Dear Miss Damen

Manuscript ID BMJ.2015.031028 entitled “Systematic review of prediction models for cardiovascular disease risk in the general population: too many models, too few validated and often insufficient reporting”

Thank you for sending us your paper. We sent it for external peer review and discussed it at our manuscript committee meeting. We recognise its potential importance and relevance to general medical readers, but I am afraid that we have not yet been able to reach a final decision on it because several important aspects of the work still need clarifying.

We hope very much that you will be willing and able to revise your paper as explained below in the report from the manuscript meeting, so that we will be in a better position to understand your study and decide whether the BMJ is the right journal for it. We are looking forward to reading the revised version and, we hope, reaching a decision.

dr. Wim Weber
European editor, The BMJ
wweber@bmj.com

https://mc.manuscriptcentral.com/bmj?URL_MASK=c94c4b65df3c49acbe2d85f6e978015

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

These comments are an attempt to summarise the discussions at the manuscript meeting. They are not an exact transcript.

Members of the committee were: Elizabeth Loder (Chair), Tim Cole (Statistics advisor), Jessamy Bagenal, José Merino, Georg Rögglar, Tiago Villanueva, Wim Weber.

Decision: Put points

Detailed comments from the meeting:

Yours is an exhaustive systematic review of CVD prediction models. It makes a persuasive case that there are too many, often of poor quality. But readability of the review suffers somewhat from the same problem, with many models to describe. You have designated three classes: new models, validations of existing models, and extensions with new predictor(s). Reviewer 1 makes a good case to omit new predictor studies, as requirements are less than for other categories. This would focus the paper better and shorten it.

Please change the title to reflect what research you did, it should not announce the findings.

First, 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.

In your response please provide, point by point, your replies to the comments made by the reviewers and the editors, explaining how you have dealt with them in the paper.

Comments from Reviewers

Reviewer: 1

Recommendation:

Comments:
A very thorough investigation that shows that the situation is much more chaotic than I had thought. However, I am not convinced that there is a way, given sufficient “head-to-head” comparisons, of finding one, or a few, superior risk-assessment method(s) from among those available; this seems to be the unspoken belief driving the paper. I guess a lot of people have their doubts about, eg QRISK2, despite its adoption in NICE guidelines, but at least it’s a rough-and-ready, intuitively correct, locally-acceptable approach (as all public-health measures can only be) to a national treatment protocol. (I am thus certainly of the authors’ opinion that no more time should be spent on refining existing models).
Is it worth publishing? Yes, if you want authoritative evidence for what most cardiologists (and a lot of patients) believe. No, if you hope it will lead to a better, tidier, more rational, etc, future.

Additional Questions:
Please enter your name: John Walsh
Job Title: Patient representative
Institution: -
Reimbursement for attending a symposium?: No
A fee for speaking?: No
A fee for organising education?: No
Funds for research?: No
Funds for a member of staff?: No
Fees for consulting?: No
Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No
Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No
If you have any competing interests (please see BMJ policy) please declare them here:
Reviewer: 2
Recommendation:
Comments:
In their review of CVD prediction models, Damen et al. have undertaken an ambitious task of describing an extremely heterogeneous field of research. Adding importance to the subject is the amount of resources invested in this area of science.

Although, the field of the article is important, the article in itself is of moderate importance. However, the data described in the article is interesting because it describes well the poor quality of most of the research articles in this field. Nevertheless, it probably will not help most readers of a general medical journal to make better decisions as most of the suggestions derived from the conclusions are directed to those who design and supervise the general population-based studies on which most risk prediction models are based. Most often their work is limited, not by scientific aims or lack of vision, but by practical problems such as funding etc. The manuscript adds to existing knowledge by providing and update on the field but offers no big surprises.

The research question seems clearly defined and well answered and the overall design of study is appropriate. The outcome measure is to provide descriptive data and mostly this is done adequately. The manuscript is well written and the methodology behind the literary research and the review itself seems sound.

