

Descriptive Statistics: Reporting the Answers to the 5 Basic Questions of Who, What, Why, When, Where, and a Sixth, So What?

Thomas R. Vetter, MD, MPH

Descriptive statistics are specific methods basically used to calculate, describe, and summarize collected research data in a logical, meaningful, and efficient way. Descriptive statistics are reported numerically in the manuscript text and/or in its tables, or graphically in its figures. This basic statistical tutorial discusses a series of fundamental concepts about descriptive statistics and their reporting. The mean, median, and mode are 3 measures of the center or central tendency of a set of data. In addition to a measure of its central tendency (mean, median, or mode), another important characteristic of a research data set is its variability or dispersion (ie, spread). In simplest terms, variability is how much the individual recorded scores or observed values differ from one another. The range, standard deviation, and interquartile range are 3 measures of variability or dispersion. The standard deviation is typically reported for a mean, and the interquartile range for a median. Testing for statistical significance, along with calculating the observed treatment effect (or the strength of the association between an exposure and an outcome), and generating a corresponding confidence interval are 3 tools commonly used by researchers (and their collaborating biostatistician or epidemiologist) to validly make inferences and more generalized conclusions from their collected data and descriptive statistics. A number of journals, including *Anesthesia & Analgesia*, strongly encourage or require the reporting of pertinent confidence intervals. A confidence interval can be calculated for virtually any variable or outcome measure in an experimental, quasi-experimental, or observational research study design. Generally speaking, in a clinical trial, the confidence interval is the range of values within which the true treatment effect in the population likely resides. In an observational study, the confidence interval is the range of values within which the true strength of the association between the exposure and the outcome (eg, the risk ratio or odds ratio) in the population likely resides. There are many possible ways to graphically display or illustrate different types of data. While there is often latitude as to the choice of format, ultimately, the simplest and most comprehensible format is preferred. Common examples include a histogram, bar chart, line chart or line graph, pie chart, scatterplot, and box-and-whisker plot. Valid and reliable descriptive statistics can answer basic yet important questions about a research data set, namely: “Who, What, Why, When, Where, How, How Much?” (*Anesth Analg* 2017;125:1797–802)

I keep six honest serving-men;

(They taught me all I knew)

Their names are What and Why and When

And How and Where and Who.

Rudyard Kipling (1902), “The Elephant’s Child” from

Just So Stories

Descriptive statistics are specific methods basically used to calculate, describe, and summarize collected research data in a logical, meaningful, and efficient way. Descriptive statistics are reported numerically in the manuscript text and/or in its tables, or graphically in its figures.^{1–3}

From the Department of Surgery and Perioperative Care, Dell Medical School at the University of Texas at Austin, Austin, Texas.

Accepted for publication August 4, 2017.

Funding: None.

The author declares no conflicts of interest.

Reprints will not be available from the author.

Address correspondence to Thomas R. Vetter, MD, MPH, Department of Surgery and Perioperative Care, Dell Medical School at the University of Texas at Austin, Health Discovery Bldg, Room 6.812, 1701 Trinity St, Austin, TX 78712. Address e-mail to thomas.vetter@austin.utexas.edu.

Copyright © 2017 International Anesthesia Research Society

DOI: 10.1213/ANE.0000000000002471

As insightfully observed by Grimes and Schulz,⁴ “Descriptive studies often represent the first scientific toe in the water in new areas of inquiry. A fundamental element of descriptive reporting is a clear, specific, and measurable definition of the disease or condition in question. Like newspapers, good descriptive reporting answers the 5 basic W questions: who, what, why, when, where...and a sixth: so what?” This basic statistical tutorial discusses the following fundamental concepts about descriptive statistics and their reporting:

- Total study sample size versus study group sizes
- Study sample point estimate
- Frequency, percentage, ratio, and proportion
- Measures of the central tendency of data
- Measures of the variability or dispersion of data
- Confidence interval (CI) as measure of the precision of a point estimate
- Graphically displaying different types of data

TOTAL STUDY SAMPLE SIZE VERSUS STUDY GROUP SIZES

An understanding of statistics is predicated on the key distinction between a study sample and the population from which it

is selected, in that conventional inferential statistics allows one to extrapolate from the descriptive statistics of a sample to infer something or to make conclusions about the study population.^{3,5-7}

For example, in a clinical study like a randomized controlled trial, the statistical analysis of the descriptive data from a random sample of patients is used to make inferences and conclusions about the larger population and similar other, future patients.⁵

Prospectively determining the appropriate study sample size and the study group or subsample sizes is crucial in conducting a valid and ethical quantitative research study.⁸⁻¹⁰ The mechanics of doing so will be the subject of a future statistical tutorial.

