Manuscript ID BMJ.2015.027121 entitled "Development and validation of risk prediction equations to estimate future risk of blindness and lower limb amputation in patients with diabetes: cohort study."

Thank you for provisionally accepting our paper. We have revised the paper as suggested, highlighting the revisions in the text in blue. We have also provided a clean copy and the PICO.

****Report from The BMJ's manuscript committee meeting****

Detailed comments from the meeting:
First and foremost, please revise your paper to respond to all of the comments by the reviewers. Their reports are available at the end of this letter, below.

Please also respond to these additional comments by the committee:

*Editors pointed out that these conditions are not seen very often in the UK anymore, but are much more common in other parts of the world (including the US) and thus this does have wider clinical relevance. (The units for HBA1c are also different in the US, which should be considered in the reporting of these data.) These are also very serious complications and thus are of interest to clinicians.
Authors' response: We have added a reference to make it clear that the UK now uses the Standard International (SI) unit of millimoles of HbA1c per mole of Hb (mmol/mol) instead of the percentage[1]. We have also added a reference to the conversion chart between percentage and mmol/mol.[2]

*Editors wondered about the clinical usefulness of the risk prediction, however, since there's not much that can be done to prevent these complications.
Authors' response: We have added a reference to section 4.2 with evidence that blindness can be preventable by screening and treatment of retinopathy[3]. There is also some evidence for targeted programs to prevent amputation as highlighted in section 4.2 of the discussion on clinical implications.

*We would like to see a hyperlink to the risk calculator in the paper itself.
Authors' response: There was already a link on page 3 and in section 3.5.3. We have added another one to section 3.4.

Reviewer: 1 Eva Pagano

Comments:
The Authors have developed and externally validated risk prediction equations to quantify the absolute risk for blindness and lower limb amputation in patients with diabetes, at 10 years.

The methodological approach adopted is well designed and applied. Equations, with a double external validation, show robust results. The methods description is clear and easy to follow.
Reasons to conduct the study and provide patients and doctors of a prediction tool are well discussed. Data sources are reliable and currently used in observational studies.
Results are of practical interest for clinicians. The identified predictive tool could be used for better stratifying patients according to their individual risks.
Authors' response: Thank you for these comments.

As general comment, I would stress in discussion the problems of generalisability to other countries with different epidemiological patterns.
Authors' response: we have added a sentence to emphasize this in section 4.4 of the discussion.
I have got only minor comments.

Abstract

- Measurement: Mortality is listed beside hospital and GP electronic records as a source to measure incident diagnoses of blindness and amputation. This choice seems weird. How mortality record can contribute in identifying amputations and, specially, blindness? As clearly stated in Methods, ONS provides data on survival, for censoring observation. If ONS is also used for identifying diabetes cases and complications, it should be clearly stated. The role of ONS for identifying the study outcomes, if any, should be clarify. (Same comment for the paragraph 2.2 in Methods).

Authors' response: We have clarified in section 2.2 that we used ICD-10 codes to identify cases from either the primary or underlying cause of death as recorded on the linked ONS mortality record. For example, a patient may have had a fall as the primary cause of death but the underlying cause may be recorded as blindness. We have also highlighted the use of the ONS mortality record to identify incident cases in the abstract.

- Methods (pag. 3): among the listed measures I would substitute sensitivity with accuracy, as we are interested in both sensitivity and specificity.

Authors' response: - we have amended the abstract to add specificity in addition to sensitivity. In section 2.5 we mention specificity and sensitivity and the results are presented in table 5. We think it's clearer to use the precise statistical terminology such as sensitivity and specificity rather than terms such as accuracy.

Introduction

Beside the cited UKPDS model, other models provide estimate of risk of microvascular complications. For example, the CORE Diabetes Model provides simulation of amputation and several models (CORE diabetes model, EAGLE model and Sheffield Diabetes model) provide risk estimates for retinopathy rather than blindness. Such literature should be mentioned as relevant background. Accordingly, in Discussion (4.1 Key findings), the estimated equations cannot be considered the first in predicting blindness and amputation. Such statement should be rephrased to better position the study contribution within the existing evidence. Mostly, to provide a tool to easily calculate the mentioned risks.

Authors' response: We have included the following text in the discussion. "There are three economic models are based on the DCCT and UKPDS studies. The CORE diabetes model[4 5]

and the Sheffield diabetes model[6] are both based on equations derived from the DCCT trial and the UKPDS study. The EAGLE model[7] is based on equations derived from UKPDS and the DCCT as well as the Wisconsin Epidemiological Study of Diabetic Retinopathy. The CORE model predicts risk of amputation[7] whilst the CORE, EAGLE and Sheffield models predict retinopathy rather than blindness".