The foremost problem of the manuscript is the heterogeneity of the subject it aims to describe. Mostly the conclusions that the authors propose are valid but they seem to be based on assumptions that all developed models (as defined by the authors) and incremental value assessments have been genuinely published for the purpose of competing with the pre-existing juggernauts dominating the field (for example SCORE, Framingham risk score and QRISK) or at least trying to offer a valid option for clinical work. In clinical practice, the existing validated models (SCORE, Framingham risk score and QRISK) are already implemented and no controversy exists of their replacement by new risk prediction models (with no validation).

Major concerns:
1. Could the authors describe more accurately the articles (and their purpose) currently defined as “ describing the development of new risk factor models”? One of the main concerns is that the models labelled “developed models” actually intended to compete with or outperform existing and well-established models (SCORE, PRS QRISK2)? If not, this could explain much of the shortcomings in the reporting described in the current report. As the authors present that only 46% of the complete regression formula was published and only 22% were internally validated. These facts are reflected in the fact that only 64% of the developed models were externally validated. If these models are indeed developed to compete with pre-existing models in clinical use, the main conclusion of the current review could be simply that the methodology of most published studies (focusing on developing new risk models) should be improved.

2. The conclusion that “Future research should focus on assessing the validity and impact of existing models…” (abstract, page 5, line 130) seems to restricting. According to the manuscript, most developed models (66%) included the same set of major predictors (age, smoking, blood pressure, and blood cholesterol measurements) of CVD. It is reasonable to presume that examining the risk coefficients (effects) of these major risk factors in specific populations (if validated properly) can lead to more accurate prediction of CVD than using previous coefficients developed for generalization to all populations (within a continent as in SCORE) or developed in completely different ethnic population. Many variables (even the most common ones) used in risk prediction can vary in their definition between populations and they can also have different sized effects in different populations. The registry data maintained globally is highly heterogeneous in quality but in most countries with high accuracy of registry data, the applicability of such data can be very high for risk prediction. With constantly improving data management and electronic patient registries, could it be possible that specific population-level, easily available and monitored risk prediction models, can out-perform the old risk models based on fixed coefficients and baseline hazard?

3. The authors have also searched for articles that describe the incremental value of a new predictor. The problem is that such a publication rarely even tries to specifically present new models for validation and/or to compete with existing validated models. However, in such case it is often a better option to use the best fit model in the background than try to implement pre-existing models with bad fit (due to any reason). Why not exclude these models from the current review?

4. The problem with including models that were used to calculate incremental value assessments of added risk markers is that they are often (due to practical reasons related to the publication demands) limited in their presentation. They distort the statistics describing the quality of reporting. For example, in abstract, the authors report (page 5, line 124): “…and for 92 (25%) crucial information was missing. Are these publications (of incremental value assessments) included in this figure presented in abstract? Currently in results section, these models (for incremental value assessment) are described only very briefly. I would recommend excluding these article from the review or not including them with summary statistics in abstract.

5. As the authors concede, the literary research is already lagging behind. However, they justify the now 2.5 year interval between the literary search and current date by stating that: “including these articles is unlikely to change our main conclusions and recommendations”. If the purpose of this review was to provide a comprehensive view on the subject, I feel an update
would be in order. Most of the valid conclusions could be made even without the information provided by the current manuscript. In the present manuscript the authors cite the work of Beswick et al. (ref. 12) form the year 2008 (not from ten years ago as the authors claim). How have the conclusions changed since? If the last 2.5 years from June of 2013 to this day are not necessary for updating the message of the present manuscript, how has the field changed from 2008?

Minor issues:

1. How many developed models have incorporated the use of statins or other medications such as blood pressure lowering medication or family history of CVD.