When reporting the descriptive findings of a research study, it is very important to provide (a) the total sample size (ie, the total number of subjects sampled) and (b) the study group sizes (ie, the number of subjects sampled in each study group or subsamples) in the text, tables, and figures.^{11,12}

SAMPLE POINT ESTIMATE

Based on a sample of data that is randomly collected at a given point in time, an estimated value can be generated. Such a sample point estimate is a single value that can be used to validly estimate the corresponding population parameter.^{3,13-15} A series of repeated, random (unbiased) samples of data collected from the same underlying population would be expected to generate different yet equally valid point estimates of the corresponding population parameter.¹³⁻¹⁵

FREQUENCY, PERCENTAGE, RATIO, AND PROPORTION

As discussed in the previous statistical tutorial, some demographic and clinical characteristics can be parsed into and described using separate, discrete categories. Such categorical data can be either dichotomous (2 categories) or polytomous (more than 2 categories).^{2,16-19}

Dichotomous and polytomous categorical data can be described as (a) the raw counts or absolute frequencies (eg, "50") of the categories or (b) the percentages or relative frequencies (eg, "50%") of the categories. Relative frequencies can also be reported as ratios (eg, "50:50"), proportions (eg, "50/100"), or decimals (eg, "0.50").

Provide Numerators and Denominators for All Reported Percentages or Proportions

When reporting any observed percentage or proportion in the abstract, text, and/or tables of a manuscript, the authors should provide the corresponding numerator and denominator.

MEASURES OF THE CENTRAL TENDENCY OF DATA

The mean, median, and mode are 3 measures of the center or central tendency of a set of data.^{5,14,20-22} When a set of data is truly normally distributed, its distribution forms or follows a bell-shaped curve, and its mean, median, and mode are all at the same point, at the center of the distribution.²³ However, this is seldom the case, but instead these 3 values vary (Figure 1).

Mean

The mean is likely the most widely known descriptive statistic. Dating back to their earlier days as high-achieving

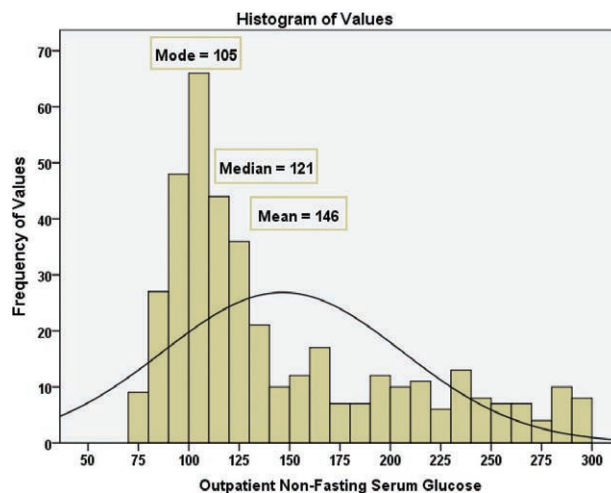


Figure 1. A hypothetical example of a histogram displaying a discordant mean, median, and mode values for a skewed, nonnormally (non-Gaussian) distributed data set.

students, promising, future researchers were aware of their average score in a class and their grade-point average, as well as their score versus the average score on a standardized aptitude test. All of these average values were an arithmetic mean, which is defined as the total sum of the values divided by the number of observations (sample size).²¹⁻²⁴

The symbol for the mean of a sample is \bar{x} , which is pronounced "x bar." The symbol for the mean of a population is the Greek letter μ , which is pronounced "mew" and spelled "mu" (in English).²⁰

The mean is typically reported for continuous (interval or ratio) data that have a normal (Gaussian) distribution.¹⁹ The mean can be greatly affected by or is very sensitive to outlying values ("outliers"), especially if they are extreme.²¹

Median

The median is a middle number or a value in a sample or population. It is the number above and below which there are an equal numbers of data points. For example, it is the midpoint in a set of test scores, which marks the 50th percentile of their data distribution. If there is an even number of values, the median is the average of the 2 middle ones.²⁰⁻²⁴ While there is no one agreed on symbol for the median, it can be notated as P_{50} or Mdn .^{20,23}