Methods

For international readers, the Egton Medical Information System and the Read codes are not straightforward. Please, provide a concise explanation.
Authors' response: We have added a sentence to explain the EMIS system in section 2.1. We have highlighted the sentence about Read codes which includes a reference to a Wikipedia page explaining more about Read codes.

The acronym OPCS is not fully explained at its first use.
Authors' response: we have explained OPCS-4 now in section 2.2 and included a reference with a hyperlink with more details about the coding system.

In the predictor variable identification, is there any cut-off time point for variables based on the latest information recorded in the primary care record before entry to the cohort?
Authors' response: We didn't use a cut off – it was based on the latest information recorded in the record as described in section 2.3

Methods to identify the clinical events used as predictors (atrial fibrillation, congestive cardiac failure,….) have not been clarified. I would not recommend a detailed list of the criteria/codes, but a generic reference to the data sources used and the time frame searched. Number of years since diabetes diagnosis is also a variable of interest and details on the source should be added (see comments on results).
Authors' response: We have highlighted the text with this information in section 2.3.

Derivation of the models:

- Typo in 2.4 (second line): "for" is repeated. Authors' response: We have corrected this.
- Reasons to carry out 10 imputation should be clarified as for readers, especially clinicians, it is not meaningful. Authors' response: We have clarified the reason in the text (to improve statistical efficiency) and have added a reference[8].
- The same for the criteria used for retaining variables (<0.80 and >1.20). Authors' response: One is a 20% reduction and the other is a 20% increase so they are equivalent in that respect. With a large dataset smaller hazard ratios might not be clinically relevant and might result in too complicated a model.
- What is intended with "plausible" in examining interactions? Authors' response: interactions which are similar in direction for both men and women and consistent with the literature.

Results

The high percentage of newly diagnosed patients is not easily justifiable. A comment should be provided. Still, from the table is not clear if there are missing data on this variable.
Authors' response: We had an open cohort so people who developed diabetes during the

follow up period were eligible for entry to the cohort from that point, meaning the percentage of newly diagnosed patients is higher than if we had assembled a cohort at a single timepoint. This was done to maximize the study population and also provide information on the effect of duration of diabetes. All patients had a diagnosis date so there is no missing data for this variable.

When listing the amount of missing data for the different variable, cholesterol ratio should be added.
Authors': This information was already in Table 1. We have now added it to section 3.2 of the results.

Is there any explained reason for the difference in recording ethnicity and HBA1c among QResearch and CPRD? Provide a comment.
Authors: There are differences in the computer systems used which may affect data recording- QResearch is derived from the EMIS system and CPRD is derived from the In Practice System (INPS).

The paragraph 3.3 Primary outcome of amputation and blindness: Description in the text of absolute number of events is not of great interest. I suggest to describe the relative measures and to highlight the relevant differences, where present (e.g. lower blindness among men in CPRD versus QResearch). Authors: We have added a sentence to section 3.3. "The rate of blindness in men was lower in CPRD (2.33 per 1000 person years) than in both QResearch cohorts (3.03 per 1000 person years)".

3.4.1. Lower limb amputation: The effect of smoking should be described. Authors' response: We have added the following text to section 4.1. "Increasing levels of smoking were associated with increased risk of amputation with the association being more marked amongst women than men. For heavy smokers compared with non-smokers, there was a 1.9 fold increase in risk of amputation for women and a 1.3 fold increased risk for men"

3.5.1Validation results for QResearch are not so similar to those of the CPRD cohort, as stated. Even if a formal statistical difference is present only for amputation (ROC statistic both in women and men), all the point estimates of the indicators are lower in QResearch. Authors should comment this result even if it does not change the overall discussion. Authors – We have highlighted this in section 3.5.1.

The last paragraph of Methods (beginning of pag.15) should be detached from the previous by adding a specific title (e.g. 3.6 Implementation). Authors' response: We agree and have added a new heading 3.6 for implementation.

Discussion
Overall, I recommend a more exhaustive reference to the literature in the field of predicting models. The last sentence of the 4.1 Key findings could be changed into "the fist tool for predicting…". Authors – we have made this change.

4.3 Comparisons with the literature, last sentence "However, as previously reported,…" Data seem not to support the statement as smoking was not included in the final models. Rather, it was included in the amputation models. Could it be a typing error?
Authors' response: smoking was included in the amputation model but not the blindness model. We have adjusted the wording to make this clearer.

Figure 4 and 5. Are not easy to understand. Please add the y-axis label. Author: We have added a label to the y axes for both graphs.