Jussi Hernesniemi
MD, PhD
University of Tampere

Additional Questions:
Please enter your name: Jussi Hernesniemi
Job Title: Physician
Institution: University of Tampere
Reimbursement for attending a symposium?: No
A fee for speaking?: No
A fee for organising education?: No
Funds for research?: No
Funds for a member of staff?: No
Fees for consulting?: No
Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No
Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No
If you have any competing interests (please see BMJ policy) please declare them here:

Reviewer: 3
Recommendation:

Comments:
Members of my family have heterozygous familial hypercholesterolemia (HeFH) and I have thought quite a bit about risk assessment for the primary prevention of cardiovascular disease. HeFH is a disease characterized by high risk of premature heart disease, but the risk is highly variable, with some untreated persons suffering myocardial infarction as early as their 20s or 30s and others not until late in life (e.g., 70s or older) or not at all. A number of studies have been done to try to discover the factors associated with premature heart disease in HeFH, but no model has been developed that explains more than a small part of the variance in risk and no validated risk calculator exists. As a result, most guidelines call for all HeFH patients to be treated starting in childhood, which results in overtreatment for some patients.

In addition, I am familiar with some of the risk calculators available for the general population and have used them myself to calculate my risk of an event (I do not have HeFH myself so am eligible to use most of these risk calculators). I have noticed that the various calculators vary with respect to the predictive variables that are included and the endpoints being assessed, as well as in the level of risk predicted. All in all, our ability to predict who will have a first cardiovascular event in the future is limited, meaning that any person wishing to lower their risk of such an event through interventions such as blood pressure and cholesterol lowering is rolling the dice in terms of whether they will benefit. I am also aware that many studies have been done in recent years on the effect of adding additional predictors to Framingham or other commonly used risk calculators and that the additional predictors usually only have a small additional effect, if any, over traditional risk factors alone. One possible exception is coronary artery calcium, but its use is this context is controversial. My impression has been that this unsatisfactory state of affairs is mainly due to the inherent unpredictability of the human cardiovascular system, rather than to a failure of research methods and reporting. For example, acute myocardial infarction results from sudden plaque rupture and clot formation, which partially or completely blocks the flow of blood through a coronary artery. Currently, our ability to predict the timing of such a sudden event is quite limited. Improvement in our ability to predict such events is would be very helpful to patients and their physicians as it would allow greater targeting of preventive interventions such as medications and lifestyle changes, and less under and overtreatment.

The authors of this systematic review have made an attempt to survey the field of prediction modeling for cardiovascular disease risk and outline some of the major shortcomings and difficulties in the field. They found quite a few, and make some specific suggestions in terms of improving methods of modeling, reporting, and research focus. For example, they suggest that the field would benefit from the formulation of a guidance with clear definitions of the relevant outcomes, predictors, and prediction horizons. I find myself unable to form a definite opinion on their suggestion that less emphasis be placed on the development of new CVD prediction models. In particular, my impression is that most validated models are based mainly on
Caucasian and North American or European populations and that models are needed for the less developed countries and for non-Caucasian populations.

In general, this paper seems mainly aimed at researchers in the field of cardiovascular risk modeling. As written, I do not think it will be useful to most patients or GPs. To make it more user-friendly, I suggest that definitions of technical terms such as discrimination, calibration, and validation be provided. In addition, summary points would be helpful. A bit more explanation for some of their conclusions is needed as well. Parts of the paper read like an editorial but I found myself relatively unpersuaded by some of their points, such as the need to refrain from developing new models, but willing to listen if more reasons were given. It also would be helpful to know more about how and why the current state of affairs developed, e.g., file drawer effect, perverse incentives, funding priorities, etc.