The median is typically reported for ordinal data or continuous data that do not have a normal (Gaussian) distribution.¹⁹ The median is not unduly affected by outlying values ("outliers"), unless they are excessive.²¹

Mode

The mode is the discrete number or integer that occurs most commonly or frequently in the data set. Most data distributions have only 1 mode and are thus referred to as unimodal. If a data distribution has 2 modes (2 "frequency peaks"), it is referred to as bimodal. There is no symbol for the mode.²⁰⁻²⁴

MEASURES OF THE VARIABILITY OR DISPERSION OF DATA

In addition to a measure of its central tendency (mean, median, or mode), another important characteristic of a

research data set is its variability or dispersion (ie, spread).²⁵ In simplest terms, variability is how much the individual recorded scores or observed values differ from one another.²⁶

The most general measure of variability is the total range, namely, the absolute difference between the minimum (lowest, smallest) and the maximum (highest, largest) recorded values.^{22,25,26}

Of note, this observed total range of recorded scores or observed values should be distinguished from the possible range of scores of values for an applied test or measurement instrument (eg, 0–10 on the 11-point numerical rating scale for unidimensional pain intensity). Authors should always provide the possible range of values for any applied test or measurement instrument when describing it in the methods section of their manuscript.

Standard Deviation Accompanies a Mean

The standard deviation (SD) is typically reported for a mean. The SD represents the average amount of variability in a set of values or scores, and practically or roughly speaking, the SD is the average distance from the mean.^{25,26}

The larger the SD, the larger the average distance each data point is from the mean of the distribution of the overall data set, and hence, the more dispersed or spread apart the entire data set.²⁶

The SD of a sample is inversely related to the square root of the sample size (actually, $n - 1$), so the larger the sample size, the relatively smaller the SD.^{25–27} However, if the data have a normal (Gaussian) distribution, 68.3% of the recorded and analyzed values always fall within ± 1 SD of the corresponding mean, 95.4% within ± 2 SD of that mean, and 99.7% within ± 3 SD of that mean.²⁸

The standard deviation is typically abbreviated as “SD” or “s” for a sample variable and as the Greek letter σ for a population parameter. The SD is expressed in the same units as the primary data.²⁷ The SD can be reported with the sample mean as its plus/minus value (eg, 77.2 ± 12.5) or as its absolute value (eg, $77.2 [12.5]$).

As an aside, the SD is technically the square root of the variance, another measure of the variability of a data set, but the variance is not conventionally reported mainly because its units are expressed as the square (x^2) of the primary data.^{25–27}

Interquartile Range Accompanies a Median

The interquartile range (IQR) is another commonly reported measure of the variability or dispersion (spread) of the values in a data set. The IQR is typically reported for a median.

As its name implies, the IQR includes the range of values between the 25th percentile (the first quartile, Q_1) and the 75th percentile (the third quartile, Q_3).^{22,25,27} Operationally, all of the individual recorded scores or observed values are divided into 4 equally sized quartiles (“buckets”), and the IQR includes the 2 middle quartiles.²⁵

The interquartile range is typically abbreviated as “IQR.” The IQR is expressed in the same units as the primary data.²⁷ The specific values for the 25th percentile (Q_1) and the 75th percentile (Q_3) are actually reported (eg, “IQR: 35, 65”).

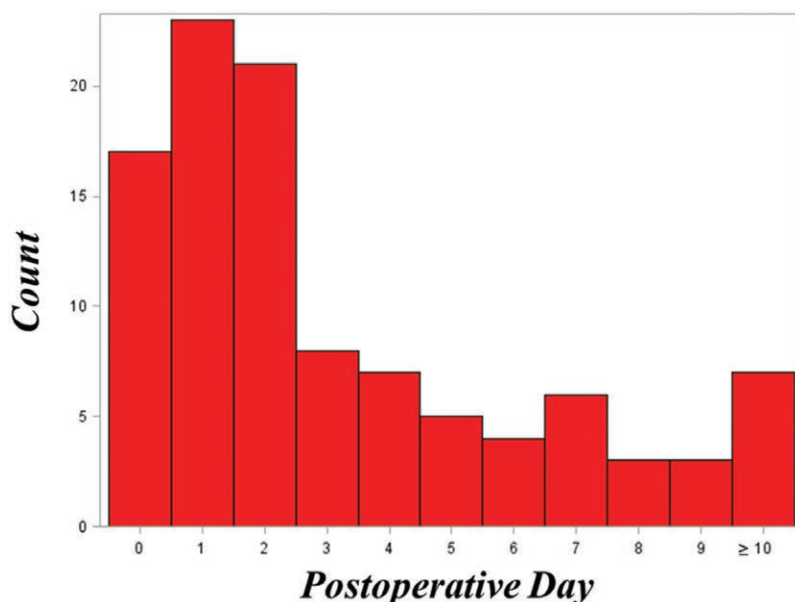
Authors can also report the absolute range of the recorded scores or observed values for a variable, especially if it highlights an important aspect and/or provides additional insight.

