Author's response to reviews

Title: Validation of a model to investigate the impact of modifying cardiovascular disease (CVD) risk factors on the burden of CVD: the Rotterdam Ischemic heart disease and Stroke Computer simulation model (RISC) model.

Authors:

Bob JH van Kempen (b.vankempen@erasmusmc.nl)
Bart S Ferket (b.ferket@erasmusmc.nl)
Albert Hofman (a.hofman@erasmusmc.nl)
Ewout Steyerberg (e.steyerberg@erasmusmc.nl)
Erser B Colkesen (colkesen@zonnet.nl)
Matthijs S Boekholdt (S.M.Boekholdt@amc.uva.nl)
Nicholas J Wareham (nick.wareham@mrc-epid.cam.ac.uk)
Kay-Tee Khaw (kk101@medschl.cam.ac.uk)
Myriam MH Hunink (m.hunink@erasmusmc.nl)

Version: 2 Date: 23 July 2012

Author's response to reviews: see over
Reviewer: Simon Capewell

Reviewer's report:

Major Compulsory Revisions

1. It is not clear how the model, which seems to be a closed cohort simulation comprising age groups >50, could be used for exploring the effects of interventions on CVD burden in the general population. Please clarify.

One could argue that interventions on CVD burden in the general population should start at younger ages than 45 —although most guidelines recommend interventions from 40-50 years of age on. We more explicitly mentioned the usefulness of our model with regard to age in the discussion on page 11, which you also mentioned in your second point. The Rotterdam Study and EPIC study, both closed cohort population based studies, are not necessarily a valid representation of the general population. However, individuals were not selected based on risk factors or disease status (conditional on age), so do not represent a selected population with regard to disease risk or risk profiles a priori. We extended the description of the limitation these selected populations have on the usefulness of our model on page 11 of the discussion.

2. I think that the authors need to discuss more the limits that the age range imposes in their model usefulness. Although they included most of the relevant age groups for the CVD problem, underestimation of premature mortality, particularly among men could be a concern.

We more extensively discussed the limitation of the usefulness of our model with regard to age on page 11 of the discussion section.

3. Interesting difference in performance appear when applied to the EPIC data. Although differences in baseline risk and the calibration procedure itself may explain the discrepancies, could be a possible alternative explanation be differences in censoring in the observed data?? Please clarify

For the EPIC data, follow-up on mortality (CVD or non-CVD) was complete in the sense that “All individuals have been flagged for mortality at the UK Office of National Statistics, with vital status ascertained for the entire cohort” (Eur Heart J (2007) 28 (22): 2770-2777.doi: 10.1093/eurheartj/ehm390)

4. It will be VERY helpful to see the performance of the model presented by gender and age.

We thank the reviewer for his suggestion and have included an additional analysis which comprises the simulated versus the observed CVD mortality during years 1-13 for individuals stratified by tertiles of age x gender. We included graphs 11 and 12 in the appendix, extended the methods section on page 6 and 7, the results section on page 9 and included an additional paragraph in the discussion section on page 10. We choose not to provide information on (parameter) uncertainty around the simulated values as we did for the complete cohort as inherently smaller numbers would lead to wider intervals, which could lead to a too optimistic representation.

5. The authors appropriately addressed the lack of some variables in the EPIC cohort to fully reproduce the model. They suggest that the incremental value of these items is probably limited, and I agree with this. Could the model then be specified with the more limited set of risk factors and still perform well? This certainly will increase its potential for use in different populations and settings. It would be helpful to see this simpler analysis.

We agree with the reviewer that a simpler analysis would increase the potential of the RISC model to be adapted and validated in other populations, without losing too much performance. However the modification of the RISC model into a more simple model is something we intend to do, but takes a
respectable amount of time and work, and will not be subject to the current paper. We have extended the discussion section on page 12 with regard to this remark.

6. The authors discuss model validation using a framework that only explores part of the concept. There is a more comprehensive framework published by Kopec et al (see BMC Public Health 2010, 10:710 doi:10.1186/1471-2458-10-710) that covers many of the validation exercises the authors conducted. It will be really useful if the authors can address or discuss some of the other validation aspects that are important and suggested by Kopec et al in their paper.

We thank the reviewer for this suggestion. A number of aspects from the more comprehensive framework by Kopec et al were covered by an earlier paper by Nijhuis et al (Nijhuis Med Decis Making. Mar-Apr 2006;26(2):134-144): conceptual model, parameters, computer implementation, face validity, internal consistency. We additionally addressed between-model validity on page 12 of the discussion section. In essence, we covered analyzing the consequences of modeling based decisions in the penultimate paragraph of the discussion on page 12.

Minor Essential revisions:

7. The penultimate sentence in the 3rd paragraph of the discussion citing references 22 and 31 seems to be incomplete.

We thank the reviewer for pointing out and have changed it accordingly.

Reviewer: Peter Scarborough
Reviewer’s report:
This is a very good article. The authors are clear about their objectives and the reporting of the methods and the results are clear. I only have a few discretionary revisions and one minor essential revision for the authors to consider:

Discretionary revisions:

1. The RISC model relies on data from a number of input variables that are not commonly collected outside of major epidemiological studies (e.g. fasting glucose, plasma creatinine level, ankle-brachial index). Presumably this limits the use of the model to specific populations, and the model may not be effective at estimating risk of CVD in general national populations. The authors could possibly mention this as a limitation at some point in the discussion.

We agree with the reviewer that a simpler analysis would increase the potential of the RISC model to be adapted and validated in other populations, without losing too much performance. We have extended the discussion section on page 12 with regard to this remark.

2. It is not clear why both PTCA and CABG are included as measures of acute MI in the RISC model - although they may help identify MI they are clearly not acute MI. This seems to be a problem as it means that the MI estimates cannot be compared with EPIC Norfolk (the other reason given, that EPIC only records hospitalised MIs and deaths, does not seem relevant as the vast majority of MIs in the UK will result in either hospitalisation or death). Again, the authors may want to comment on this in the discussion.

PTCA and CABG are not included as measures of acute MI but are included in our definition of CHD. The RISC model was designed to simulate CHD as a combined endpoint, and not MIs (and other CHD events) separately. This was done in accordance to most clinical trials outcomes, which more often than not use CHD (as a combined endpoint) as their primary outcome. We addressed this in the discussion on page 11

3. It does not seem clear to me why the authors include the external validity results for the non-recalibrated RISC model. It seems obvious to me that the model should be recalibrated to the baseline level of risk in the EPIC cohort.
Although it is obvious to the informed reader, we feel it is illustrative for the process of model validation, which contributes to the aim of our paper.

4. The authors could drop the reference to utility parameters on page 5, as QALY estimates are not included in the results.

We perhaps misunderstood the reviewer’s remark here, but on page 5 (or elsewhere in the manuscript) we have not referred to QALYs. On page 5, we refer to the design of the Rotterdam study (15), the analysis of uncertainty and heterogeneity in simulation models by Groot-Koerkamp et al (ref 16,17)

5. On page 7 the analysis technique is described. 100 sets of parameters are drawn. For each set of parameters, 2000 sets of risk factor profiles are drawn. For each set of risk factor profiles, 200 random walks are simulated. This seems like there were 100x2000x200 =40,000,000 simulations run for each analysis. Have I understood that correctly?

Yes. We included an additional sentence on page 7 of the methods.

Minor essential revision:
1. On page 10, the authors refer to 6 major models, of which 3 have not been validated, 2 have only been internally validated and 2 have performed an external validation. 3+2+2=7, not 6.

There was overlap between the types of validation: one of the models was only externally validated (Gunning-Schepers), so 6 is correct. We removed ‘only’ in this sentence since this is incorrect.