Additional Questions:
Please enter your name: Marilyn Mann
Job Title: patient reviewer
Institution: NA
Reimbursement for attending a symposium?: No
A fee for speaking?: No
A fee for organising education?: No
Funds for research?: No
Funds for a member of staff?: No
Fees for consulting?: No
Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No
Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here:

Reviewer: 4
Recommendation:
Comments:
Systematic review of prediction models for cardiovascular risk in the general population: too many models, too few validated and often insufficient reporting

Originality
Whilst Beswick and colleagues (Beswick AD, Brindle P, Fahey T, Ebrahim S. A Systematic Review of Risk Scoring Methods and Clinical Decision Aids Used in the Primary Prevention of Coronary Heart Disease (Supplement). London: Royal College of General Practitioners., 2008) have published a systematic review the conclusion from that report “The generalisability of risk scores in different populations cannot be assumed and there is currently little evidence that they are effective in improving health outcomes. Future research needs to be directed at refining the accuracy of prediction models and, most importantly, examining ways of turning them into effective clinical tools.” indicates that the objective and conclusions of the submitted paper are substantially different. I therefore believe it does add to the published literature and should be seen as original work.

By highlighting the plethora of prediction models and documenting their shortcomings, particularly around external validation, the authors provide much need guidance to researchers (and funders) on what should be the priorities for future studies. Secondarily the results would be of interest to clinicians and policy makers with an interest in cardiovascular health who are potentially the end users of any developed models. They need to understand the limitations and differences between prediction models and be able to utilise ones appropriate for their populations.

Scientific reliability
The research objective is clearly defined.

Study design
The authors have undertaken a systematic review, albeit a descriptive one, and this is appropriate for the study question/objective.

Introduction
The introduction is brief, acknowledges previous published research and provides good justification for the current study.

Methods
The methods are clearly and adequately described. Although they have only searched from 2004 to 2013 I think their reliance on a previous review is acceptable given the nature of the topic. Having an end point is also important and they do acknowledge, when discussing the study limitations in the discussion (line 436), the large number of subsequent publications that have occurred after 2013. I agree with them that it is unlikely these publications would change their main conclusions or recommendations.

Their search strategy, screening and data extraction methods are consistent with other studies.

I agree that descriptive analysis was the best approach for the available data and current methods.

Results

The results were very interesting (surprising) even for a researcher/ and clinician who works in the field. For me there were several take home messages that I can use in my research, teaching and clinical work. Firstly the sheer number of studies, the fact that only a few have had extensive external validation, the wide variation in predictive performance, the inconsistency in outcome measures, the limited use of assessing the incremental value of predictors (despite the large number of predictors used) and the paucity of validated prediction models in some geographic areas.

Discussion

The discussion and conclusions are well written and consistent with the results. Although long I believe it would be hard to reduce its length and still do justice to the extensive results. The authors make sensible and logical recommendations that could potentially impact cardiovascular health care across the globe. I believe the messages are clear.

The limitations are acknowledged and discussed. It is important that the authors reference the limited evidence about the effect of risk prediction tools clinically and in particular on actual CVD incidence.

References

Appear to be up to date.

Abstract/summary/key messages/What this paper adds

Reflect accurately the paper, results and discussion.

Additional point

Clearly the authors have enough results and recommendations for one paper but I did wonder if, by using the collected data, if they could make recommendations about which prediction models did not require further validation in specified populations (North America/Europe) and which were most promising, based on current research, for populations that had previously been neglected (for instance Africa/Asia/certain subgroups).

Prof Nigel Stocks
University of Adelaide

Additional Questions:
Please enter your name: Nigel Stocks
Job Title: Head Discipline of General Practice
Institution: University of Adelaide
Reimbursement for attending a symposium?: No
A fee for speaking?: No
A fee for organising education?: No
Funds for research?: No
Funds for a member of staff?: No
Fees for consulting?: No
Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No
Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No
If you have any competing interests (please see BMJ policy) please declare them here:

Reviewer: 5
Recommendation:

Comments:
I enjoyed reading this paper which gives a comprehensive and systematic review of the literature concerning CVD prediction models. This work is timely, as such a review was last carried out by Beswick et al over 10 years ago. The findings are consistent with that review - highlighting a lack of validation and of impact on health outcomes despite a multitude of models.
Whilst this current paper comments on the lack of impact (Conclusions in the Abstract) we need to be clear that this systematic review included model development and validation studies, and studies of incremental variable inclusion, but not impact studies. It benefits from the more recent 2014 publication of the CHARMS checklist.

In terms of methodology the paper is of high quality, although I thought for today's practice a stronger justification might have been given for excluding models predicting stroke events in atrial fibrillation (AF). At the time of the Beswick review, which was concerned anyway with CHD rather than broader CVD outcomes, use of risk scores for stroke in AF were less common, but in the time since then they have become a very important issue in clinical practice and the associated literature might have added significantly to this work.

Finally, whilst I agree that the number of prediction models seems excessive, the authors reach this conclusion without very much discussion on why this is so and what the motivation is to produce yet more models (apart from producing more research papers of course!) But there might have been a case for new models either because 1) the population characteristics or setting differ in ways relevant to model adequacy; and 2) the data source itself might determine the suitability of the modelling approach. There is, after all, a huge difference between the data used to derive the Framingham equation and that used to support the derivation of QRISK2. One is based on relatively small but high quality prospectively collected cohort data and the other on routinely collected data from UK primary care. The differences between the models are not therefore just due to differences in the populations concerned. And because the wide use in practice of Framingham across diverse populations different from the source cohort has attracted much criticism, it is perhaps not surprising that people have felt justified in modifying it regionally, or indeed in starting from scratch.

I hope these comments are useful.

Tim Holt

Additional Questions:
Please enter your name: Tim Holt
Job Title: Senior Clinical Research Fellow
Institution: Oxford University
Reimbursement for attending a symposium?: No
A fee for speaking?: No
A fee for organising education?: No
Funds for research?: No
Funds for a member of staff?: No
Fees for consulting?: No
Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No
Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No
If you have any competing interests (please see BMJ policy) please declare them here:

**Information for submitting a revision**

Deadline: Your revised manuscript should be returned within one month.

How to submit your revised article: Log into http://mc.manuscriptcentral.com/bmj and enter your Author Center, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision.

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript using a word processing program and save it on your computer. Once the revised manuscript is prepared, you can upload it and submit it through your Author Center. When submitting your revised manuscript, you will be able to respond to the comments made by the reviewer(s) and Committee in the space provided. You can use this space to document any changes you make to the original manuscript and to explain your responses. In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the reviewer(s). As well as submitting your revised manuscript, we also require a copy of the manuscript with changes highlighted. Please upload this as a supplemental file with file designation 'Revised Manuscript Marked copy'. Your original files are available to you when you upload your revised manuscript. Please delete any redundant files before completing the submission.

When you revise and return your manuscript, please take note of all the following points about revising your article. Even if an item, such as a competing interests statement, was present and correct in the original draft of your paper, please check that it has not slipped out during revision. Please include these items in the revised manuscript to comply with BMJ style (see: http://www.bmj.com/about-bmj/resources-authors/article-submission/article-requirements and http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists).
Items to include with your revision (see http://www.bmj.com/about-bmj/resources-authors/article-types/research):

1. What this paper adds/what is already known box (as described at http://resources.bmj.com/bmj/authors/types-of-article/research)

2. Name of the ethics committee or IRB, ID# of the approval, and a statement that participants gave informed consent before taking part. If ethics committee approval was not required, please state so clearly and explain the reasons why (see http://resources.bmj.com/bmj/authors/editorial-policies/guidelines.)

3. Patient confidentiality forms when appropriate (see http://resources.bmj.com/bmj/authors/editorial-policies/copy_of_patient-confidentiality/)

4. Competing interests statement (see http://resources.bmj.com/bmj/authors/editorial-policies/competing-interests)

5. Contributorship statement+ guarantor (see http://resources.bmj.com/bmj/authors/article-submission/authorship-contributorship)

6. Transparency statement: (see http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/transparency-policy)

7. Copyright statement/licence for publication (see http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse)

8. Data sharing statement (see http://www.bmj.com/about-bmj/resources-authors/article-types/research)

9. Funding statement and statement of the independence of researchers from funders (see http://resources.bmj.com/bmj/authors/article-submission/article-requirements).

