N**egative result rules **OUT** a disease.
  - *Example:* An HIV test with 98% sensitivity means that, when a disease is present, it will be detected 98% of the time.
  - *Example:* Consider a test that was positive 100% of the time regardless of the presence or absence of disease. It would technically have 100% sensitivity because it would be positive in all the patients with disease (but would be clinically useless because it would be positive in all the patients without disease, too). Therefore, the sensitivity is not the whole picture when it comes to test characteristics; specificity is also important.
- **Specificity:** Given the disease is absent, the probability that the test will be negative. Stated another way, specificity is the ability of a test to remain negative in the absence of disease. It is defined as Tn/(Tn + Fp) or Tn/(total number of people without disease). Tests with high specificity are useful to confirm a diagnosis because there are few false positives. A highly specific test can, therefore, rule in the disease. Consider the mnemonic **SP-P-IN**—for a test that is **SP**ecific, a **P**ositive result rules **IN** a disease.
  - *Example:* An HIV test with 98% specificity means that, when a disease is absent, the test will be negative 98% of the time.

SN-N-OUT: a test that is SeNsitive, when Negative rules OUT a disease

SP-P-IN: a test that is SPecific, when Positive rules IN a disease

In general, there is a tradeoff between sensitivity and specificity. For example, changing the cutoff value for an "elevated" serum calcium level will change the test's ability to detect a diseased population with hyperparathyroidism. Raising the cutoff value will increase the specificity (fewer false positives), but will also decrease the sensitivity (more false negatives) (Fig. 1-4). As another example, if the random blood sugar cutoff for the diagnosis of diabetes were moved from 200 mg/dL to 1000 mg/dL (obviously extreme), then the test would be very specific because, of course, anyone with a blood sugar over 1000 mg/dL likely has diabetes (very few false positives). However, it would be very insensitive because people with diabetes with random blood sugar readings of 900 mg/dL would have false negatives.

Increasing the diagnostic cutoff will make a test more specific but less sensitive.

Table 1-1 Example of a 2 × 2 Test Result for Calculating Biostatistics

		The "Truth"	
		Yes	**No**
Table	**Yes**	**A** True +	**B** False +
	No	**C** False –	**D** True –

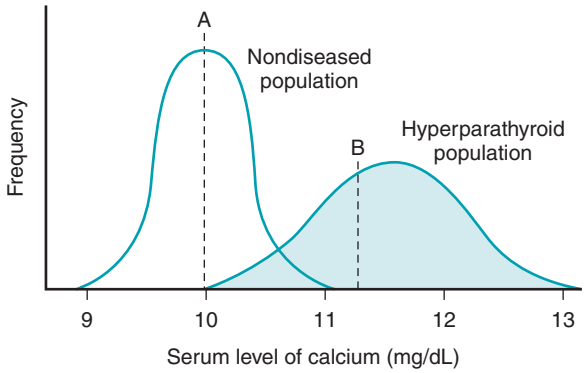


FIGURE 1-4 Increasing the diagnostic cutoff (moving it to the right) will make the test more specific but less sensitive. In this example, if the cutoff point is at A, then the test will be 100% sensitive because all patients with hyperparathyroidism will be detected. At cutoff A, however, the test is very nonspecific because many healthy people will be incorrectly classified as having hyperparathyroidism. If the cutoff point is moved to B, the test will be 100% specific because all positive tests will be true positives. At cutoff point B, however, the test is not sensitive (many cases of hyperparathyroidism will be missed). *(From Jekel JF, Katz DL, Wild JG, Elmore DMG.* Epidemiology, Biostatistics, and Preventive Medicine. *3rd ed. Philadelphia: Elsevier; 2007.)*

Positive Predictive Value (PPV): Given the test is positive, the probability that the disease is present. PPV = Tp/(Tp + Fp) or Tp/(total number of positive tests). For example, if a computed tomography (CT) scan has 98% specificity for appendicitis, then given a positive finding of appendicitis, the patient will truly have the disease 98% of the time.

Negative Predictive Value (NPV): Given the test is negative, the probability that the disease is absent. NPV = Tn/(Tn + Fn) or Tn/(total number of negative tests). For example, if an HIV test has 98% NPV, then given a negative test, the patient will truly be HIV negative 98% of the time.