As the outcomes identification was mainly based on Read codes (used by UK general practice), I am not familiar with them and I was not able to judge the selection provided by the Authors.

Reviewer: 2 Gary Collins

Comments:
The authors are treading a well-worn path having developed numerous risk scores from the QRESEARCH database. Risk scores for blindness and lower limb amputation have been developed, internally validated and externally validated on very large datasets.

The methodology is strong and the authors adhering to recommended practices in developing and validating risk scores. The authors have also carefully followed the TRIPOD reporting guidelines for prediction models.

My comments are minor.

Abstract. Make it clear that the patients come from the UK.

Authors' response: We have added UK to the abstract.

Page 9/Page 10: Multiple imputation. Some more information would be useful, whether in the main paper or in supplementary material, that includes what variables were included, any transformations etc.
Authors' response: We have added this detail to section 2.4.

Page 10. Confirm that the Area Under the Receiver Operating Characteristic Curve has been calculated for survival data (and not for binary outcomes).
Authors' response: Originally we had calculated the ROC value based on binary outcomes but have re-calculated it using the Harrell's C statistic for survival data and combined values across the imputed datasets. We have clarified this in section 2.5 of the methods.

Page 10. Why are thresholds of 10% and 20% chosen? Is this for illustrative purposes only? If these are thresholds recommended by the authors then some rationale would be needed. Alternatively, the authors could produce net benefit curves (decision curve analysis, Vickers et al 2006; Med Decis Making) to examine this.

Authors' response: These were chosen for illustrative purposes only. We have added a sentence to clarify this in the results (section 3.5.3) and discussion (section 4.4)

Page 11. Whilst missing data is mentioned for particular variables, how many of the cohort had complete data (or conversely how many had at least 1 missing variable) in the development cohort.
Authors' response: Of the 454,575 patients in the derivation cohort, 266,142 (58.6%) had missing data for at least one of the predictor variables. We have added this to section 3.2.

Page 12. Whilst missing data is mentioned for particular variables, how many of the cohort had complete data (or conversely how many had at least 1 missing variable) in the validation cohort.
Authors' response: Of the 142,419 patients in the QResearch validation cohort, 83,403 (58.6%) had missing data for at least one of the predictor variables. Of the 206,050 patients in the CPRD validation cohort, 166,648 (80.9%) had missing data for at least one of the predictor variables. We have added this to page section 3.2.

Page 12. Whilst the authors mention the availability of the risk score at the start of the manuscript. It would also seem a natural place to discuss this in the Results section, preferably with a brief explanation as to why the model is not published in the paper. I know the reasons as to why this is the case, but the average reader won't. The authors are one of the few teams I am aware that actively maintain and update their models on an annual basis, and this should be highlighted as it is a particular strength of these models/group.

Authors' response: We have added these points to section 3.4 of results and section 4.4 of the discussion as suggested.

References

1. Barth JH, Marshall SM, Watson ID. Consensus meeting on reporting glycated haemoglobin (HbA1c) and estimated average glucose (eAG) in the UK: report to the National Director for Diabetes, Department of Health. Diabet Med 2008;**25**(4):381-2.
2. Drugs and Therapuetics Bulletin. Change in units for HbA1c Secondary Change in units for HbA1c 2010. http://dtb.bmj.com/site/about/HBA1C_chart_Feb_10.pdf.
3. Bachmann MO, Nelson SJ. Impact of diabetic retinopathy screening on a British district population: case detection and blindness prevention in an evidence- based model. Journal of Epidemiology and Community Health 1998;**52**(1):45-52.
4. Palmer AJ, Roze S, Valentine WJ, et al. The CORE Diabetes Model: Projecting Long-term Clinical Outcomes, Costs and Costeffectiveness of Interventions in Diabetes Mellitus (Types 1 and 2) to Support Clinical and Reimbursement Decision-making. Current Medical Research and Opinion 2004;**20**(sup1):S5-S26.
5. Foos V, Palmer JL, Grant D, et al. PRM58 Long-Term Validation of the IMS CORE Diabetes Model in Type 1 and Type 2 Diabetes. Value in Health;**15**(7):A470.
6. Waugh N, Scotland G, McNamee P, et al. Screening for type 2 diabetes: literature review and economic modelling. Health Technol Assessment. London: HTA, 2007:125.

7. Mueller E, Maxion-Bergemann S, Gultyaev D, et al. Development and validation of the Economic Assessment of Glycemic Control and Long-Term Effects of diabetes (EAGLE) model. Diabetes Technol Ther 2006;**8**(2):219-36.
8. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011;**30**(4):377-99.