Reviewer's report

Title: Two distinct do-not-resuscitate protocols leaving less to the imagination: an observational study using propensity score matching

Version: 3

Date: 8 July 2014

Reviewer: Joseph Massaro

Reviewer's report:

Major Compulsory Revisions:
1. Re: authors response #4 to previous major compulsory revisions: For the calculation of standardized differences, see http://www.tandfonline.com/doi/abs/10.1080/03610910902859574#.U7vXIMtOMW8. Here, they discuss the standardized difference and the cutoff value of 10%. Also please see “Austin (2008) : Assessing balance in measured baseline covariates when using many-to-one matching on the propensity score. Pharmacoepidemiology and Drug Safety. 2008; 17: 1218-1225.” Though this latter manuscript’s primary objective is to discuss standardized differences for N:1 matching, it provides a background of assessing balance diagnostics using 1:1 matching, and provides further references for assessing balance at baseline in 1:1 matching.

I’m not exactly sure how STATA procedure “pbalchk” works, but it appears (and I may very well be wrong) that it basically is considering groups well matched if a z-statistic between the two groups is between -2 and 2; this basically is the same as presenting a p-value to assess balance/imbalance.

It is fine if the authors present any type of standardized difference they feel comfortable with, as long as they provide the brief details in the manuscript of how it is calculated. The bottom line is that some sort of a standardized difference approach to assess balance/imbalance is more useful and generally more accepted than a p-value approach might be.

Minor Essential Revisions:
1. Re: authors response #3 to previous major compulsory revisions: The approach the authors used to match on the propensity score is acceptable. However, one of the points of my original comment was to ask the authors to provide details of the propensity matching in the manuscript. Specifically, please mention if a “greedy” algorithm or “optimal” algorithm was used and please specifically mention that no caliper was used.

2. Re Authors’ Response #3 to the previous minor essential revisions: While the central limit theorem may allow you to analyze skewed data using a parametric approach (and one could argue that, for very highly skewed variables, a parametric approach may still not be appropriate even with a sample of size 88), the means can still give a biased estimate of central tendency. Please present
median and quartiles (in addition to mean and standard deviation) of potentially skewed variables such as ICU Stay, etc.

Discretionary Revisions:
1. Re: authors Response #5 to previous major compulsory revisions: I appreciate the sensitivity analyses the authors have added, and it is useful. But I apologize -- I was not clear in my previous comment regarding the sensitivity analysis. I was actually suggesting that the authors provide an analysis comparing, for example, the 88 DNRCC with the entire cohort of 2051 Non-DNR, adjusting for propensity score. Often, in non-randomized trials, when matching is performed as the primary analysis, an adjusted analysis on all participants is provided as a secondary analysis. However, I will consider this a discretionary revision.

The following are comments for which no further revisions required:
1. Re: authors’ response #1 to previous major compulsory revisions: I agree with the author’s response and manuscript revision.
2. Re: authors’ response #2 to previous major compulsory revisions: I am fine with the author’s response. There are several trains of thoughts of whether conditional or unconditional is most appropriate, so if the authors feel an unconditional approach is more acceptable, then that is what they should use.
3. Re: authors’ response #6 to previous major compulsory revisions: I am fine with the authors response.
4. Re: authors’ response #1 to previous essential minor revisions: I am fine with the authors’ response.
5. Re: authors’ response #2 to previous essential minor revisions: I am fine with the authors’ response.

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

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