Author's response to reviews

Title: The accuracy of General Practitioner workforce projections

Authors:

Malou Van Greuningen (m.vangreuningen@nivel.nl)
Ronald S Batenburg (r.batenburg@nivel.nl)
Lud FJ Van der Velden (l.vandervelden@nivel.nl)

Version: 3 Date: 19 June 2013

Author's response to reviews: see over
Dear Editors,

I am enclosing herewith the most recent revised version of our manuscript (MS: 1195208522923906) ‘The accuracy of General Practitioner workforce projections’ for publication in Human Resources for Health.

We are pleased that the modifications made in the first revision of the manuscript fulfilled the reviewers’ suggestions. We are thankful to the statistical expert for pointing out some important and appreciated suggestions to further improve the article. We have thoughtfully taken into account these latest comments and made some modifications. The explanation of what we have changed in response to the concerns of the reviewer is listed, point by point, in the following pages.

We believe that the comments have been very useful to improve the manuscript. We hope that all changes included in the revised paper fulfil the reviewers suggestions and makes the manuscript acceptable for publication in Human Resources for Health soon.

Looking forward to hearing from you soon,

Kind regards,

Malou van Greuningen

Corresponding author is:
Malou van Greuningen
NIVEL
P.O. Box 1568
3500 BN Utrecht
The Netherlands
Phone: +31 30 2729 795
Email: m.vangreuningen@nivel.nl
Authors’ reply to reviewers’ reports

Additionally, as requested by some of the reviewers, the editors have asked an expert to review the statistical methodologies.

Overall, your statistical methodologies need to be described and justified more, particularly since the methods used will probably be unfamiliar to an average reader. Second, you are confirming and rejecting three hypotheses, but you do not state the statistical criteria for doing so. Therefore, here is some more specific comments that the editors would like to be addressed by the authors:

? When describing the Dutch model on p. 8, you should give an example prediction with numbers. The equations on that page do not have all the elements included in Figure 1's model. Need to explain.
We discuss the equation more thoroughly on p. 8. The inflow-part of the model contains several other sub elements, such as return on training, inflow from abroad, labour market return. We have also illustrated the equation with an example prediction with numbers on p. 8 and 9.

? You state that the Dutch GP workforce projection model is back tested, but back tested is not adequately described. You reference two articles by Dowd, which are probably unfamiliar to your readers.
In the methods section we described what backtesting is, in more detail (p. 7).

? You did not sufficiently motivate why "back testing" is the best method to test the accuracy of the Dutch GP forecasting model.
In the methods section we added a motivation for choosing back testing as a method to test the accuracy (p. 7). This is also discussed in the discussion section (p. 14).

? In the prediction literature, there is a bias-variance trade-off e.g., your predictive model might result in a low bias (but very high variance) or a high bias (and very low variance). This was not discussed. For example, see Hastie et al., The Elements of Statistical Learning, and Springer.
The projection errors are mainly caused by bias and not by variance. The variance is low, because data of all Dutch GPs is used to make projections. The projection error is mostly bias, caused by differences between the past GP workforce and the current and future GP workforce. This is also discussed on p. 15.

? The first hypothesis, the longer the projection period, the lower the accuracy, is confirmed. It is unclear what their confirmation/rejection criteria were from a statistical standpoint. It seems as though it was based on visual inspection of Table 1.
It was indeed based on visual inspection of Table 1. As a consequence of the experts’ remarks we have reconsidered this. We have decided to test the differences in accuracy statistically, according to the three hypotheses stated in the article. In the methods section (p. 10/11) and results section (p. 12/13) we discuss these tests.

? Your confirmation is well known in forecasting, so you should discuss how the magnitude of the inaccuracy increases with the projection period.
We discuss this on p. 12 in the results section.

? Table 1 is confusing to follow. Needs more description.
We have added a descriptive legend to the table, in which the contents is explained.
The second hypothesis, the longer the base period, the higher the accuracy, is not confirmed. Again, like hypothesis 1, what were the statistical criteria? See the answer to the question about hypothesis 1.

Since it was not confirmed, it seems counter intuitive, because you are saying more data reduces accuracy. You state that this is because older data are less representative for GPs in the target year. It is not clear what you mean by that e.g., is it how long they stay in workforce, hours per week, etc.?

According to projections based on base periods including older data, we expect the GPs to flow out of the workforce at an earlier age than is observed. If the base period is longer and thus contains also older data, the outflow is overestimated and the total number of GPs is thus underestimated. It is in this light that older data is less representative for GPs in the target year, because current GPs stay in the workforce longer. This is also the case for labour market return (is higher for current GPs). This is also discussed on p. 14.

For third hypothesis, again, same comments as above. See the answer to the question about hypothesis 1.