- Once the result of a diagnostic test is known, this statistical information is most useful. For example, if a diagnostic test for a pulmonary embolism (PE) is positive, the question is: "Given a positive result, what is the actual probability a PE is present?"
- However, PPV and NPV vary depending on the prevalence of the disease in a population. They must be used with caution if your patient population is not identical to the one studied. Sensitivity (and specificity) will not be affected because the ratio of Tp/(total number of people with disease) will not change for a given test, but the ratio of Tp/(total number of positive tests) will vary because an area with higher prevalence will have a higher number of positive tests.

Sens = Tp/(Tp + Fn)
Spec = Tn/(Tn + Fp)
PPV = Tp/(Tp + Fp)
NPV = Tn/(Tn + Fn)

Calculations of Risk Measures

2 × 2 Table: You can also draw a 2 × 2 table whenever faced with calculating risk or odds. To remember how to set up the table, think "outcome over exposure" (i.e., the outcome is more important than the exposure). Intuitively, "yes" precedes "no" on both axes. Note that the exposure can be "good," such as a beneficial treatment, or "bad," such as a carcinogen or harmful medication. A 2 × 2 table for exposure can be set up similar to a 2 × 2 table for diagnostic tests, where the exposure can refer to a known characteristic, an observed exposure, or an assigned treatment (Table 1-2).

TABLE 1-2 Example of a 2 × 2 Table for Calculating Risk or Odds

		Outcome	
		Yes	**No**
Exposure	**Yes**	**A**	**B**
	No	**C**	**D**

The incidence (or prevalence in the setting of cross-sectional studies) is calculated for each group from the 2 × 2 table. The probability (p) that the event will occur in the exposed group is given by **risk in exposed = a/(a + b),** and in the unexposed (control) group it is given by **risk in unexposed = c/(c + d)**.

- *Example:* One hundred patients are treated with a statin and five suffer a myocardial infarction (MI). What is the incidence of MI in the group treated with ("exposed to") a statin?
  - *Solution:* a/(a + b) = 5/100 = 0.05 = 5%.

Risk Difference (RD): The difference between the two groups. RD = risk in exposed − risk in unexposed, or vice versa. There are several ways to express the risk difference.

- **Absolute risk reduction (ARR):** The reduction in incidence associated with a treatment. ARR = risk in control group − risk in treatment group.
  - *Example:* In one study, 5% of patients on a statin suffered an MI, whereas 9% of those on placebo suffered an MI. What is the ARR of statins?
    - *Solution:* ARR = 9% − 5% = 4%. The use of statins in this population will reduce the number of MIs by 4%.
- **Attributable risk:** The increase in disease incidence associated with an exposure. Attributable risk = risk in exposed − risk in unexposed.
  - *Example:* In one study, 9% of patients exposed to asbestos developed bronchogenic carcinoma, and 2% of those without exposure developed bronchogenic carcinoma. What is the attributable risk of asbestos?
    - *Solution:* ARR = 9% − 2% = 7%; that is, asbestos exposure increased the incidence of bronchogenic carcinomas by 7%.
- **Number needed to treat (NNT):** The number of patients required to receive an intervention before an adverse outcome is prevented (e.g., death, MI). **NNT = 1/ARR.** The opposite concept is **number needed to harm (NNH),** used for interventions or exposures that may be detrimental (e.g., radiation exposure). **NNH = 1/attributable risk.**
  - *Example:* In one study, the ARR of statin therapy is calculated at 4%. What is the NNT?
    - *Solution:* NNT = 1/.04 = 25. Twenty-five patients would need to be treated with statins to prevent one MI.
- **Relative risk (RR, risk ratio):** The ratio of incidence in the two groups. **RR = risk in exposed/risk in unexposed.** For a negative outcome, a ratio greater than 1 indicates a harmful treatment/exposure and a ratio less than 1 indicates a beneficial treatment/exposure, whereas a ratio of 1 is a null effect.
  - *Example:* In one study, 50% of diabetic patients developed heart disease, compared with 10% of a control population. What is the relative risk of diabetic patients developing heart disease?
    - *Solution:* RR = 0.5/0.1 = 5. Diabetic patients are 5 times more likely to develop heart disease than nondiabetic patients.
- **Relative risk reduction (RRR):** The percentage of diseases prevented by a treatment. **RRR = (risk in unexposed − risk in exposed)/risk in unexposed = ARR/baseline risk.** For harmful exposure, the equivalent concept is excess relative risk = (risk in exposed − risk in unexposed)/risk in unexposed.
  - *Example:* Five percent of patients on a statin suffered an MI, whereas 9% of those on placebo suffered an MI. What is the RRR of being on a statin?
    - *Solution:* RRR = (0.09 − 0.05)/0.09 = 0.44 = 44%. That is to say, in the statin group, MIs were reduced by 44% relative to those in the control group.
- **Odds:** The ratio of the probability of the outcome to the probability of not having the outcome. **Odds = p/(1 − p).**
- **Odds ratio:** A comparison of event rates between exposed and unexposed groups, calculated using odds instead of probabilities. It is the **odds of an event in the exposed group divided by the odds of the event in an unexposed group**. In a 2 × 2 table, it is calculated by: OR = (a/b)/(c/d) = ad/bc, thus simply cross-multiplication of the values in the 2 × 2 table. Odds ratios are somewhat unintuitive; however, one can simplify understanding by considering that OR > 1 implies increased likelihood of an event in the exposed group, OR < 1 implies decreased likelihood of an event in the exposed group, and OR = 1 implies no difference between the exposed group and the control group. **ORs are used instead of risk ratios in case-control studies** because the risk ratio cannot be calculated from the study data owing to purposeful oversampling of cases in the study design. The OR will approximate the relative risk if the outcome is rare.

