EDITORIAL | Cores of Reproducibility in Physiology

Cores of Reproducibility in Physiology (CORP): Advancing the corpus of physiological knowledge

Peter D. Wagner
University of California, San Diego La Jolla, California
Submitted 29 November 2016; accepted in final form 29 November 2016

This volume of the Journal of Applied Physiology marks publication of our first CORP article (1). It is a statistical essay on best practices to minimize the probability of studies coming up with either false positive or false negative conclusions. CORP (Cores of Reproducibility in Physiology) is an APS-wide initiative responsive to a growing concern in biological research. This concern is the frequency with which published outcomes cannot be replicated by other investigators.

Although such concerns may have initially arisen with data from antibody-based molecular techniques, they are equally problematic at all levels of physiological research. It is not a problem per se for two purportedly identical studies to disagree, because if both are adequately described, the reason(s) for discrepancies can usually be found, and this alone may considerably advance the field. It is when the outcomes are in doubt because of poor descriptions or statistical errors that we waste our time, resources, and dollars and put future research (and even clinical care) at risk.

There are many reasons for failure to replicate the work of others. Differences in technical aspects may be the most obvious, including those that can be ascribed to differences in hardware, in software, in reagents, in experimental subject types, or in experimental protocols. But a reason universal to all is failure to work within the bounds of good statistical practice.

In the current article (1), Dr. Curran-Everett focuses on perhaps the most fundamental point in statistical considerations: minimizing the uncertainty of one’s conclusions. He was not asked to invent anything new but to show the way to reduce the probability of a conclusion being uncertain or even wrong. Both when concluding that there is no difference between groups when there really is a difference (false negative) and the converse, namely, concluding that there is a difference between groups when there really is none (false positive).

Even using the simplest experimental design—comparison of two groups at a single time point—there is much more to analysis than the P value itself. The article is all about the relationships among 1) differences in average data between groups, 2) variance in those data, and 3) numbers of subjects studied (be they cells, tissues, organs, or intact systems). When you understand those relationships, proper experimental design is facilitated and uncertainty reduced.

One of the most informative yet more chilling elements of statistical consideration is expressed in Table 2 of the article (1). It is, for me, a major take home message. This table is saying the following. Suppose I accept P < 0.05 as the magic line in the sand, and I publish a paper revealing results at this significance level. Pretty good, yes? Only a 5% chance my conclusion is wrong. Now another research group—using EXACTLY the same protocol, reagents, experimental subject types and numbers, hardware, and software—repeats my experiment with equal expertise. Having another group confirm the findings is the essence of the term “reproducibility.” There is an equal probability that their attempt to duplicate my work will fail (they will find P > 0.05) or succeed (they will find P < 0.05). Think about that. There is no more than a 50/50 chance my competitor will agree with me, although I found P < 0.05. Is that good enough? Is 50/50 a good enough risk for your stock market investments? Remember, the risk is based only on chance, because all else was identical between the two studies. Even were the line in the sand drawn at P < 0.01, Table 2 (1) says that there is a substantial—27%—chance my competitor will disagree with me (she will find P > 0.05) for the wrong reasons: chance, rather than because I was wrong and she was right. And so on.

In the months ahead we are looking forward to several more invited CORP submissions. As you can see, these are directed at specific experimental procedures.

- Hydration and temperature assessment.
- Analysis of single muscle fiber contractile characteristics.
- Why $\dot{V}O_{2max}$?: $\dot{V}O_{peak}$ is no longer acceptable.
- Noninvasive assessment of endothelial function.
- NIRS: What can the latest technologies do and what are the limitations?
- EMG and force/voluntary activation measures.
- Measurement of respiratory impedance in animal models of pulmonary disease.
- Measurement of force and calcium transients using intact fibers from mammalian skeletal muscle.
- Measurement of reactive oxygen species in skeletal muscle using fluorescence probes.
- Blood, plasma volume, and RBC mass measurement.
- Measurement of hypoxic and hypercapnic ventilatory responses.

All CORP articles are invited, so should you have an idea for a CORP article on a specific topic or methodology relevant to the domain of Journal of Applied Physiology,
please send the idea to me (jappleic@ucsd.edu) along with the name(s) of the best author(s) for the article. The Associate and Consulting Editors will consider each proposal seriously.

AUTHOR CONTRIBUTIONS

P.D.W. drafted manuscript; P.D.W. edited and revised manuscript; P.D.W. approved final version of manuscript.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCE