Author’s response to reviews

Title: PRIMEtime CE: a multistate life table model for estimating the cost-effectiveness of interventions affecting diet and physical activity.

Authors:

Adam Briggs (adambriggs@doctors.org.uk)
Linda Cobiac (linda.cobic@dph.ox.ac.uk)
Jane Wolstenholme (jane.wolstenholme@dph.ox.ac.uk)
Peter Scarborough (peter.scarborough@dph.ox.ac.uk)

Version: 1 Date: 29 Oct 2018

Author’s response to reviews:

BHSR-D-18-00710

Comparing the cost-effectiveness of public health interventions to improve diet and physical activity using the PRIMEtime CE multistate life table model.

Adam Briggs; Jane Wolstenholme; Peter Scarborough

Following reviewers’ and editor’s comments, we are responding with two papers:

Paper 1


and

Paper 2

A D M Briggs, J Wolstenholme, P Scarborough. Estimating the cost-effectiveness of salt reformulation and increasing access to leisure centres in England, with PRIMEtime CE model validation using the AdViSHE tool
Response to reviewers

Editor Comments:

The paper as it stands without prior publication of the model seems problematic. As Reviewer #1 reasonably points out, the submitted manuscript should be broken down to two papers, unless it can be presented as a single article (I personally find it somehow hard to achieve successfully within the word limit of the journal) with the model preceding and the empirical application following. Please pay particular attention while addressing points 1, 3 and 6 of the same reviewer, which refer to the information provided for the model and the justification with respect to the research question. Please address all comments made by both reviewers thoroughly.

We agree with your suggestion to consider resubmitting this manuscript as two separate papers.

The first (referred to as paper 1 throughout this document) describes the model in detail, illustrating its main strengths and limitations. This includes addressing many of the points suggested by the reviewers. We have added Linda Cobiac to the authorship of this paper. As suggested by reviewer 1, we have added a more of a description of PRIMEtime, which Linda developed.

The second paper (referred to as paper 2 throughout this document) describes the two CE analyses which formed a large part of the original paper. In this second paper, we use the two CE analyses to validate PRIMEtime CE using the AdViSHE tool. Due to the level of detail included in the model validation, we have put much of the information regarding the comparison of PRIMEtime CE with other similar models into the supplementary data. We think that this means that the main text isn’t swamped with detail and therefore sticks to the high level points regarding model validation, and that readers can then refer to the supplementary materials for the detail.

Should peer-review be positive, we would suggest that these two papers are published alongside one another.

Reviewer reports:

Petros Pechlivanoglou (Reviewer 1): This is a well written paper describing a simulation model that could estimate the value for money of public health interventions in the UK as well as the cost-effectiveness of two hypothetical interventions on salt reduction and increase in physical activity. The model is comprised of a number of steps and relies heavily on a number of assumptions. In fact, I think that the study is so ambitious that it is trying to explain too many things at the same time. I believe that the work would benefit greatly if separate publications for
the model the PRIMEtime CE model and the CE of interventions were to be pursued. In that way, a lot of the material currently in supplementary appendices that have vital information to the structure of the model would be possible to be presented in the main text of the manuscript and properly peer-reviewed. As the authors mention in the limitations section, there are many assumptions in the model that affect the final estimates. These assumptions would need to be clearly presented in the main manuscript rather than the appendix. Below I provide specific comments throughout the manuscript:

Thank you for this suggestion. We agree and suggest publishing two separate manuscripts. This has meant we have been able to move much of the detailed discussion of the model to the main text of paper 1. We have then used this opportunity to add a substantial section on model validation to paper 2, including a detailed discussion in the supplementary material analysing results of PRIMEtime CE alongside three other models: the UK Health Forum model, the IMPACT CHD model, and the Markov model used by Frew et al.

1. The rationale for focusing on these two interventions is missing. Is it that the interventions serve only as an illustration? Have there been identified because of their potential for impact? what is the literature base for the effectiveness and cost-effectiveness of these two interventions? e.g. are there any other studies that look at the CE of salt reformulation?