Risk in exposed = a/(a + b)
Risk in unexposed = c/(c + d)

ARR = control group risk − treatment group risk
Attributable risk = risk in exposed − risk in unexposed
NNT = 1/ARR
RR = Risk in exposed/risk in unexposed
RRR = ARR/risk in unexposed
Odds ratio = ad/bc

- *Example:* Investigators conduct a case-control study to evaluate the risk for lymphoma due to CT scans. One hundred people are selected with lymphoma, and 100 people without lymphoma are selected as controls. Five patients with lymphoma had received a CT scan at some point, whereas only 2 patients without lymphoma had ever received a CT scan. Calculate the odds ratio.
  - *Solution:* First refer to the 2 × 2 table in Table 1-3. OR = (a/b)/(c/d) = ad/cb = 2.6. The odds of having lymphoma in those who received a CT scan are 2.6 times those of patients who never had a CT scan.

TABLE 1-3 Example of a 2 × 2 Table for Odds Ratio

		Outcome	
		Yes	**No**
Exposure	**Yes**	**A** = 5	**B** = 2
	No	**C** = 95	**D** = 98

STATISTICAL TESTS AND SIGNIFICANCE

Reliability is a measure of the consistency of a test, that is, the likelihood that, upon repetition, it will deliver the same results in the same situation. Reliability decreases as random error increases in a test. **Validity** is the ability of a test to measure what it is intended to measure. Reliability does not necessarily imply validity. For example, a test may reliably measure serum concentrations of vitamin D; however, this does not inherently mean it is a valid predictor of a disease such as osteoporosis.

Accuracy: ability to obtain "true results"
Precision: ability to replicate results

Inter-rater reliability reflects to what degree test results will vary depending on who is administering the test. For example, body temperature has good inter-rater reliability when a thermometer is used, but poor inter-rater reliability when simple touch is used.

Accuracy: Analogous to validity and a measure of a test's ability to obtain "true" results (Fig. 1-5B).

Precision: Analogous to reliability and a measure of a test's ability to replicate results (Fig. 1-5C).

Categorical Data: Data with a fixed number of nominal categories, or data that has been grouped as such (e.g., race, gender, or living/dead).

Continuous Data: Data that can take any value within a range (e.g., height or weight).

Categorical data: nominal categories
Continuous data: can take any value within a range

Statistical Tests for Data Analysis

See (Table 1-4) for a summary of statistical tests.

