**Reviewer’s report**

**Title:** Economic evaluation of Type 2 Diabetes Prevention Programmes: Markov model of low- and high-intensity lifestyle programmes and metformin in participants with different categories of intermediate hyperglycaemia

**Version:** 0  **Date:** 03 Aug 2017

**Reviewer:** Tony Blakely

**Reviewer's report:**

ECONOMIC EVALUATION OF TYPE 2 DIABETES PREVENTION PROGRAMMES: DOES THE TYPE OF PRE-DIABETES AND INTENSITY OF INTERVENTION MATTER?

This paper is very good, and well-positioned in terms of current questions and uncertainties. The conclusion that these lifestyle interventions are (moderately) effective and cost effective, but are no panacea for reducing diabetes incidence rates, is important and needs making - and the authors explicitly state this in the Conclusion emphasizing the need for population-wide prevention. Point well made. The methods and approach, be it from working with advisory groups through to actually modelling, seems strong - although I have a number of suggestions for improvement.

Major comments

1. The Markov model structure itself is very simple - and indeed the authors acknowledge in the Discussion that other disease states need adding. Noted. Keeping with the simple model structure is probably sufficient for the questions at hand, and the authors have worked hard to get the 'best' input parameters. My key quibble here though is heterogeneity by age. E.g. "Prevalence of IFG, IGT and at risk HbA1c was extracted from a UK-based study [27] and the annual probability of transitioning to T2DM was obtained from a meta-analysis, with different transition probabilities assumed for IFG, IGT and HbA1c [28]." Did these (and other) transition rates vary by age? This is not clear, and needs describing. If there was no variation in transition parameters by age, why not and what bias might this cause?

2. The detailing of the process against the AdViSHE protocol is good - very useful and reassuring. But I have a problem with the overall calibration and validation of the model. Simply saying "D4: Validation against empirical data: Have the model outcomes been compared to empirical data? No, we did not have access to alternative data sets." is not sufficient in my view. As I understand it, the authors have search hard for RRs and such like to disaggregate the population by states and rates of transition. But does it all come together as a whole to match the UK total population? There is a clue that is does not. The total QALYs per subject in Table 3 differs (quite markedly) between IGT and the IGF/HbA1c arms. I would expect the total QALYs in the comparator arm to be very close, across IGT, IFG and HbA1c arms. Something is amiss here, I think. Having parameterized the models for IGT, IFG and HbA1c arms, I think the authors need to then run them with no interventions on and ensure rates of DM incidence and mortality, etc, generated by the three models agree with each other and what is known to be occurring in UK population itself (including short term projections).
3. How are baseline rates of DM incidence inputted to model? And what are they? (The emphasis in the input parameter tables is on RRs, but the starting baseline rates need stating as well.)

4. There could be better clarity as to what comparator is used when. For example, the following sentence in the Abstract: "At the current NICE willingness to pay threshold of £20,000/QALY, intensive lifestyle interventions were cost-effective with ICERs of £3,707/QALY and £11,219/QALY for IGT and IFG respectively." Follows a sentence where comparator is null. I think this sentence with ICER means intensive compared to pragmatic - and indeed using the numbers provided by the authors in T.3 I got an ICER for intensive compared to pragmatic of f 818, close to 3707 given rounding.

5. Related, I think the results could be more clearly presented. Showing the CERs and ICERs for one intervention compared to another in Table 3 would help. And actually putting all the values on one CE Plane too would help the reader see the results (even if just for expected value analyses of all interventions - see comment below about uncertainty).

6. I note the authors preferred to place the PSA as appendix only analyses. I am not convinced! I think it important to convey to the reader the uncertainty in the main paper. Therefore, it would be good to hear why the authors are not routinely giving uncertainty intervals for the main study findings in Table 3, as cited in the Results, and even in the Abstract? Further, the input parameter tables should include the CI or uncertainty intervals around the input parameters that are subsequently probabilistically drawn from in the PSA (including the parametric form, e.g. beta, normal etc). Finally, I note the authors state that insufficient data exists to put uncertainty distributions around some input parameters, so deterministic sensitivity analyses were used only for these parameters …. Which means the plots in Appendix 5 are underestimating the scatter of points - yes? Should all parameters have probabilistic uncertainty in the PSA? (I think "yes").

7. Extending this, the authors usefully use their PSA and (presumably) a NMB approach to generate CEACs and report the probability of being cost effective at a £20,000 threshold. Namely, p.16: "Intensive lifestyle interventions were cost-effective relative to no intervention in 80.5%, 70.6% and 69.5% of simulations ..." Can the probabilities of cost effectiveness be given for intensive compared to pragmatic, not intensive compared to nil?

8. The costs for DM states come from ref 36. I think these are what I would call excess costs. That is, the costs of people living with pre DM. But inside the model, with the intervention turned on, people are living longer, and therefore generating more 'unrelated health care costs' to the system. From what I see, these unrelated health costs are not included, and should be. It will slightly increase net costs for interventions, slightly worsening cost effectiveness, I suspect. If I am wrong in my assumptions, then improved description is needed.

9. Evaluations of screening programmes should - usually - include the costs and consequences of the screening test itself. I do not think this study has included the costs of population-wide testing for IGT, IFG and HbA1c.... Which is more than just the test costs, but programme costs. And ideally the complications and 'harms' of such a programme (e.g. anxiety from being labelled, etc). Rather, I think the authors have started from the (magically!) identified risk population. There is an argument to do so; e.g. in my country CVD risk assessment over the age of 50 is very high and 'routine' primary care, and cost utility analysis may will not include this population-screening programmatic cost. But it is a (serious) moot issue - many activities in primary care are actually a screening programme, yet not considered as
such and arguably not evaluated as such. Some comments from the authors on this would be useful, I think.

Finally, overall this paper is very interesting and useful. I hope the above comments help improve the paper. I look forward to seeing it in print.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Yes

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable
**Declaration of competing interests**

Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal