When to Adjust Alpha During Multiple Testing

Mark Rubin[1]

1 University of Newcastle

Abstract

In this new paper (Rubin, 2021), I consider when researchers should adjust their alpha level (significance threshold) during multiple testing and multiple comparisons.

In this new paper (Rubin, 2021), I consider when researchers should adjust their alpha level (significance threshold) during multiple testing and multiple comparisons. I consider three types of multiple testing (disjunction, conjunction, and individual), and I argue that an alpha adjustment is only required for one of these three types.

There's No Need to Adjust Alpha During Individual Testing

I argue that an alpha adjustment is *not* necessary when researchers undertake a single test of an individual null hypothesis, even when many such tests are conducted within the same study.

For example, in the jelly beans study below, it's perfectly acceptable to claim that there's "a link between green jelly beans and acne" using an unadjusted alpha level of .05 given that this claim is based on a *single test* of the hypothesis that green jelly beans cause acne rather than *multiple tests* of this hypothesis.

Retrieved from https://xkcd.com/882/
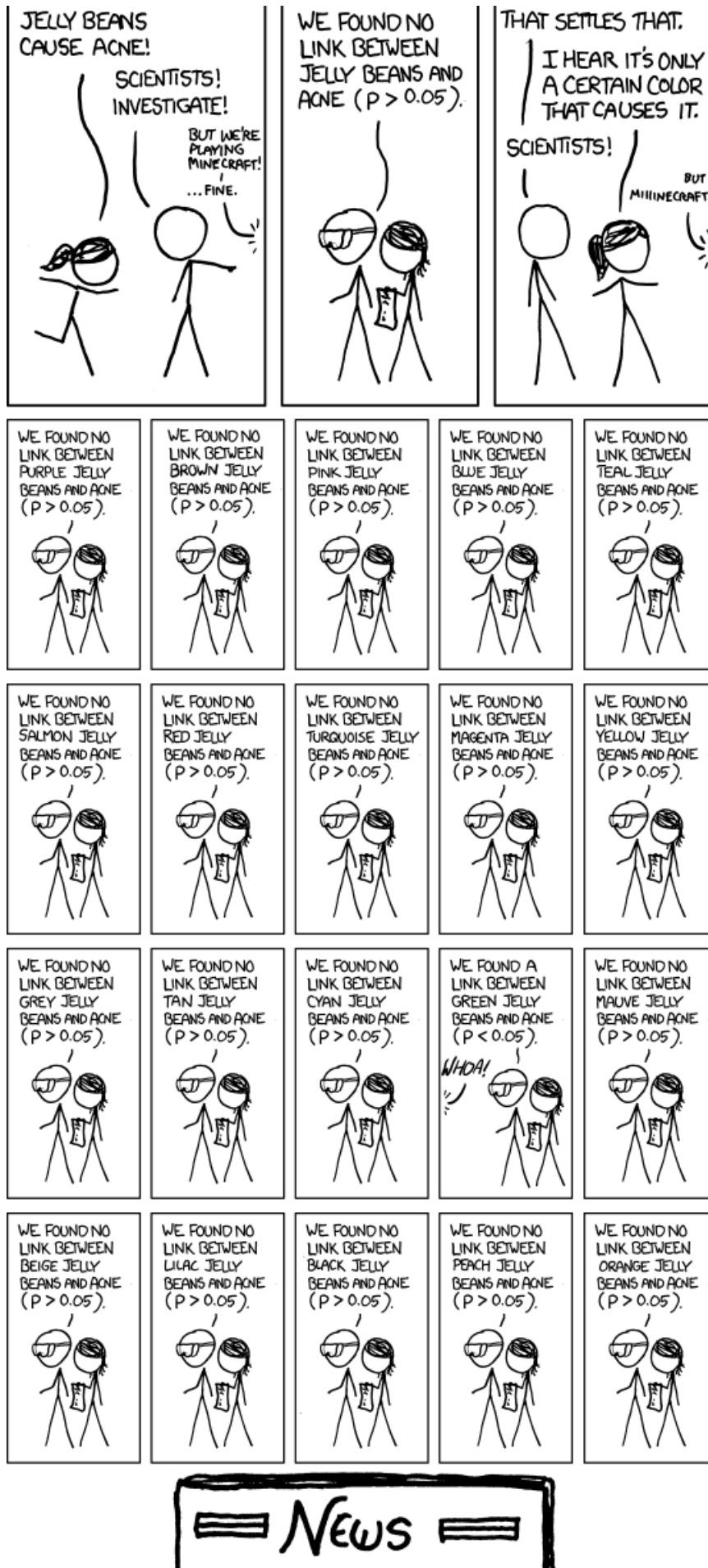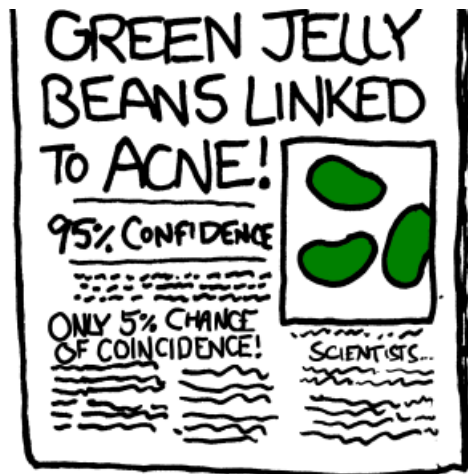
For a list of quotes from others that are consistent with my position on individual testing, please see Appendix B here.