Two-Sample t-Test: Compares the means (continuous data) of two groups and determines whether there is a difference between the means based on a predetermined level of significance (α). For example,

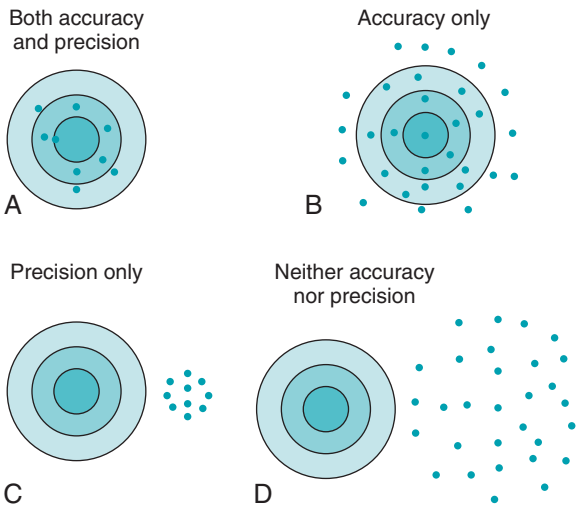


FIGURE 1-5 Precision versus accuracy. *(From Jekel JF, Katz DL, Wild JG, Elmore DMG.* Epidemiology, Biostatistics, and Preventive Medicine. *3rd ed. Philadelphia: Elsevier; 2007.)*

Table 1-4 Summary of Statistical Tests

NAME	NUMBER OF GROUPS	TYPE OF DATA	COMMENT
Two-sample t-test	2	Continuous	Compares 2 means
ANOVA	3+	Continuous	Compares 3+ means
Chi-square test	2+	Categorical	Compares 2+ groups
Correlation coefficient	2	Continuous	Measures covariance

is there a significant difference in the average systolic blood pressure (continuous data) between men and women (categorical data)? As the name implies, only two groups can be compared in a two-sample t-test.

ANOVA (ANalysis Of VAriance): Serves a similar function to a t-test but can compare more than two groups. For example, is there a significant difference in the average systolic blood pressure (continuous data) of Asians, Caucasians, and African Americans (categorical data)?

Chi-Square Test: Compares proportions between groups of categorical data. There is no limit to the number of variables being compared. For example, is use of antibiotic x, y, or z (categorical) associated with a difference in survival from sepsis (categorical)?

Correlation Coefficient (*r*): The degree to which two continuous variables change together in a linear fashion. The correlation coefficient ranges from −1 to 1. 1 implies a perfect correlation, 0 implies no correlation, and −1 implies a perfect inverse correlation. For example, there is a positive correlation between height and forced expiratory volume (FEV), which increases as height increases. The square of the correlation coefficient is the coefficient of determination (R^2), which takes on a value between 0 and 1, and is a measure of how much the change in the dependent (y) variable is determined by the change in the independent (x) variable.

Causality

Although tempting, it is important to never assume that an association implies causation (i.e., "a" causes "b"). This simplification is often not the case because of the following:

- **Inverse causation**: "b" could actually be causing "a." One might conclude that acquired immunosuppression predisposes people to lentivirus infection; however, the lentivirus (HIV) is actually the cause of the immunosuppression (AIDS).
- **Confounding**: A third variable "c" could be affecting both measured variables "a" and "b." For example, gingivitis may not cause diabetes, but lack of access to health care may predispose someone to both conditions.
- **Other sources of bias** (see below).

Correlation does not imply causation.

Hypothesis Testing

Null hypothesis (H_o): The hypothesis of no association between two variables. For example, "a given treatment has no effect" or "two groups have identical risks despite different exposures." When constructing an experimental design, one attempts to statistically accept or reject the null hypothesis. The null hypothesis is paired with the **alternative hypothesis (H_a),** which takes the opposite assertion. For example, in a study comparing metronidazole to placebo for the treatment of giardiasis, the null hypothesis asserts that there would be no increased rate of resolution, whereas the alternative would assert that metronidazole hastens recovery.

Alpha (Type 1) Error: Rejecting the null hypothesis when it is true, creating a "false alarm." Mnemonic: Think **A** for **a**larm and type 1 because A is the first letter of the alphabet. For example, a study finds the use of vitamin C improves recovery from a URI when, in fact, it does not. Alpha is sometimes called the level of significance because it is the predetermined level below which the differences are considered unlikely to be due to chance alone and the null hypothesis is rejected. Usually, the alpha is set at 0.05.