CI AS MEASURE OF THE PRECISION OF A POINT ESTIMATE

Testing for statistical significance,^{29–31} calculating the observed treatment effect size (or the strength of the association between an exposure and an outcome),³² and generating a corresponding CI^{33,34} are 3 tools commonly used by researchers (and their collaborating biostatistician or epidemiologist) to validly make inferences and more generalized conclusions from their collected data and descriptive statistics.⁷ A number of journals, including *Anesthesia & Analgesia*, strongly encourage or require the reporting of pertinent CIs.^{35,36} The important concept of the CI will be introduced here and expanded on in a future statistical tutorial.

As noted above, a series of repeated, random samples of data collected from the same underlying population would be expected to generate different yet equally valid

Figure 2. Example of a histogram. Histogram of timing of postoperative stroke. Reproduced with permission from Hsieh et al.⁴⁸



point estimates (eg, sample means, \bar{x}_1 , \bar{x}_2 , \bar{x}_3 , etc) of the corresponding true population parameter (eg, the population mean, μ).^{13–15} Likewise, a series of clinical trials would each be expected to generate a different observed treatment effect (eg, the reduction in the frequency or incidence of postoperative myocardial infarction). Which of these sample point estimates is the most accurate (“correct”) estimate of the corresponding population parameter? One cannot ever be certain, but this is where a CI comes into play.

A CI can be calculated for virtually any variable or outcome measure in an experimental, quasi-experimental, or observational research study design.^{37,38} Specifically, a CI can be calculated for a single observed mean or proportion.^{39,40} It can also be calculated for an observed difference between 2 subgroups or subsample means or proportions, which is referred to in a clinical trial as the observed treatment difference or effect.⁴¹

Generally speaking, in a clinical trial, the CI is the range of values within which the true treatment effect in the population likely resides.^{42,43} In an observational study, the CI is the range of values within which the true strength of the association between the exposure and the outcome (eg, the risk ratio or odds ratio) in the population likely resides.^{44,45}

The CI is formally defined statistically as follows: “if the level of confidence is set at 95%, it means that if data collection and analysis could be replicated many times, the CI should include within it the correct value of the measure 95% of the time.”⁴⁴ Stated another way, if, for instance, one performed the data sampling process from the same population 100 separate times, generating 100 distinct sample point estimates, and then 100 corresponding CIs are calculated, 95% of these intervals are expected to contain the true (correct) value.^{15,43}

The confidence interval is typically abbreviated as “CI.” A CI can be calculated for any degree of confidence. While the 95% CI is traditionally and most commonly reported, a 90% CI or a 99% CI may be more relevant and instead reported. No matter its level (0%–100%) and its range or

width, a CI has a lower boundary (limit) and an upper boundary (limit).^{39,44}

Counterintuitively, for the same data set and its given sample size, the lower boundary (limit) and upper boundary (limit) of its more precise CI (eg, for its 99% CI versus its 95% CI) are wider apart.^{7,43,46}

There is an inverse relationship between width of the CI and sample size, so the larger the relative sample size, the narrower and more precise the CI.^{7,43,46} Thus, a prime motivator for conducting a systematic review and then combining (“pooling”) the identified individual study data via a meta-analysis is to create a larger sample size and to generate a more precise pooled estimate of the treatment effect.⁴²

GRAPHICALLY DISPLAYING DIFFERENT TYPES OF DATA

There are many ways to graphically display or illustrate different types of data, “A picture is really worth a thousand words.”⁴⁷ While there is often latitude as to the choice of format, ultimately, the simplest and most comprehensible format is preferred.

