I linked to the preliminary draft: “Petitions in scientific argumentation: dissecting the request to retire statistical significance” (Hardwicke & Ioannidis 2019)

In response to signatures for Amrhein, Greenland et.al., 2019: fallacies of negative results lead them to recommend “retiring statistical significance”

In my editorial, I’d like to get in some of the recommendations in my blogpost
https://errorstatistics.com/2019/06/17/the-2019-asa-guide-to-p-values-and-statistical-significance-dont-say-what-you-dont-mean-some-recommendations/

The six principles of ASA I (2016):

I. P-values can indicate how incompatible the data are with a specified statistical model.
II. P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
III. Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.
IV. Proper inference requires full reporting and transparency.(P-values can’t be interpreted without knowing about multiple testing, data-dredging)

V. A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.
VI. By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis.

ASA II

♦ *Don't conclude anything about scientific or practical importance based on statistical significance (or lack thereof). (p. 1)*

♦ *No p-value can reveal the plausibility, presence, truth, or importance of an association or effect. (p.2)*

♦ *A declaration of statistical significance is the antithesis of thoughtfulness. (p. 4)*

♦ *Whether a p-value passes any arbitrary threshold should not be considered* ***at all*** *when deciding which results to present or highlight. (p. 2, my emphasis)*

♦ *It is time to stop using the term "statistically significant" entirely. Nor should variants such as "significantly different," "p < 0.05," and "nonsignificant" survive. (p.2)*

♦ *"Statistically significant"– don't say it and don't use it. (p. 2)*

The relationship between ASA I and II?

Principle 1 from ASA I. P-values can indicate how incompatible the data are with a specified statistical model.

The smaller the p-value, the greater the statistical incompatibility of the data with the null hypothesis, if the underlying assumptions used to calculate the p-value hold. This incompatibility can be interpreted as casting doubt on or providing evidence against the null hypothesis ..." (ASA I p. 131)

An indication of how incompatible data are with the absence of a relationship *would* be an indication of the presence of the relationship;

Principle 1 (from ASA I) doesn't appear to square with the first bulleted item I listed (from ASA II):

(1) "Don't conclude anything about scientific or practical importance based on statistical significance (or lack thereof)" (p.1, ASA II).

Either modify (1) or erase Principle 1.

But if you erase all thresholds for incompatibility there are no tests, and no falsifications, even of the statistical kind.

My understanding (from Ron Wasserstein) is that this bullet is intended to correspond to Principle 5 in ASA I – that P-values do not give population effect sizes.

But it is now saying something stronger

So, my first recommendation is:

Replace (1) with:
"Don't conclude anything about **the** scientific or practical importance **of the (population) effect size** based **only** on statistical significance (or lack thereof)."

Either that, or simply stick to Principle 5 from ASA I :
“A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.” (p. 132)

This statement is, strictly speaking, a tautology, true by the definitions of terms: probability isn’t itself a measure of the size of a (population) effect.

My second friendly amendment concerns the second bulleted item:

(2) No *p*-value can reveal the plausibility, presence, truth, or importance of an association or effect. (p. 2)

Focus just on "presence": it would seem to follow that no P-values[5], however small, even from well-controlled trials, can reveal the *presence* of an association or effect–and that is too strong.

Again, we get a conflict with Principle 1 from ASA I.

My second recommendation is to replace (2) with:

"No p-value **by itself** can reveal the plausibility, presence, truth, or importance of an association or effect.

Without this or a similar modification, the ability of any other statistical quantity or evidential measure is likewise unable to reveal these things.

This leads to my third bulleted item from ASA II:

(3) A declaration of statistical significance is the antithesis of thoughtfulness… it ignores what previous studies have contributed to our knowledge. (p. 4)

Surely they do not mean to say that anyone who asserts the observed difference is statistically significant at level p invariably ignores all previous studies, background information

Good inquiry is piecemeal: There is no reason to suppose one does everything at once in inquiry

My third recommendation is to replace (3) with (something like):

"**failing to report anything beyond** a declaration of statistical significance is the antithesis of thoughtfulness."