Reviewer's report

Title: The Null Hypothesis Significance Test in Health Sciences Research (1995-2006): Statistical Analysis and Interpretation

Version: 1 Date: 8 February 2010

Reviewer: Lorena Rios

Reviewer's report:

Major Compulsory Revisions

1. At the end of the introduction section there are several expressions that can result ambiguous to the reader. Please define explicitly what do you mean by the following expressions:
   a) “it is interesting to examine patterns of use of the NHST”… What do you mean by “patterns of use”? 
   b) “To identify what is missing in the process of eradicating….” Do you have any hypotheses for this? Or any suggestions on how to solve the problem? 
   c) “In addition to assessing the adherence to the guidelines proposed by JCMJE”. What elements of the guidelines? 
   d) “To examine whether these patterns differ between English and Spanish speaking worlds”. Why do you want to examine this? Is there any reason for thinking that there is a difference? 
   e) What is meant by substantive significance? Do you mean clinical importance? 
   f) “The author is letting a computer do the thinking…” In what way?

2. The paper does not state the objectives clearly. Please, provide a clear definition of the objectives or research question at the end of the introduction section.

3. In the first paragraph of the Material and Methods section, the manuscript describes how journals were selected. Please justify the journal selection, particularly:
   a) How did you determine the degree of influence of the journals? 
   b) Have any of the selected journals formally endorsed the ICMJE guidelines?

4. Why did you decide to take 60 papers for each biennium and journal?

5. How did you classify papers as descriptive, analytical or theoretical? What criteria did you use?

6. It is not clear from the methods section if descriptive and theoretical papers were excluded from the analysis. Please, provide specific inclusion and exclusion criteria. It would be very helpful to add a table or a flow chart like the one
recommended by the CONSORT statement showing the selection/sampling process of all articles, number excluded with the reasons for exclusion and number of papers analyzed.

7. Studies usually perform multiple tests on multiple outcomes with multiple p values and CI. Probably, not all the p values will be significant and therefore not all will be misleadingly reported. It is unclear what type of data on all these issues were abstracted for the purpose of this study.

8. The operational definition used to label a paper as having the problem of significance fallacy is not clearly stated in the methods section. Please, explain how did you operationalize its assessment.

9. Who did the data abstraction? What were their qualifications? Was it done in duplicate? How was the reliability of the data abstraction process assessed?

10. At the end of the methods section, you mention a pilot study of your data abstraction. It would be very informative if you provide information on the number of papers evaluated in the pilot study. Please, provide also the kappa values for each of the items you evaluated in duplicate. In the description, it is not clear what the kappa value of 0.78 corresponds to. Please, also provide CI for the kappa values.

11. Combining Bayesian CI and classical CI seems inappropriate – given the goal of the study - which is to show how studies inappropriately perpetuate the use of NHST. This issue is not applicable to the Bayesian paradigm. This issue needs further clarification.

12. In the Results section, 2nd paragraph it is stated that most clinical articles ignore the ICMJE recommendations. Assessing the adherence to ICNJE recommendations is not explicitly stated as an objective of the study. There is no description on how this was done in the methods section either. Therefore, it is hard to understand the criteria used to make this judgment. Please, clarify.

13. In the Results section, 3rd paragraph it is stated that the percentage of articles referring to significance in an ambiguous or incorrect way is declining over time. What do you mean by incorrect or ambiguous? What is correct? The ambiguity of these expressions is probably related to the lack of an operational definition for significance fallacy. Please, clarify.

14. In the discussion section, 1st paragraph: Why do you think it is puzzling? Please, elaborate further on this.

15. In the discussion section, at the end of the 1st paragraph: the expression “ritualistically and mindlessly as the dominant doctrine” seems unjustified.

16. In the discussion section, 3rd paragraph it is stated: “the trend analysis shows that.....” The paper does not present a trend analysis, it rather presents a descriptive analysis by periods.

17. Perhaps, the study could address the issue of determining factors that can
explain the observed results. This would help the scientific community to think about possible solutions.

18. In the discussion or the introduction sections, there is no mention of any other studies that have explored the pattern of use of p values and CI in the medical literature. If there is any, you should cite that work and also compare your results with previous studies. If you did not find any study similar to your work, you could state it in your manuscript.

19. In the discussion of the limitations of your work, you should mention the subjectivity in the evaluation of the presence of significance fallacy in the articles, the lack of an assessment in duplicate for all the papers (if this is the case), and how you dealt with these problems.

20. Please, provide a conclusion containing the main message of your manuscript at the end of the discussion section.

Minor Essential Revisions

1. Material and Methods: the 3rd paragraph begins “49 of the 1092 papers...”. When a sentence starts with a number, please write it in words (Forty nine instead of 49).

2. Material and Methods: 4rd paragraph, last sentence: it reads “indicators of error” and it should say: indicators of margin of error.

3. Please, correct table 5, where categories are named in Spanish instead of English.

4. In the discussion section, the 2nd paragraph reads: “not long ago, when researchers did observe...”. It seems that it should say: “not long ago, when researchers did NOT observe...”. Please, check this.

5. On table 3, the 4th column is labeled fallacy of signification. It should say significance fallacy.

6. On Table 2, columns are labeled IC instead of CI.

7. An all tables, please correct the display of CI using the standard on reporting (1): Example: % (95%CI) 14.1 (9.3 to 18.9)

8. Please, define acronyms below all tables.

9. In the tables, it is not clear if the denominators correspond only to analytical papers or the numbers also include theoretical and methodological papers. Please, clarify.

10. Regarding the English language, the paper needs thorough editing using shorter sentences and the active voice. Please, also correct some outstanding grammatical errors.

11. Regarding the references, please use the BMC Journal bibliography style.
and be consistent. For example, some journals have been abbreviated while others have been written in full. The citation of websites has sometimes the word “accessed” and sometimes not, the issue number is sometimes present and sometimes not, etc.

References:

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Not suitable for publication unless extensively edited

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:
I declare that I have no competing interests.