More detail on this point has been added to paper 2. The interventions were selected following a discussion about priorities with local stakeholders. There are various other modelling studies that have looked at the CE of salt reformulation, and one study that has investigated the Be Active physical activity intervention. These are discussed in considerable detail in paper 2.

2. The estimates are presented as inclusive of 95% confidence intervals. Given the fact that a number of the input parameters did not include an estimate of uncertainty around them, or that the shape of the uncertainty for some estimates had to be imputed (e.g. through a triangular distribution) I am doubtful as to whether this uncertainty represents the real range of uncertainty around the estimates. In addition, no information on the methods underlying the uncertainty analysis is being described in the main manuscript. I would suggest that you move any discussion on methods of uncertainty analysis in the main manuscript

The methods for the uncertainty analyses have been included in the main manuscripts of both papers. Paper 1 includes a description of how uncertainty is included in PRIMEtime CE, and paper 2 includes a description of how uncertainty in each of the interventions is included (either in the main text or the supplementary material).
3. There is very little information on the effectiveness of the interventions in the manuscript. Regarding the salt reformulation, it is assumed that the intervention will bring salt consumption to the 2017 salt targets, however no assumption is being made on the compliance to such an intervention. In fact if I understand correctly the implicit assumption is that of perfect compliance. In addition, the governmental and industry costs are assumed to be lasting for three years, however, implementation of the salt reform would likely require long-term monitoring to ensure that the reform is actually taking place.

Detailed discussion of the limitations of the modelling of the two intervention scenarios has been added to the additional data file of paper 2.

4. Furthermore, the salt reformulation intervention needs to be placed better in context of the other studies on CE of salt reduction where a recent sys review indicated that the vast majority of past studies detected cost savings. doi:10.1017/S1368980017000593ss

This is discussed in much more detail in paper 2, and forms part of the model’s validation.

5. The Be Active intervention is informed by data that are quite different in nature by the salt reduction intervention. In particular, the BE Active intervention is accompanied by a measured relative effectiveness estimate that is directly associated to the intervention and not a hypothetical change to the effect modifier (salt or physical activity). In other words, this means that the observed finding in the Be Active intervention is more likely to be lower, given possible compliance issues but at the same time closer to a real world estimate. This discrepancy amplifies the notion that comparison across public health interventions are harder to be made given the heterogeneity of the input data.

We agree. This is a challenge for all comparisons of public health intervention and not something that is unique to PRIMEtime CE. The intention of modelling the impact of health and costs of public health interventions is to allow decision makers to consider what to invest in based on some, rather than no, data. The quality of the data should, where possible, be reflected in the uncertainty estimates. We have added comments to the discussion of papers 1 and 2 to reflect this point.

6. Information on the PRIMEtime multistate lifetable model is minimal. The authors cite the supplementary file of a different paper, which itself does not describe in great detail the structure of the model exactly, as well as the sources used for the model to be fitted. How well did the PRIMEtime model predict the incidence of disease in the UK? Is it a validated model? In my
opinion the multistate model and how this becomes a CE model needs to be article no 1 before the evaluation of the interventions.

On the basis of this comment we have converted the manuscript into two separate papers, with the first being a more detailed description of the PRIMEtime model and the second being a detailed description of the interventions simulated and PRIMEtime CE validation.

7. In the results section the authors report a negative ICER for the salt reformulation. A negative ICER can be misleading as it can imply either cost-savings or a dominated strategy. Also, the magnitude of a negative ICER is not easily interpretable. Please consider switching to a net benefit framework.

We agree this is misleading and have instead presented the return on investment.

8. The authors describe an intervention that is applied on a closed cohort. Given that the impact of the intervention is likely to affect more cohorts that the one alive when the decision of implementation will be taken, would there be value in applying the model on an open cohort instead?

The use of a closed cohort is consistent with other public health economic models. The intention of PRIMEtime CE is to allow the comparison of interventions, with the relative impact being as important as the absolute cost per QALY. Therefore, we do not think that the additional data and assumptions required to convert it into an open cohort justify the additional effort required and uncertainty it would generate.

