

## Editorial

### Words and Meaning in Scientific Reporting: Consecutive, Prospective, and Significant

Seth S. Leopold MD

Received: 21 June 2013 / Accepted: 1 July 2013 / Published online: 12 July 2013  
© The Association of Bone and Joint Surgeons® 2013

Scientific reporting is a tricky business. Seemingly innocent differences in the usage of common words can send misleading messages. Journal editors should be your advocates (we try), but it is important, as a thoughtful reader, to fend for oneself.

Some incorrect word choices occur through simple error, and some words have been misused so commonly for so long that they may even be losing their original meanings. But some words are misused purposefully, with the intention of causing readers to mistake a low-horsepower study for something more. Three of the most common “offenders” are the words consecutive, prospective, and significant.

**Consecutive.** The correct usage of this word describes the enrollment into a study of all patients for a particular diagnosis during a defined period of time. Used correctly, the term conveys that selection bias did not affect the decision of which patients to include. This is an important idea, and it is precisely for this reason that misuse of this word in clinical research is so troublesome.

---

The author certifies that he, or any members of his immediate family, has no commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

The opinions expressed are those of the writers, and do not reflect the opinion or policy of *CORR*® or the Association of Bone and Joint Surgeons®.

---

S. S. Leopold (✉)  
Clinical Orthopaedics and Related Research, 1600 Spruce Street,  
Philadelphia, PA 19103, USA  
e-mail: sleopold@clinorthop.org

When consecutive is misused, it actually obscures the effect of selection bias on the study in question. A report describing “consecutive patients who received my novel operation for ulnar-sided wrist pain” would allow me to treat any number of patients in other ways, perhaps with different procedures or even nonsurgically, and cherry-pick the best participants, such as nonsmokers, or those with no workers compensation claims, to receive my new technique. Using the word consecutive in this way distracts readers from what might indeed be very selective inclusion criteria, and this usage is likely to cause readers to overestimate the benefits of the treatment in question.

In my experience, this word rarely is used correctly in orthopaedic research.

**Prospective.** A prospective study poses its questions before any patients are enrolled or treated. Stated another way, a prospective study begins, and its ethical review takes place, before the clinical outcomes of interest have occurred, and before any data are recorded.

By contrast—and this misuse is common—a study that queries a prospectively maintained institutional database for outcomes that have already occurred is a retrospective study, not a prospective one. For many reasons, this distinction is not merely semantic. Such studies are prone to data dredging; if I query a database for 60 endpoints at the  $p < 0.05$  level, pure chance alone would provide me with about three findings that I could present as “statistically significant,” but that significance could well be spurious and misleading. In addition, the quality of the data collection tends to be better in truly prospective studies, the inclusion criteria tend to be clearer (and easier to maintain), the questions more focused, and observer bias tends to be less severe or at least easier to ascertain, as authors usually have predefined who will assess the endpoints of interest.

Retrospective studies remain the most common studies in our literature, and I believe they have an important role to play [2], but describing them as prospective is misleading and harmful.

**Significant.** In its purest statistical sense, a finding is significant when a statistical test rejects a claim of no difference between groups, the null hypothesis, at a certain level of probability, the p value. The p value estimates the probability that an effect as extreme as that observed might have occurred by chance only. Unfortunately, this term has taken on a life of its own, one not intended by those who first used it [3]. Using threshold values sometimes is simply the wrong approach, and the topic is considerably more nuanced than we generally appreciate [1]. But more importantly, there are reasons—beyond deviation from the original intent of the concept’s developers—that should cause us to approach the topic of “significance” with great caution.

Many clinicians are uncomfortable with the topic of statistics. Because of that, having an authoritative-looking threshold (the concept of significance) is embraceable, yet it can cause us not to look as deeply as we should at important, intuitive concepts that do not require a background in math or statistics. For example, most orthopaedic studies are small, with borderline significance. In such studies, if the outcome of even one or a few patients were to differ, the “all-important” p value would transition from significant to not. This is hardly the basis for sound clinical decision-making. Conversely, in large studies, a low

p value often distracts readers from the much more important question of whether small observed differences in effect size are clinically important, even if the differences are statistically significant. Finally, studies sometimes indicate that one treatment resulted in scores that were “higher, but not significantly higher,” and others claim “a trend that approached significance.” View these assertions with skepticism. They often violate the very premise on which the significance tests the authors performed are based, namely that the observed result should not likely be the result of chance.

The less frequently the word significant is used in a manuscript, the clearer that paper usually is.

Words matter, and the way we use them in scientific reporting affects how we care for our patients.

**Acknowledgments** The author would like to acknowledge Lee Beadling BA, Clare M. Rinnac PhD, and Raphaël Porcher PhD for their thoughtful suggestions, which improved this manuscript.

## References

1. Falissard B. Statistics in brief: when to use and when not to use a threshold p value. *Clin Orthop Relat Res.* 2012;470:315–316.
2. Leopold SS. Editorial: let’s talk about level IV: the bones of a good retrospective case series. *Clin Orthop Relat Res.* 2013; 471:353–354.
3. Neyman J, Pearson ES. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character.* 1933;231:289–337.