ARIC Manuscript Proposal #3503 (Revised)

PC Reviewed: 12/10/19	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: The value of additional risk factors in addition to current prediction rules to better predict 10-year cardiovascular risk

b. Abbreviated Title (Length 26 characters): flexible risk prediction

2. Writing Group:

Writing group members:

- Department of Vascular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands: Drs. Steven Hageman, Prof Frank Visseren, Dr. Jannick Dorresteijn.
- Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands: Prof. Rolf Groenwold
- Julius Center Research Program Cardiovascular Epidemiology, Utrecht, The Netherlands: Prof. Y.T. van der Schouw
- Johns Hopkins Bloomberg School of Public Health (ARIC and MESA-studies): Dr. Kunihiro Matsushita, Dr. Michael Blaha,
- Michael E DeBakey Veterans Affairs Hospital and Baylor College of Medicine (ARIC): Dr. Vijay Nambi
- University of Cambridge School of Clinical Medicine (EPIC-Norfolk study): Prof. Kay-Tee Khaw, Prof. Nick Wareham
- German Institute of Human Nutrition Potsdam, Germany (EPIC-Potsdam study): Prof. Matthias Schulze
- University of Duisburg-Essen (Heinz-Nixdorf Registry): Prof. Raimund Erbel

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. SH **[please confirm with your initials electronically or in writing]**

First author: **S.H.J. Hageman, MD**

Address: University Medical Center Utrecht
Department of Vascular Medicine
Internal mail no F02.126
Phone: + 31 88 75 556 53
E-mail: S.H.J.Hageman-4@umcutrecht.nl

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).

Name: Kunihiro Matsushita
Address: Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health
Welch Center for Prevention, Epidemiology, and Clinical Research
2024 E. Monument St., Suite 2-600 (Rm 2-602), Baltimore, MD 21287
Phone: (443) 287-8766

Fax: (410) 367-2384
E-mail: kmatsus5@jhmi.edu

3. Timeline: We aim to have a full manuscript approximately 6-9 months after acquiring all necessary data.

4. Rationale:
The use of prediction models for estimating cardiovascular risk is recommended by European and American guidelines[1,2]. Several prediction models are available in the primary prevention to estimate lifetime treatment benefit like the life-CVD model[3] or to calculate 10-year cardiovascular risk, e.g. the SCORE-model[4] and the ASCVD pooled cohort equation[5].These models are widely-used and practical because they use easy to measure and generally available risk factors to calculate 10-year cardiovascular risk. In clinical practice however, often other risk factors are known apart from those in the prediction model, for example family history, body mass index (BMI), estimated glomerular filtration rate (eGFR), albuminuria, social-economic status, coronary calcium score, ankle/brachial-index, etc. Both the ESC and the ACC/AHA guidelines acknowledge several risk factors that may enhance risk prediction and recommend those to be considered in patients with intermediate risk or a risk close to a treatment threshold, although no clear solution is given on how to deal with these predictors[1,2]. The goal of the current analysis is not to select on possible predictors in order to develop a new model, but to evaluate a complete list of clinically available predictors on their added value on top of a basic model. Addition of several factors will increase predictive accuracy. For the other variables, clinical applicability of prediction models will benefit from the result that the presence or absence of a variable would not change the predicted risk.

5. Main Hypothesis/Study Questions:
Aims
1. To quantify the relative effect of additional risk factors in addition to a basic model for the prediction of 10-year cardiovascular mortality risk in apparently healthy people.
2. To validate the accuracy, calibration and reclassification potential of combining a basic prediction model with a variable number of additional risk factors for the prediction of 10-year cardiovascular mortality in apparently healthy people.

6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

Study population
Patients originate from several European and North-American cohorts: the Atherosclerosis Risk in Communities (ARIC, 4th visit as a baseline, N=11,656),[6] Multi-Ethnic Study of Atherosclerosis (MESA)-study (N=6,814),[7] European Prospective Investigation into Cancer (EPIC-NL, N=40,011), EPIC-Norfolk(N=21,453),[8] EPIC-Potsdam (N=27,548),[10] Heinz Nixdorf Recall (N=4,200),[9] and the primary care database CPRD (N= 4,425.016).[11] Patients at least 40 years old will be included. Exclusion criteria are prior CVD or heart failure or missing data on one of the original SCORE variables (age, gender, smoking, blood pressure, total cholesterol).

Predictors

To the predictors of the original SCORE-model (age, gender, smoking, blood pressure, cholesterol level), we will add the following predictors, depending on availability per cohort: albuminuria, ankle-brachial index, atrial fibrillation, auto-inflammatory disease, BMI/central obesity, carotid IMT and plaque, CVD family history defined as and a positive history of premature (prior to age 60) myocardial infarction (MI) in either parent, DM, education level, eGFR, ethnicity, hs-CRP, hs-troponin, menopausal age/status, Lp(a), NT-proBNP, number of medications, obstructive sleep apnea syndrome, physical activity, , waist circumference.

Statistical analysis

Using the SCORE-predictors as stated before, a Cox proportional hazards model will be created. We will use a Cox model instead of a Weibull model as was originally used for SCORE,[4] because the Cox model does not have the same assumptions about the shape of the survival curve. Also, as most clinical studies use Cox models, using a Cox model will allow our methodology to be applied to other studies more easily. For optimal fit, we will refit the model for every dataset.

The additional predictors will be added using the "naïve approach",[12,13] which gives predictions based by multiplying baseline survival with additional hazard rates instead of using regression modelling to predict individual risks. The independent hazard rates and population frequencies will be calculated from the cohorts. These, and the baseline individual predicted risk from the SCORE-parameters will be used in the following formula: *individual predicted risk^(hazard ratio/population relative risk),* where the population relative risk is equal to *(prevalence of a factor)*HR of the factor + (1-prevalence).* All variables will be dichotomized or categorized for easier clinical use.

