Principles of Statistics

What the Sports Medicine Professional Needs to Know

Bryan L. Riemann, PhD, ATC, FNATA[a,*],
Monica R. Lininger, PhD, LAT, ATC[b]

KEYWORDS

• Clinical meaningfulness • *P* values • Data interpretation • Confidence intervals
• Effect sizes • Statistical power • Minimal detectable difference
• Minimal important difference

KEY POINTS

- Statistical significance reflects the influence of chance on the outcome, whereas clinical meaningfulness reflects the degree to which the study results reported are relevant to sports medicine practice.
- When statistically significant differences are revealed, confidence intervals and effect sizes can be used to enhance the practical interpretation of the research results.
- Absolute reliability characteristics, such as the minimal detectable change, determine the extent of error around a measurement and, when coupled with an appropriate minimal important difference estimate, can assist in triangulating clinically meaningful changes in patients undergoing treatment.
- In the circumstance of nonstatistically significant results, evaluation needs to occur to determine if the study had adequate statistical power prior to concluding no difference or association exists.

Along with many disciplines in medicine and allied health, the evidence-based practice (EBP) movement has prompted practitioners in the field of sports medicine to have a better competency in understanding research. As new procedures, methods, and understanding are studied with the results presented in research studies, practicing sports medicine professionals are faced with evaluating both statistically significant and clinically meaningful benefit along with whether the results are pertinent to their patients. Unquestionably, interpreting statistical findings as part of the research

Disclosure Statement: No conflicts of interest pertaining to this work.
[a] Sports Medicine, Biodynamics and Human Performance Center, Georgia Southern University, 11935 Abercorn Street, Savannah, GA 31419, USA; [b] Physical Therapy and Athletic Training, Athletic Training Education Program, Northern Arizona University, PO Box 15094, Flagstaff, AZ 86011, USA
* Corresponding author.
E-mail address: briemann@georgiasouthern.edu

Clin Sports Med 37 (2018) 375–386
https://doi.org/10.1016/j.csm.2018.03.004
sportsmed.theclinics.com

evaluation process can be a daunting challenge for many practitioners. Few practitioners enter a field, such as sports medicine, wanting to develop and possess an extensive expertise in statistics. Fortunately, once command over some of the nomenclature and few key concepts, along with understanding the more common statistical procedures, is attained, practitioners can often begin to evaluate research. The purpose of this article is to provide minimal essentials a sports medicine professional needs to know about interpreting statistics and research results to facilitate the incorporation of the latest evidence into practice. Topics covered include the difference between statistical significance and clinical meaningfulness; effect sizes and confidence intervals; reliability statistics, including the minimal detectable difference (MDD) and minimal important difference (MID); and statistical power. To begin the discussion, some of the common research and statistical terms are presented in **Table 1**.

ROLE OF STATISTICS IN THE RESEARCH PROCESS

Research, the process of acquiring new knowledge through systematic data collection procedures followed by controlled and critical analysis of the data, is one of the foundations of EBP. It is the analysis of data part of this definition for which statistics become necessary. One of the first tasks after data are collected is to summarize the data so they may be reduced into smaller and more interpretable chunks. Descriptive statistics are those indicators that are used to portray the data in more interpretable chunks. The challenge is that no single descriptive statistic can represent an entire data set as a single value. For example, the mean change in range of motion can be described for 2 groups as 10°; however, closer inspection of the individual participants may yield a different perspective (**Table 2**). Thus, at minimum, the optimal approach is to provide some indicator of the data set center and the extent of variation around the center. The measurement scale of the data dictates the appropriate descriptive statistics to use (**Table 3**).

In research, it is not feasible to study all members of a defined population. Because research involves selecting a sample from a target population, assumptions and hypotheses must be examined to determine whether the obtained results are tenable or if they could have simply occurred due to chance. Additionally, a researcher often wants to generalize or make predictions about the population from which a sample was obtained. A population is defined as all individuals who meet the inclusion and exclusion criteria for a specific study. When a sample is drawn from the population, sampling error is likely incurred because it is probable that the sample does not perfectly represent (ie, duplicate) a population. Thus, estimates about what exists in the entire population based on the sample differ from the true reality. The magnitude of sampling error is attempted to be decreased through using optimal research design elements, including random sampling and random allocation (assignment), sufficiently sized samples, and reliable outcome measures. Despite the efforts to decrease sampling error, some uncertainty always exists; the challenge for researchers and practitioners alike is how to evaluate whether the difference/relationship results are real versus the likelihood that they occurred based on chance (ie, sampling error). By providing a *P* value, inferential statistics attempt to provide some indication regarding the extent to which chance might explain the results.

INTERPRETING *P* VALUES

Interpretation of the resulting *P* values from inferential statistics is an area of frequent confusion. Two perspectives on using *P* values have been described, significance testing and hypothesis testing. Over the years the 2 perspectives have often become