Common examples of graphics include a stem-and-leaf plot, histogram (Figure 2), bar chart (Figure 3), pie graph (Figure 4), line graph (Figure 5), scatter plot (Figure 6), and box-and-whisker plot (Figure 7).^{22,47,53,54} Of note, with few exceptions, pie charts are generally suboptimal as the same information can better presented visually with a bar chart.⁵⁶

The box-and-whisker plot is specifically intended to graphically display nonnormally (non-Gaussian) distributed continuous (interval or ratio) data.^{19,24} The box-and-whisker plot displays the following: (a) the median (Mdn) value (solid line within the box); (b) the IQR (2 ends of the box); (c) either the 5th–95th percentile values or the smallest and largest observations (2 lines with caps extending from the box, the so-called whiskers); and (d) any extreme

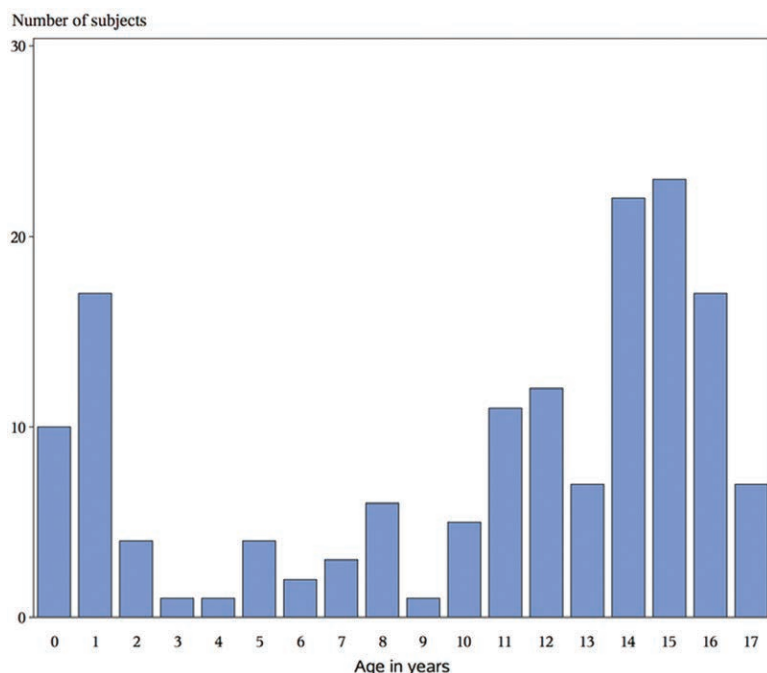


Figure 3. Example of a bar chart. The distribution of subjects by year of age. Reproduced with permission from Spielberg et al.⁴⁹

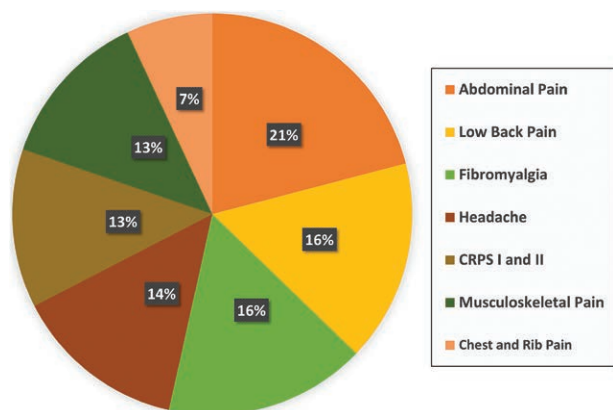


Figure 4. Example of a pie graph. The most common outpatient pediatric pain locations or diagnoses. CRPS indicates complex regional pain syndrome. Data were derived from Vetter.⁵⁰

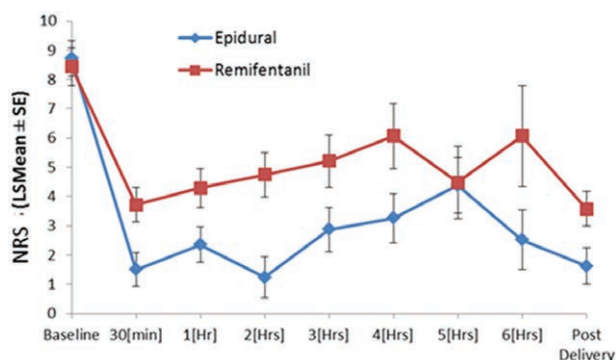


Figure 5. Example of a line graph. The figure depicts the model estimated mean (with error bars) changes from baseline for both study groups at each time point. NRS indicates numerical rating scale. Reproduced with permission from Stocki et al.⁵¹