10. Patient involvement statement (see http://www.bmj.com/about-bmj/resources-authors/article-types/research).

11. Please ensure the paper complies with The BMJ's style, as detailed below:

a. Title: this should include the study design eg “systematic review and meta-analysis.”

b. Abstract: Please include a structured abstract with key summary statistics, as explained below (also see http://resources.bmj.com/bmj/authors/types-of-article/research). For every clinical trial - and for any other registered study- the last line of the abstract must list the study registration number and the name of the register.

c. Introduction: This should cover no more than three paragraphs, focusing on the research question and your reasons for asking it now.

d. Methods: For an intervention study the manuscript should include enough information about the intervention(s) and comparator(s) (even if this was usual care) for reviewers and readers to understand fully what happened in the study. To enable readers to replicate your work or implement the interventions in their own practice please also provide (uploaded as one or more supplemental files, including video and audio files where appropriate) any relevant detailed descriptions and materials. Alternatively, please provide in the manuscript urls to openly accessible websites where these materials can be found.

e. Results: Please report statistical aspects of the study in line with the Statistical Analyses and Methods in the Published Literature (SAMPL) guidelines http://www.equator-network.org/reporting-guidelines/sAMPL/. Please include in the results section of your structured abstract (and, of course, in the article's results section) the following terms, as appropriate:

i. For a clinical trial: Absolute event rates among experimental and control groups; RRR (relative risk reduction); NNT or NNH (number needed to treat or harm) and its 95% confidence interval (or, if the trial is of a public health intervention, number helped per 1000 or 100,000.)

ii. For a cohort study: Absolute event rates over time (eg 10 years) among exposed and non-exposed groups; RRR (relative risk reduction.)

iii. For a case control study: OR (odds ratio) for strength of association between exposure and outcome.

iv. For a study of a diagnostic test: Sensitivity and specificity; PPV and NPV (positive and negative predictive values.)

v. For a systematic review and/or meta-analysis: Point estimates and confidence intervals for the main results; one or more references for the statistical package(s) used to analyse the data, eg RevMan for a systematic review. There is no need to provide a formal reference for a very widely used package that will be very familiar to general readers eg STATA, but please say in the text which version you used.

f. Discussion: To minimise the risk of careful explanation giving way to polemic, please write the discussion section of your paper in a structured way. Please follow this structure: i) statement of principal findings of the study; ii) strengths and weaknesses of the study; iii) strengths and weaknesses in relation to other studies, discussing important differences in results; iv) what your study adds (whenever possible please discuss your study in the light of relevant systematic reviews and meta-analyses); v) meaning of the study, including possible explanations and implications for clinicians and policymakers and other researchers; vi) how your study could promote better decisions; vi) unanswered questions and future research

g. Footnotes and statements

Online and print publication: All original research in The BMJ is published with open access. Our open access policy is detailed here: http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse. The full text online version of your article, if accepted after revision, will be the indexed citable version (full details are at http://resources.bmj.com/bmj/about-bmj/online-access-publishing-model). The print and iPad BMJ will carry an abridged version of your article. This abridged version of the article is essentially an evidence abstract called BMJ pico, which
we would like you to write using the template downloadable at http://resources.bmj.com/bmj/authors/bmj-pico. Publication of research on bmj.com is definitive and is not simply interim “epublication ahead of print”, so if you do not wish to abridge your article using BMJ pico, you will be able to opt for online only publication. Please let us know if you would prefer this option. If your article is accepted we will invite you to submit a video abstract, lasting no longer than 4 minutes, and based on the information in your paper’s BMJ pico evidence abstract. The content and focus of the video must relate directly to the study that has been accepted for publication by The BMJ, and should not stray beyond the data.

END

Date Sent: 05-Feb-2016