To be clear, I'm not saying that an alpha adjustment is *never* necessary. It is necessary when *at least one* significant result would be sufficient to support a *joint* hypothesis that's composed of several constituent hypotheses that each undergo testing (i.e., disjunction testing). For example, an alpha adjustment would be necessary to conclude that "jelly beans of one or more colours cause acne" because, in this case, a single significant result for at least one of the 20 colours of jelly beans would be sufficient to support this claim, and so a familywise error rate is relevant.

Studywise Error Rates are Not Usually Relevant

I also argue against the automatic (mindless) use of what I call *studywise error rates* — the familywise error rate that is associated with all of the hypotheses that are tested in a study. I argue that researchers should only be interested in studywise error rates if they are interested in testing the associated joint studywise hypotheses, and researchers are not usually interested in testing studywise hypotheses because they rarely have any theoretical relevance. As I explain in my paper, "in many cases, the joint studywise hypothesis has no relevance to researchers' specific research questions, because its constituent hypotheses refer to comparisons and variables that have no theoretical or practical basis for joint consideration."

Sometimes it doesn't make sense to combine different hypotheses as part of the same family!

For example, imagine that a researcher conducts a study in which they test gender, age, and nationality differences in alcohol use. Do they need to adjust their alpha level to account for their multiple testing? I argue "no" unless they want to test a studywise hypothesis that, for example: "*Either* (a) men drink more than women, (b) young people drink more than older people, *or* (c) the English drink more than Italians." If the researcher does not want to test this potentially atheoretical joint hypothesis, then they should not be interested in controlling the associated familywise error rate, and instead they should consider each individual hypothesis separately. As I explain in my paper, "researchers should not be concerned about erroneous answers to questions that they are not asking."
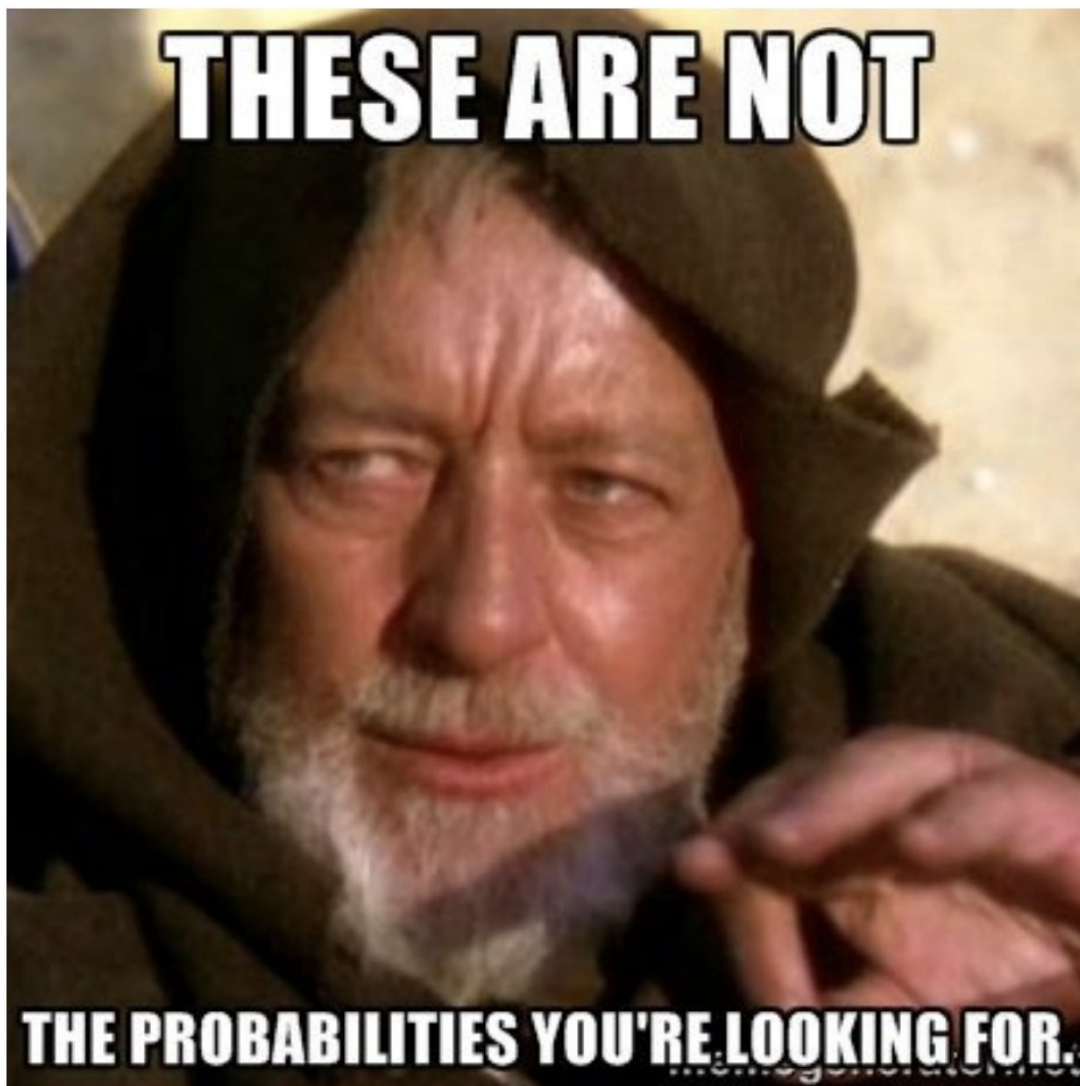
For a list of quotes that support my position on studywise error rates, please see Appendix A here.

For further information, please see:

Rubin, M. (2021). When to adjust alpha during multiple testing: A consideration of disjunction, conjunction, and individual testing. *Synthese*. https://doi.org/10.1007/s11229-021-03276-4 or Open Access

For more of my work in this area, please see:

https://sites.google.com/site/markrubinsocialpsychresearch/replication-crisis