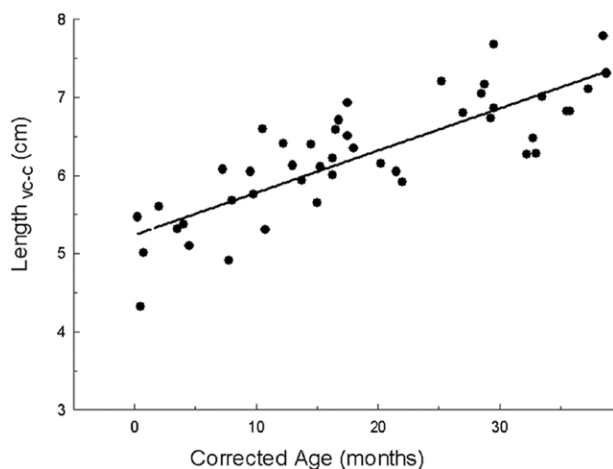


Figure 6. Example of a scatter plot. Relationship between the distance between the vocal cords and carina (length VC-C) and corrected age in 46 children ($r = 0.84$). VC-C indicates vocal cords-carina. Reproduced with permission from Sirisopana et al.⁵²

outliers located beyond the so-called lower fence ($Q1 - 1.5 \times [IQR]$) or the “upper fence” ($Q3 + 1.5 \times [IQR]$).^{22,24}

The reader is referred to the most current version of the publication and style manuals of the American Medical Association¹² and the American Psychological

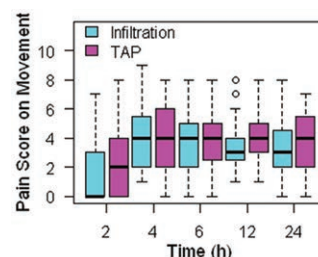


Figure 7. Example of a box-and-whisker plot (box plot). Box plots of pain scores on movement assessed using an 11-point numerical rating scale (0 = no pain and 10 = the worst possible pain) over time after cesarean delivery (transversus abdominis plane block). TAP indicates transversus abdominis plane. Reproduced with permission from Tawfik et al.⁵⁵

Association^{11,57,58} for in-depth discussions of how to display or illustrate research data tabularly or graphically.

CONCLUSIONS

Who? What? Why? When? Where? How? How Much? These are 7 questions kids learn in grade school or when first learning a foreign language. Like other baby boomers, no matter how rusty my trusty French may be, I take comfort that I can still remember these 7 basic words: Qui, quoi, pourquoi, quand, où, comment, and combien. They cover the basics and help you understand most situations and contexts.⁵⁹ And so it is with descriptive statistics! ■■

DISCLOSURES

Name: Thomas R. Vetter, MD, MPH.

Contribution: This author wrote and revised the manuscript.

This manuscript was handled by: Jean-Francois Pittet, MD.

REFERENCES

1. Salkind NJ. Statistics or sadistics? It's up to you. In: *Statistics for People Who (Think They) Hate Statistics*. 6th ed. Thousand Oaks, CA: Sage Publications, 2016:5–18.
2. Urdan TC. Introduction to social science research principles and terminology. In: *Statistics in Plain English*. 4th ed. New York, NY: Routledge, Taylor & Francis Group, 2017:1–12.
3. Mendenhall W, Beaver RJ. An invitation to statistics. In: *Introduction to Probability and Statistics*. Boston, MA: CL-Wadsworth, 2013:1–6.
4. Grimes DA, Schulz KF. Descriptive studies: what they can and cannot do. *Lancet*. 2002;359:145–149.
5. Motulsky H. From sample to population. In: *Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking*. New York, NY: Oxford University Press, 2014:22–28.
6. Salkind NJ. Significantly significant. In: *Statistics for People Who (Think They) Hate Statistics*. 6th ed. Thousand Oaks, CA: Sage Publications, 2016:177–196.
7. Urdan TC. Statistical significance, effect size, and confidence intervals. In: *Statistics in Plain English*. 4th ed. New York, NY: Routledge, Taylor & Francis Group, 2017:73–91.
8. Motulsky H. Choosing the sample size. In: *Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking*. New York, NY: Oxford University Press, 2014:216–229.
9. Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. *Lancet*. 2005;365:1348–1353.
10. Pandis N, Polychronopoulou A, Eliades T. Sample size estimation: an overview with applications to orthodontic clinical trial designs. *Am J Orthod Dentofacial Orthop*. 2011;140:e141–e146.
11. American Psychological Association. *Publication Manual of the American Psychological Association*. 6th ed. Washington, DC: American Psychological Association; 2009.

