Reviewer's report

Title: The influence of measurement error on calibration, discrimination, and overall estimation of a risk prediction model

Version: 1 Date: 14 March 2012

Reviewer: Polyna Khudyakov

Reviewer's report:

While I am generally supportive of this manuscript, I do have some remarks that I think the authors would need to address. So, I proposed to do an essential revision addressing my comments below.

Major Compulsory Revisions

1. I found that section “Methods” was unclear and hard to follow. I think that this section must be significantly revised.

2. Page 6, line 23: “height (m) and weight (kg) were generated as Gaussian random variables with means and variances obtained from the derivation cohort.” As I understand the values of the parameters for this Gaussian distribution are presented in Table 1, but this is mentioned only on page 11 at the beginning of the Results section.

3. Page 7, line 1: it was mentioned that “… specified values for the regression coefficients” were used, but the reader can only guess that these values are actually presented in Table 2.

4. Page 7, line 5: “BMI is calculated as the ratio of height in meters (m) to weight in kilograms (kg)”, but as far as I know, the BMI is defined as the body weight divided by the square of height. The definition similar to mine appears on page 10, but in Table 1 we see again “(m/kg^2)”.

5. Page 9, line 13: “in this simulation wi will represent various levels of over- and underestimation”. It would be more helpful for readers to understand the contents if the authors provide the values of wi.

6. Page 9, last sentence: “Simulated values of height and weight were generated by applying a random component to the standard deviation of the variable”. It will be helpful, if you be more specific here. I think you have to explain the random component that was applied here and in which way it was done. Moreover, the formulas for Height and Weight, written on page 10, do not have sense.

7. Overall, the formulas written in the paper have different format and notations, which makes the reading hard to follow. In particular, formulas on page 9 are presented using Y and wi, but on page 10 the formulas are written using Height, Weight, M1 and M2. I also do not understand why the definition of square of BMIobserved was presented here. In addition, I think it is a mistake to add index to n in the first formula on page 10.

8. Page 10, line 8: “*note that …” I do not understand why this sentence starts
with * and ends with comma and point. Moreover, I don’t understand how this sentence is connected to the text that was written before and also the text that following after this sentence.

Minor Essential Revisions
1. Page 4 line 9: the sentence needs a grammatical correction.
2. Page 6, line 22: it is written “To randomly generate a person with diabetes, …”. By this sentence the reader could understand that you are targeting your population to the diabetes patients only. I suggest you to revise this sentence and to add information about the percent of people with diabetes in your population.
3. Page 11: I suggest you to revise the titles of subsections in the section “Results”, because in the current form the Results section looks too short and the subsection “Random Error” can be considered as a new section.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

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.