9. The comparison to a 20/30K threshold is somewhat irrelevant once societal costs are taken into consideration.

Yes, this is a valid point however this threshold has been used to benchmark results and aid their comparison.

10. There are a number of sentences in the results section (e.g. line 337) that belong to the methods instead of the results.

There have been substantial changes to the methods and results section of paper 2 which should now address these issues.
10. I think it would be informative for the reader if you were to report a per person cost and QALY estimate next to the totals (e.g. line 296 table 2)

For the context of using the results to illustrate the model and for its validation, we have not added this level of detail, particularly given how we have expanded paper 2.

11. In line 261 in the sensitivity analysis, would you consider sensitivity analysis on the utility estimates? given the long term nature of these chronic conditions a small decrement difference in utility might have important implications to the CE estimates.

It would be possible to include sensitivity analyses around the utility estimates assuming that there were particular aspects of the utilities that were thought important to test (i.e. the age related decrement, or the utility decrement related to specific diseases). At the moment, PRIMEtime CE includes all the disutilities as estimated by Sullivan et al., which we think is the most honest way of interpreting the Sullivan et al. data. It is certainly possible within the model to “turn off” certain decrements, such as the age-related decrement, and we have added this option to the text in paper 1. In paper 2 we describe the model’s validation and as part of this we change the utilities used to match those used in other models. We discuss the implications of this in the results and discussion.

12. Along the same lines as above, the results of the sensitivity analysis in l.374 should in my opinion be in the main manuscript and not in the supplementary materials.

We have left these in the supplementary due to the volume of information in the main text of paper 2. However, we refer to these results in the discussion.

13. In line 191, how does this method capture costs in the presence of multi-morbidity do you double or triple count costs? is there a way to avoid it?

This costing method does not account for comorbidity. It is useful in that it ensures that the entire envelope of NHS expenditure is accounted for, however estimating disease specific patient costs in multimorbid patients is extremely challenging (there are very limited data on this), plus it would be inconsistent with the modelling method where diseases are assumed to be independent of one another.

We have updated the limitations section of the discussion in paper 1 to try and make this limitation clear both for costs and utilities.
Lennert Veerman (Reviewer 2): This paper uses a lifetable model to compare the potential health and economic impact of two preventive interventions in England. It is well-written and addresses an important topic that is of interest to the readership of BMC Health Services Research. It highlights the large differences in the likely return on investment between various options for prevention of non-communicable disease, and in the comparison with the study by Emma Frey et al, the enormous influence that choices in data and model structure can have on the findings, which elegantly points to the need for further development of guidelines and for comparative studies. I have a number of comments, but am generally very positive about this paper.

General comments

My first comment is that throughout the paper, negative ICERs are reported. That is a big no-no according to standard economic evaluation methods (see Drummond's textbook). Negative ICERs make no sense: an intervention that saves 1 life year and saves GBP10,000 has the same ICER as one that saves 100 life years and saves GBP 1,000,000, but clearly the latter is preferable. Since both health and costs are benefits here, they should properly be added, not one divided by the other. (And then there is the issue that the ICER for an intervention that costs 1 life year and costs GBP10,000 would also have the same ICER!) Generally, negative ICERs are reported as 'dominant' or 'dominated'. I would suggest relying more on the 'return on investment' (RoI) metric, and using those for both interventions. I don't see why that couldn't also be reported for the 'Expanding Be Active' scenario, even though it does not turn out to be cost-saving.

We agree and in paper 2, have reverted to reporting return on investment only.

Except that, secondly, the RoI seems to have been calculated only on the monetary components of the results, ignoring the health gains. That is unusual: the health gains represent an economic value, which is usually included in the RoI. Else, it is an incomplete metric. If you want to stick with that, it should be made clear in the methods and perhaps also in the discussion. Alternatively (and preferably), the health gains can be included. This requires quantifying the value of a QALY in monetary terms.

