UC San Diego

UC San Diego Previously Published Works

Title

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

Permalink

<https://escholarship.org/uc/item/1613j7vd>

Journal

Journal of applied physiology (Bethesda, Md. : 1985), 122(1)

ISSN

8750-7587

Author

Wagner, Peter D

Publication Date

2017

DOI

10.1152/japplphysiol.01046.2016

Peer reviewed

Editorial

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

Peter D. Wagner, MD
University of California, San Diego
La Jolla, CA 92093-0623
858-534-4190
jappleic@ucsd.edu

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

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

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

47 In the present article (1), Dr. Curran-Everett focuses on perhaps the most fundamental point in
48 statistical considerations: minimizing the uncertainty of one's conclusions. He was not asked to
49 invent anything new, but to show the way to reduce the probability of a conclusion being
50 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: a) differences in average data between groups, b) variance in those data, and c) 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 even though 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, since all else was identical between the two studies. Even if the line in the sand was 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 VO_{2max} ?: VO_{2peak} is no longer acceptable
- Non-invasive 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.

96

97 **REFERENCES**

- 98 1. Curran-Everett D. CORP: Minimizing the chances of false positives and false negatives.
99 *J Appl Physiol*, DOI: 10.1152/jappphysiol.00937-2016.