For all predictors the hazard rates will be calculated in all cohorts where they are available. The results will be pooled to a single hazard rate. As the availability in cohorts differs per predictor, a different data pool may be used for different predictors. In all cohorts, the updated model will be assessed on test accuracy (c-statistic), calibration (predicted versus observed risk plots) and reclassification potential (net reclassification index). Model performance will be assessed both for addition of all predictors separately, but also for applying multiple multiplication factors simultaneously. Additional predictors that have too much correlation will not be added together. As the availability itself of additional predictors may carry predictive value, the naïve approach will be validated in the routine care database CPRD. The results of the naïve approach will be compared with a reduced modelling strategy, in which the complete model is refitted based on all different combinations of availability of predictors. All analyses will be performed in R-Statistic Programming (R Foundation for Statistical Computing, Vienna, Austria).

7.a. Will the data be used for non-CVD analysis in this manuscript? ____ Yes __X_ No

8.a. Will the DNA data be used in this manuscript? __X__ Yes ____ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = "No use/storage DNA"? ___X_ Yes ____ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status. ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/aric/mantrack/maintain/search/dtSearch.html

___X___ Yes _____ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?

MP1823 - Interaction of chronic kidney disease with classical cardiovascular risk factors. Pooled analysis of general population cohorts; Matsushita, K
In this proposal, an extensive analysis is performed assessing the effect of albuminuria and eGFR on the risk of renal and cardiovascular outcomes. Detailed analyses are performed including non-linear relations and interactions with other risk predictors. In comparison to this previous proposal, the current proposal aims to evaluate a complete list of potential risk modifiers and to study the methodology regarding the addition of several predictors. The lead author of this proposal is included in the current proposal and will make sure there would be no major overlap between manuscripts from MP1823 and a manuscript from the current proposal.

MP3297 - Estimating absolute treatment effect of blood pressure lowering therapy for individual elderly patients; T.I. De Vries
In this proposal, which is aimed at the prediction of risk and treatment benefit in the elderly, a similar methodology is used in order to improve predictions in top of a basic model. The current proposal has a more methodological scope, assesses a broader range of additional variables and is aimed at the general population without cardiovascular disease. The research proposal is from the same research group as the current proposal.

MP3008 – Individualized estimation of cardiovascular disease-free survival in primary prevention to facilitate personalized treatment with lipid, antihypertensive, and antiplatelet therapy: A collaborative analysis of three community-based cohorts; N.E.M. Jaspers
Similar to the current proposal, this recently published proposal is aimed at the prediction of cardiovascular disease for the general population. An externally validated model was developed to estimate CVD-free life-expectancy. In the current proposal, we will not develop a new model but rather aim to improve the accuracy and flexibility of existing models. The project is from the same research group as the current proposal.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? ____ Yes __X__ No

12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

13. References:

1. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of

cardiovascular risk: A report of the American college of cardiology/American heart association task force on practice guidelines. *Circulation*. 2014;129(25 SUPPL. 1):49-73. doi:10.1161/01.cir.0000437741.48606.98

2. Piepoli MF, Hoes AW, Agewall S, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice. *Eur Heart J*. 2016;37(29):2315-2381. doi:10.1093/eurheartj/ehw106
3. Jaspers NE, Blaha M, Matsushita K, et al. Prediction of individualized liftime benefit from cholesterol lowering, blood pressure-lowering, antithrombotic therapy, and smoking cessation in apparently healthy people. *Eur Heart J*.
4. Conroy RM, Pyörälä K, Fitzgerald AP, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: The SCORE project. *Eur Heart J*. 2003;24(11):987-1003. doi:10.1016/S0195-668X(03)00114-3
5. Preiss D, Kristensen SL. The new pooled cohort equations risk calculator. *Can J Cardiol*. 2015;31(5):613-619. doi:10.1016/j.cjca.2015.02.001
6. Fretz A, Schneider ALC, McEvoy JW, et al. The Association of Socioeconomic Status With Subclinical Myocardial Damage, Incident Cardiovascular Events, and Mortality in the ARIC Study. *Am J Epidemiol*. 2016;183(5):452-461. doi:10.1093/aje/kwv253
7. Qureshi WT, Michos ED, Flueckiger P, et al. Impact of Replacing the Pooled Cohort Equation With Other Cardiovascular Disease Risk Scores on Atherosclerotic Cardiovascular Disease Risk Assessment (from the Multi-Ethnic Study of Atherosclerosis [MESA]). *Am J Cardiol*. 2016;118(5):691-696. doi:10.1016/j.amjcard.2016.06.015
8. Day N, Oakes S, Luben R, et al. EPIC-Norfolk: study design and characteristics of the cohort. *Br J Cancer*. 1999;80 Suppl 1:95-103. doi:10466767
9. Schmermund A, Möhlenkamp S, Stang A, et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: Rationale and design of the Heinz Nixdorf RECALL study. *Am Heart J*. 2002;144(2):212-218. doi:10.1067/mhj.2002.123579
10. Boeing H, Korfmann A, Bergmann MM. Recruitment procedures of EPIC-Germany. *Ann Nutr Metab*. 1999;43(4):205-215. doi:10.1159/000012787
11. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol*. 2015;44(3):827-836. doi:10.1093/ije/dyv098
12. Berkelmans GFN, Read SH, Gudbjörnsaottir S, et al. *Dealing with Missing Patient Characteristics When Using Cardiovascular Prediction Models in Clinical Practice*. Poster session presented at ESC Congress 2018 August 25-29. Munich, Germany; 2018.
13. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. In: *Statistics in Medicine*. Vol 23. ; 2004:1111-1130. doi:10.1002/sim.1668