We have included this point in the paper 2’s results and discussion, including how it would potentially impact results.

The third general comment I have regards the choice of the time horizon. In the base case, this is 10 years. However, as line 116-7 points out, "the time horizon should be long enough to incorporate all important costs and effects". I would recommend adhering to the quoted NICE reference case as the base case. I understand that the stakeholders unfortunately did not have
such a long-term perspective, but this choice of time horizon does prevention no favours. For example, by this yardstick, don't bother preventing youth from taking up smoking. Very unfavourable ICER. And in this case, since the physical activity intervention benefits older people less than the salt intervention, this time horizon skews the comparison to the detriment of the former, the low 1.5% discount rate notwithstanding. If you do want to stick to this short time horizon, a stronger discussion of the implications of this choice should be included in the paper.

We have emphasized this point, and the results of the sensitivity analyses based on a life-time time horizon in the discussion in paper 2, and the impact of these decisions in terms of the model’s validation.

Specific comments

Please make clear in the main document that 'cancer' was breast and colorectal, separately modelled. (And not all cancers.)

Thank you, and changed throughout.

Line 166: Because the definition of 'case fatality' in DisMod II may not be clear to all, best define this (annual mortality rate among prevalent cases, or something similar). Else it could be mistaken for the proportion that dies within 28 days of incidence.

We have added this to the text and to table 1 in paper 1.

Line 192: Was the 'cost per prevalent case' also applied to cancers? In ACE Prevention, cost per incident case was used, as cancer-related costs mostly occur in the first years (after which cure or death has taken place, in many cases). With a short time horizon, this matters (though assigning all costs to incidence risks over-estimating gains, with a 10-year time horizon).

Yes, given the method used to calculate costs is based on the average costs across all cases in the population, there was no difference in costs used for incident or prevalent cases. Therefore, if an intervention has the effect of changing the average time between contracting a modelled disease and dying from the disease, modelled costs may overestimate (if the time with a disease lengthens) or underestimate (if the time with a disease shortens) actual healthcare expenditure. This has been made clearer in the discussion of paper 2.
Line 216: "Productivity gains and wider societal costs are included as sensitivity analyses." I thought productivity gains were part of the wider societal costs?

They are, sorry this is not clear. The first sensitivity analysis is adding in just productivity gains, and the second is all wider societal costs (including productivity gains). We have changed the text to try and ensure that this is clear.

Additional data file:

Page 5: Correct definition of MET to 3.5 ml, per minute.
Change made.

Page 7 "The parameters were unadjusted for obesity as this is assumed to act on the causal pathway": This is correct if no impact of physical activity on BMI was modelled. If that was the case, I suggest making that explicit by adding a sentence here.

We have changed the text as per this suggestion.

Table S7: These values are generally low. Even lung cancer leads to a decrement of 12% of quality of life. Is that relative to prior quality of life (which on average is less than 100% at any age), or are these percentage-points? The utility loss from breast cancer is particularly low, at 1.9%. Is there any explanation for that? To compare, the lowest value for breast cancer in the GBD 2016 set of DWs is 0.049 (0.031-0.072) for the controlled phase ('has a chronic disease that requires medication every day and causes some worry but minimal interference with daily activities'). Is this because prevalence is defined as 'ever having been diagnosed in the lifetime', in combination with a high rate of cure (without mastectomy, which comes with a DW of 0.083 in GBD)?

These are the absolute utility decrements, meaning that they are, in effect, percentage points and therefore in addition to the age related utility decrement. We have added a systematic review of breast cancer utilities to help validate the Sullivan et al. values used, and included a more detailed discussion of the utilities’ limitations.

Table S13: Again, remove negative ICERs, and add a RoI for 'Expanding Be Active' for comparison.

Changes made in paper 2.
Figures S8 and S9: What does 'Primary analysis' refer to?

We have made this clear in the results presented in paper 2.

Figure S10: Please remove negative ICERs. Best put it all in RoI. Consider doing the same for S12, or add a new graph with RoI for comparison.

Change made.