12. Iverson C, American Medical Association. *American Medical Association Manual of Style: A Guide for Authors and Editors*. 10th ed. New York, NY: Oxford: Oxford University Press; 2007.
13. Brown S. Estimating population parameters. *Stats Without Tears*. 2017. Available at: <https://brownmath.com/swt/chap09.htm>. Accessed July 7, 2017.
14. Brown S. Statistics! *Stats Without Tears*. 2017. Available at: <https://brownmath.com/swt/chap01.htm>. Accessed July 7, 2017.
15. Mendenhall W, Beaver RJ. Large sample estimation. In: *Introduction to Probability and Statistics*. Boston, MA: CL-Wadsworth, 2013:281–323.
16. Field A. Why is my evil lecturer forcing me to learn statistics? In: *Discovering Statistics Using IBM SPSS Statistics: And Sex and Drugs and Rock 'n' Roll*. Los Angeles, CA: Sage, 2013:1–39.
17. Campbell MJ, Swinscow TDV. Data display and summary. In: *Statistics at Square One*. Chichester, UK; Hoboken, NJ: Wiley-Blackwell/BMJ Books, 2009:1–13.
18. Hulley SB, Newman TB, Cummings SR. Planning the measurements: precision, accuracy, and validity. In: Hulley SB, Cummings SR, Browner WE, Grady DG, Newman TB, eds. *Designing Clinical Research*. 4th ed. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2013:32–42.
19. Vetter TR. Fundamentals of research data and variables: the devil is in the details. *Anesth Analg*. 2017;125:1375–1380.
20. Brown S. Numbers about numbers. *Stats Without Tears*. 2017. Available at: <https://brownmath.com/swt/chap03.htm>. Accessed July 7, 2017.
21. Salkind NJ. Means to an end: computing and understanding averages. In: *Statistics for People Who (Think They) Hate Statistics*. 6th ed. Thousand Oaks, CA: Sage Publications, 2016:177–196.
22. Mendenhall W, Beaver RJ. Describing data with numerical measures. In: *Introduction to Probability and Statistics*. Boston, MA: CL-Wadsworth, 2013:50–93.
23. Urdan TC. Measures of central tendency. In: *Statistics in Plain English*. 4th ed. New York, NY: Routledge, Taylor & Francis Group, 2017:13–20.
24. Motulsky H. Graphing continuous data. In: *Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking*. New York, NY: Oxford University Press, 2014:61–71.
25. Urdan TC. Measures of variability. In: *Statistics in Plain English*. 4th ed. New York, NY: Routledge, Taylor & Francis Group, 2017:21–32.
26. Salkind NJ. Vive la différence: understanding variability. In: *Statistics for People Who (Think They) Hate Statistics*. 6th ed. Thousand Oaks, CA: Sage Publications, 2016:43–57.
27. Motulsky H. Quantifying scatter. In: *Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking*. New York, NY: Oxford University Press, 2014:77–84.
28. Motulsky H. The Gaussian distribution. In: *Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking*. New York, NY: Oxford University Press, 2014:85–89.
29. Woolson RF, Kleinman JC. Perspectives on statistical significance testing. *Annu Rev Public Health*. 1989;10:423–440.
30. Hayat MJ. Understanding statistical significance. *Nurs Res*. 2010;59:219–223.
31. Pocock SJ, Ware JH. Translating statistical findings into plain English. *Lancet*. 2009;373:1926–1928.
32. McGough JJ, Faraone SV. Estimating the size of treatment effects: moving beyond p values. *Psychiatry (Edmont)*. 2009;6:21–29.
33. Fethney J. Statistical and clinical significance, and how to use confidence intervals to help interpret both. *Aust Crit Care*. 2010;23:93–97.
34. Akobeng AK. Confidence intervals and p-values in clinical decision making. *Acta Paediatr*. 2008;97:1004–1007.
35. Altman D. Estimating with confidence. In: Gardner M, Altman D, eds. *Statistics With Confidence: Confidence Intervals and Statistical Guidelines*. 2nd ed. London: BMJ Books, 2000:3–5.
36. Altman D. Confidence intervals in practice. In: Altman D, Machin D, Bryant T, Gardner M, eds. *Statistics With Confidence: Confidence Intervals and Statistical Guidelines*. 2nd ed. London: BMJ Books, 2000:6–14.
37. Vetter TR. Magic mirror, on the wall-which is the right study design of them all?-part II. *Anesth Analg*. 2017;125:328–332.
38. Vetter TR. Magic mirror, on the wall-which is the right study design of them all?-part I. *Anesth Analg*. 2017;124:2068–2073.
