Small Data is Becoming a Bigger Challenge

Herbert Hoijtink
Methods and Statistics University Utrecht

Suggested talk duration (15-60 minutes)
I propose 10 minutes for my talk and 10 minutes for discussion (or more, see the end of the talk).

Summary (max. 500 words)
This talk is based on Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.--J., … H., Hoijtink, …, & Johnson, V. E. (2017). Redefine Statistical Significance. *Nature Human Behavior*. DOI: 10.1038/s41562-017-0189-z. In this paper the authors propose to change the default P-value threshold for statistical significance from 0.05 to 0.005 for claims of new discoveries. Three reasons for this are: 1) it reduces the probability of publication bias; 2) it reduces the options to use sloppy science to obtain significant test results; and 3) it corresponds to a Bayes factor of at least 10, that is, 10 times more (or less) support for the null versus the alternative hypothesis.
Of course using a smaller alpha-level has consequences for power. For example, to detect a medium effect size with a power of .80 in an independent samples t-test, 64 persons per group are sufficient if alpha equals .05. However, with alpha equal to .005, already 107 person per group are needed. Stated otherwise, small data is becoming a bigger challenge because traditional approaches to statistical inference will be hugely underpowered.
It is often thought that the use of informative hypotheses leads to a substantial increase in power. However comparing H0: both means are equal, versus. Ha: the first is larger than the second, still needs 50 persons per group (for alpha equals .05) and 94 persons per group (for alpha equals .005 and thus a corresponding Bayes factor of at least 10). Informative hypotheses not seem to be a convincing remedy for the power problem.
The talk will be concluded with two propositions that will be discussed with the audience:
1)   We should all change the default alpha level from .05 to .005
2)   Hypothesis evaluation by means of p-values or Bayes factors should not be used in case of small data

Relevance to conference theme
I propose not the use p-values or Bayes factors in case of small data

Keywords (max. 3)
Alpha level, p-value, Bayes factor