Reviewer's report

Title: Opinion versus practice regarding the use of rehabilitation services in home care: An investigation using machine learning algorithms

Version: 2 Date: 28 April 2015

Reviewer: Lucia Sacchi

Reviewer's report:

General Comment.
This paper discusses how machine learning algorithms can be used to perform a preliminary assessment for patients that might need rehabilitation. The paper relies on standardized assessment systems. The topic presented in the paper is interesting and refining expert classification through machine learning techniques can help in supporting clinical decision for planning and organizational purposes.

However, the paper has several weaknesses, which prevent it to be published in its current form. First of all, the paper is not easy to follow. There are many aspects that need to be clarified, both related to the presentation of the used methodologies and to the application to the real data. I will detail these in the following.

Detailed comments
1. At line 149 InfoRehab is mentioned for the first time, but the authors do not explain what it is. Does it refer to the "multidisciplinary health research program" introduced in line 147?
2. How is the decision model for the rehabilitation algorithm (Figure 1) built? Is it based on expert knowledge? Does it rely on already collected data? Was it built using a specific algorithm?
3. Figure 1 is not clear. What do the abbreviations within the boxes (e.g. D5 = 2 or 8) mean? The Figure must be better explained in the text. The Figure caption contains a bit more detail but it is confusing, as it introduces some parameters that are then not shown in the Figure itself (e.g. C2a, C2b, etc)
4. Were the 239 items considered as potential predictors selected manually based on expert knowledge? Was any automatic features selection technique applied?
5. In line 212, the authors state that 239 items were selected from RAI-HC, but their target is RAI-CA. How items selected to be predictors for rehabilitation are treated if they result not available in the subset of items considered by CA?
6. In line 231, the "CA database" is introduced, and it is then used to stratify the HC population accordingly. Unfortunately, the characteristics of the CA dataset had not been introduced up to that point in the paper, and for the reader it is difficult to understand how the population was then stratified.
7. It is not clear why the HC population was taken as the control population. I
would imagine control samples to be patients who are not in need for rehabilitation, but this is probably not the case. Please better address this point.

8. Table 1 is not very informative without an initial summary of the population. What is the percentage of patient falling in each of the categories (e.g. % of males, % of ADL=1, % of patients in each age class, etc). Percentages in Table 1 look very low, and there are too many combinations for the reader to extract some information from it.

9. In Figure 2, the authors show six predictors. What are these? Which variables do they represent? Are they related to the considered problem or are they just an example? How were they selected?

10. The part related to how potential interactions are taken into account using random forests is not clear and must be better explained in the text. Does the variable importance measure account for interactions? How is this expressed? Is it a score related to a single variable or is there a score for each interaction?

11. Line 338 starts talking about Table 2 (The fourth column (effect on odds ratio)), which is though not mentioned in the previous part of the text about Figure 4. Then the text keeps referring to Figure 4, creating confusion in the reader.

12. Was the predictive performance of the model evaluated in some way? This is necessary to understand if the tool can be used for prediction and, most importantly, if the presented model could be used in practice.

13. This paper lacks some discussion of the possible use of the proposed models in practice. Would it be better selecting LASSO regression or random forests? What is the usefulness of the extracted results from a medical perspective? Can they help to better target rehabilitation resources? What are the parameters that can be used to refine expert opinion?

Abbreviations need to be explained in the text (e.g. ADL and IADL, line 131)

Line 238: followed --> follows

Line 317: this sentence doesn't contain a verb

**Level of interest:** An article of importance in its field

**Quality of written English:** Needs some language corrections before being published

**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