39. Motulsky H. Confidence interval of a proportion. In: *Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking*. New York, NY: Oxford University Press, 2014:29–42.
40. Motulsky H. Confidence interval of a mean. In: *Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking*. New York, NY: Oxford University Press, 2014:95–103.
41. Sedgwick P. Randomised controlled trials: inferring significance of treatment effects based on confidence intervals. *BMJ*. 2014;349:g5196.
42. Montori VM, Kleinbart J, Newman TB, et al. Evidence-Based Medicine Teaching Tips Working Group. Tips for learners of evidence-based medicine: 2. Measures of precision (confidence intervals). *CMAJ*. 2004;171:611–615.
43. Altman D. Confidence intervals rather than P values. In: Altman D, Machin D, Bryant T, Gardner M, eds. *Statistics With Confidence: Confidence Intervals and Statistical Guidelines*. 2nd ed. London: BMJ Books, 2000:15–27.
44. Rothman KJ. Random error and the role of statistics. In: *Epidemiology: An Introduction*. 2nd ed. Oxford: Oxford University Press, 2012:148–163.
45. Rothman KJ, Greenland S, Lash TL. Precision and statistics in epidemiologic studies. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. Revised 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins, 2012:148–167.
46. Daly L. Confidence intervals and sample sizes. In: Altman D, Machin D, Bryant T, Gardner M, eds. *Statistics With Confidence: Confidence Intervals and Statistical Guidelines*. 2nd ed. London: BMJ Books; 2000:139–152.
47. Salkind NJ. A picture is really worth a thousand words. In: *Statistics for People Who (Think They) Hate Statistics*. 6th ed. Thousand Oaks, CA: Sage Publications, 2016:59–79.
48. Hsieh JK, Dalton JE, Yang D, Farag ES, Sessler DI, Kurz AM. The association between mild intraoperative hypotension and stroke in general surgery patients. *Anesth Analg*. 2016;123:933–939.
49. Spielberg DR, Barrett JS, Hammer GB, et al. Predictors of arterial blood pressure control during deliberate hypotension with sodium nitroprusside in children. *Anesth Analg*. 2014;119:867–874.
50. Vetter TR. A clinical profile of a cohort of patients referred to an anesthesiology-based pediatric chronic pain medicine program. *Anesth Analg*. 2008;106:786–794.
51. Stocki D, Matot I, Einav S, Eventov-Friedman S, Ginosar Y, Weiniger CF. A randomized controlled trial of the efficacy and respiratory effects of patient-controlled intravenous remifentanyl analgesia and patient-controlled epidural analgesia in laboring women. *Anesth Analg*. 2014;118:589–597.
52. Sirisopana M, Saint-Martin C, Wang NN, Manoukian J, Nguyen LH, Brown KA. Novel measurements of the length of the subglottic airway in infants and young children. *Anesth Analg*. 2013;117:462–470.
53. Mendenhall W, Beaver RJ. Describing data with graphs. In: *Introduction to Probability and Statistics*. Boston, MA: CL-Wadsworth, 2013:7–49.
54. Mendenhall W, Beaver RJ. Describing bivariate data. In: *Introduction to Probability and Statistics*. Boston, MA: CL-Wadsworth, 2013:94–122.
55. Tawfik MM, Mohamed YM, Elbadrawi RE, Abdelkhalek M, Mogahed MM, Ezz HM. Transversus abdominis plane block versus wound infiltration for analgesia after cesarean delivery: a randomized controlled trial. *Anesth Analg*. 2017;124:1291–1297.
56. Hickey W. The worst chart in the world. *Business Insider: Markets* 2013. Available at: <http://www.businessinsider.com/pie-charts-are-the-worst-2013-6>. Accessed July 27, 2017.
57. Nicol AAM, Pexman PM. *Displaying Your Findings: A Practical Guide for Creating Figures, Posters, and Presentations*. 6th ed. Washington, DC: American Psychological Association; 2013.
58. Nicol AAM, Pexman PM. American psychological association. In: *Presenting Your Findings: A Practical Guide for Creating Tables*. Washington, DC: American Psychological Association, 2010.
59. 7 key questions: who, what, why, when, where, how, how much? *Consultants Mind: Thinking Through the Problem*. 2013. Available at: <http://www.consultantsmind.com/2013/03/07/7-key-questions-who-what-why-when-where-how-how-much/>. Accessed July 7, 2017.