Beta (Type 2) Error: Failing to reject the null hypothesis when it is false, creating a "missed detection." For example, a study finds that there is no significant decrease in mortality for patients who regularly exercise, when, in fact, there is.

Alpha error: false alarm
Beta error: missed detection

Power: The ability of a test to reject the null hypothesis when it is false. Otherwise stated, the probability of avoiding type 2 error. Power = 1 − beta. Power is increased with increased effect size and increased sample size.

P-Value: The probability of obtaining a test statistic (such as t-test or chi-square test statistic) as extreme or more extreme by chance alone if the null hypothesis is true and there is no bias. A P-value of less than 0.05 is usually said to be statistically significant.

Confidence Interval: A range of values around the point estimate such that, with repetition, the true value will be contained with a specified probability of 1 – α. Most often, the 95% confidence interval is reported, corresponding to an alpha level of 0.05. Increasing the sample size will narrow the confidence interval. A test in which the 95% confidence interval contains the "true" value is considered **accurate**. A test with a narrow confidence interval is considered **precise.**

If the 95% confidence interval does not contain the **null result, then there is a statistically significant difference in the groups**. The null result depends on the test. Rather than memorizing these results, think about which results would imply a difference and which would not. For ratios (RR, OR), the null value will be 1, whereas for differences, the null value will be 0.

Increasing sample size will narrow the confidence interval.

Clinical Study Design

- **Experimental study:** The investigator controls the exposure assignment. An example is randomized control trials.
- **Observational study:** The investigator observes the subjects without intervention. Cross-sectional, case-control, and cohort studies are observational studies.
- **Cross-sectional study:** Subjects are enrolled without regard to exposure and disease status, which are then evaluated simultaneously. Most often a survey. A cross-sectional study designed to determine the number of people with disease at a given time is a prevalence study.
  - *Example:* How many people in the United States have AIDS?
  - *Example:* How many people who have hyperlipidemia also currently have coronary artery disease?
- **Case-control study:** Subjects are enrolled based on disease status, one group with disease (cases) and one group without (controls), and then exposure is assessed in the two groups. Case-control studies are retrospective studies where disease status is known before exposure assessment.
  - *Example:* Select subjects with and without mesothelioma then ascertain the proportion of each group previously exposed to asbestos.
- **Cohort study:** Subjects are enrolled based on exposure status, one cohort with the exposure and one without (controls), and followed over time for the disease of interest. Subjects must be free of the disease at enrollment. The study may be prospective (disease status is not known at the time of enrollment) or retrospective (chart review). The incidence in each group can be calculated. Relative risks are calculated for effect measure estimation.
  - *Example:* Does having elevated cholesterol increase your chances of having a myocardial infarction?
    - Prospective design: Patients with and without high cholesterol are enrolled and followed over time to see if they develop heart disease.
    - Retrospective design: Patients with and without high cholesterol are identified from 10-year-old hospital records. Their charts are then reviewed through the present date to determine whether they developed heart disease.
- **Randomized control trial (RCT):** Subjects without the outcome of interest are enrolled and then randomly assigned by the investigator to either the exposed or unexposed group. The groups are followed prospectively for the outcomes of interest. The advantage of randomization is that it makes the groups similar in characteristics other than the exposure of interest. Randomization makes systematic error (bias) random. It can fail to sufficiently control bias if the sample size is small or if there is differential loss to follow-up (selection bias). Clinical drug trials are the main example of RCTs. Randomized studies cannot ethically be used to assess interventions thought to be harmful. For example, you cannot randomly assign someone to start smoking.
  - *Example:* Are patients randomly assigned to receive statins less likely to have an MI than those receiving a placebo?
- **Crossover study:** A type of prospective study, usually an RCT but possibly a cohort study, in which each patient begins in either the control or treatment group and then crosses over to the other group. In this way, every patient serves as his or her own control.

Case-control studies select participants already diagnosed with the disease of interest.

Use odds ratios in case-control studies.

Cohort studies follow participants who have an exposure and follow them over time.