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 Much?” (Anesth Analg 2017;XXX:00–00)
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.6-10 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.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.15-15

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).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.14,19 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

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

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

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.

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 $\mu$, which is pronounced “mew” and spelled “mu” (in English).20

The mean can be greatly affected by or is very sensitive to outlying values (“outliers”), especially if they are extreme.21

Mean
The mean is likely the most widely known descriptive statistic. Dating back to their earlier days as high-achieving students, promising, future researchers were aware of their
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 “σ” 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

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 (x²) 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, Q1) and the 75th percentile (the third quartile, Q3).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 (Q1) and the 75th percentile (Q3) 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 CI33,34 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 point estimates (eg, sample means, x̅₁, x̅₂, x̅₃, etc) of the corresponding true population parameter (eg, the population mean, µ).
mean, \( \mu \)). 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. Specifically, a CI can be calculated for a single observed mean or proportion. 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.

Generally speaking, in a clinical trial, the CI is the range of values within which the true treatment effect in the population likely resides. 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.

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.” 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. 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).

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. 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.

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.” 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). Of note, with few exceptions, pie charts are generally suboptimal as the same information can better presented visually with a bar chart.

The box-and-whisker plot is specifically intended to graphically display nonnormally (non-Gaussian) distributed continuous (interval or ratio) data. 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 outliers located beyond the so-called lower fence (Q1 \(- 1.5 \times \text{IQR}\)) or the “upper fence” (Q3 \(+ 1.5 \times \text{IQR}\)).
The reader is referred to the most current version of the publication and style manuals of the American Medical Association and the American Psychological Association 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. 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

E SPECIAL ARTICLE


