Supplementary Material

References – ONLINE ONLY

1. World Health Organization. Women’s Health. http://www.who.int/mediacentre/factsheets/fs334/en/ (21 October 2016)

2. Mosca L, Benjamin EJ, Berra K, Bezanson JL, Dolor RJ, Lloyd-Jones DM, Newby LK, Piña IL, Roger VL, Shaw LJ, Zhao D, Beckie TM, Bushnell C, D’Armiento J, Kris-Etherton PM, Fang J, Ganiats TG, Gomes AS, Gracia CR, Haan CK, Jackson EA, Judelson DR, Kelepouris E, Lavie CJ, Moore A, Nussmeier NA, Ofili E, Oparil S, Ouyang P, Pinn VW, Sherif K, Smith S, Sopko G, Chandra-Strobos N, Urbina E., Vaccarino V, Wenger N. Effectiveness-based guidelines for the prevention of cardiovascular disease in women--2011 update: a guideline from the american heart association. *Circulation* 2011;**123**:1243–1262.

3. Shaw LJ, Bairey Merz CN, Pepine CJ, Reis SE, Bittner V, Kelsey SF, Olson M, Johnson BD, Mankad S, Sharaf BL, Rogers WJ, Wessel TR, Arant CB, Pohost GM, Lerman A, Quyyumi AA, Sopko G, WISE Investigators. Insights from the NHLBI-Sponsored Women’s Ischemia Syndrome Evaluation (WISE) Study: Part I: gender differences in traditional and novel risk factors, symptom evaluation, and gender-optimized diagnostic strategies. *J Am Coll Cardiol* 2006;**47**:S4–S20.

4. Goff DC, Lloyd-Jones DM, Bennett G, Coady S, D’Agostino RB, Gibbons R, Greenland P, Lackland DT, Levy D, O’Donnell CJ, Robinson JG, Schwartz JS, Shero ST, Smith SC, Sorlie P, Stone NJ, Wilson PWF, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 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. *J Am Coll Cardiol* 2014;**63**:2935–2959.

5. Greenland P, Alpert JS, Beller GA, Benjamin EJ, Budoff MJ, Fayad ZA, Foster E, Hlatky MA, Hodgson JM, Kushner FG, Lauer MS, Shaw LJ, Smith SC, Taylor AJ, Weintraub WS, Wenger NK, 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2010;**56**:e50-103.

6. Rich-Edwards JW, Fraser A, Lawlor DA, Catov JM. Pregnancy characteristics and women’s future cardiovascular health: an underused opportunity to improve women’s health? *Epidemiol Rev* 2014;**36**:57–70.

7. Sattar N. Do pregnancy complications and CVD share common antecedents? *Atheroscler Suppl* 2004;**5**:3–7.

8. Bellamy L, Casas J-P, Hingorani AD, Williams DJ. Pre-eclampsia and risk of cardiovascular disease and cancer in later life: systematic review and meta-analysis. *BMJ* 2007;**335**:974.

9. McDonald SD, Malinowski A, Zhou Q, Yusuf S, Devereaux PJ. Cardiovascular sequelae of preeclampsia/eclampsia: a systematic review and meta-analyses. *Am Heart J* 2008;**156**:918–930.

10. Tooher J, Thornton C, Makris A, Ogle R, Korda A, Hennessy A. All Hypertensive Disorders of Pregnancy Increase the Risk of Future Cardiovascular Disease. *Hypertension* 2017;HYPERTENSIONAHA.117.09246.

11. Robbins CL, Hutchings Y, Dietz PM, Kuklina EV, Callaghan WM. History of preterm birth and subsequent cardiovascular disease: a systematic review. *Am J Obstet Gynecol* 2014;**210**:285–297.

12. Wu P, Gulati M, Kwok CS, Wong CW, Narain A, O’Brien S, Chew‐Graham CA, Verma G, Kadam UT, Mamas MA. Preterm Delivery and Future Risk of Maternal Cardiovascular Disease: A Systematic Review and Meta‐Analysis. *Journal of the American Heart Association* 2018;**7**:e007809.

13. Morken N-H, Halland F, DeRoo L, Wilcox A, Skjærven R. Offspring birthweight by gestational age and parental cardiovascular mortality: a population-based cohort study. *BJOG: Int J Obstet Gy* 2017;n/a-n/a.

14. Lykke JA, Paidas MJ, Triche EW, Langhoff-Roos J. Fetal growth and later maternal death, cardiovascular disease and diabetes. *Acta Obstet Gynecol Scand* 2012;**91**:503–510.

15. Li C-Y, Chen H-F, Sung F-C, Chen C-C, Lu T-H, Yang C-H, Ko M-C. Offspring birth weight and parental cardiovascular mortality. *Int J Epidemiol* 2010;**39**:1082–1090.

16. Regitz-Zagrosek V, Roos-Hesselink JW, Bauersachs J, Blomström-Lundqvist C, Cífková R, De Bonis M, Iung B, Johnson MR, Kintscher U, Kranke P, Lang IM, Morais J, Pieper PG, Presbitero P, Price S, Rosano GMC, Seeland U, Simoncini T, Swan L, Warnes CA, ESC Scientific Document Group. 2018 ESC Guidelines for the management of cardiovascular diseases during pregnancy. *Eur Heart J* 2018;**39**:3165–3241.

17. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, Cooney M-T, Corrà U, Cosyns B, Deaton C, Graham I, Hall MS, Hobbs FDR, Løchen M-L, Löllgen H, Marques-Vidal P, Perk J, Prescott E, Redon J, Richter DJ, Sattar N, Smulders Y, Tiberi M, Worp HB van der, Dis I van, Verschuren WMM. 2016 European Guidelines on cardiovascular disease prevention in clinical practice. *Eur Heart J* 2016;**37**:2315–2381.

18. Timpka S, Fraser A, Schyman T, Stuart JJ, Åsvold BO, Mogren I, Franks PW, Rich-Edwards JW. The value of pregnancy complication history for 10-year cardiovascular disease risk prediction in middle-aged women. *Eur J Epidemiol* 2018;**33**:1003–1010.

19. Stuart JJ, Tanz LJ, Cook NR, Spiegelman D, Missmer SA, Rimm EB, Rexrode KM, Mukamal KJ, Rich-Edwards JW. Hypertensive Disorders of Pregnancy and 10-Year Cardiovascular Risk Prediction. *J Am Coll Cardiol* 2018;**72**:1252–1263.

20. Selmer R, Igland J, Ariansen I, Tverdal A, Njølstad I, Furu K, Tell GS, Klemsdal TO. NORRISK 2: A Norwegian risk model for acute cerebral stroke and myocardial infarction. *Eur J Prev Cardiol* 2017;**24**:773–782.

21. Nasjonal faglig retningslinje for forebygging av hjerte- og karsykdom. Helsedirektoratet.no. https://helsedirektoratet.no/retningslinjer/forebygging-av-hjerte-og-karsykdom (17 May 2018)

22. Krokstad S, Langhammer A, Hveem K, Holmen TL, Midthjell K, Stene TR, Bratberg G, Heggland J, Holmen J. Cohort Profile: The HUNT Study, Norway. *Int J Epidemiol* 2013;**42**:968–977.

23. Holmen J, Midthjell K, Bjartveit K, Hjort PF, Lund-Larsen PG. The Nord-Trøndelag health Survey 1984-1986. Purpose, background and methods. Participation, non-participation and frequency distributions. Verdal: Senter for samfunnsmedisinsk forskning, Statens Institutt for folkehelse(SIFF). Helsetjenesteforskning; 1990 p. 1–257. Report No.: 4.

24. Holmen J, Midthjell K, Krüger Ø, Langhammer A, Holmen TL, Bratberg GH, Vatten L, Lund-Larsen PG. The Nord-Trøndelag Health Study 1995-97 (HUNT 2): Objectives, contents, methods and participation. *Norsk Epidemiologi* 2003;**13**:19–32.

25. Irgens LM. The Medical Birth Registry of Norway. Epidemiological research and surveillance throughout 30 years. *Acta Obstet Gynecol Scand* 2000;**79**:435–439.

26. Thomsen LCV, Klungsøyr K, Roten LT, Tappert C, Araya E, Baerheim G, Tollaksen K, Fenstad MH, Macsali F, Austgulen R, Bjørge L. Validity of the diagnosis of pre-eclampsia in the Medical Birth Registry of Norway. *Acta Obstet Gynecol Scand* 2013;**92**:943–950.

27. American College of Obstetricians and Gynecologists, Task Force on Hypertension in Pregnancy. Hypertension in pregnancy. Report of the American College of Obstetricians and Gynecologists’ Task Force on Hypertension in Pregnancy. *Obstet Gynecol* 2013;**122**:1122–1131.

28. Skjaerven R, Gjessing HK, Bakketeig LS. Birthweight by gestational age in Norway. *Acta Obstet Gynecol Scand* 2000;**79**:440–449.

29. Moth FN, Sebastian TR, Horn J, Rich-Edwards J, Romundstad PR, Åsvold BO. Validity of a selection of pregnancy complications in the Medical Birth Registry of Norway. *Acta Obstet Gynecol Scand* 2016;**95**:519–527.

30. Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR, White HD. Third Universal Definition of Myocardial Infarction. *Circulation* 2012;**126**:2020–2035.

31. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association* 1999;**94**:496–509.

32. Lin DY, Wei LJ. The Robust Inference for the Cox Proportional Hazards Model. *Journal of the American Statistical Association* 1989;**84**:1074–1078.

33. Hlatky MA, Greenland P, Arnett DK, Ballantyne CM, Criqui MH, Elkind MSV, Go AS, Harrell FE, Hong Y, Howard BV, Howard VJ, Hsue PY, Kramer CM, McConnell JP, Normand S-LT, O’Donnell CJ, Smith SC, Wilson PWF, American Heart Association Expert Panel on Subclinical Atherosclerotic Diseases and Emerging Risk Factors and the Stroke Council. Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American Heart Association. *Circulation* 2009;**119**:2408–2416.

34. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. *Statist Med* 2015;**34**:1659–1680.

35. Pencina MJ, D’Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med* 2004;**23**:2109–2123.

36. Pencina MJ, D’Agostino RB, D’Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;**27**:157–172; discussion 207-212.

37. Pencina MJ, D’Agostino RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 2011;**30**:11–21.

38. SAS Macros. http://ncook.bwh.harvard.edu/sas-macros.html (19 May 2018)

39. Harrell FE. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer; 2010.

40. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the Odds Ratio in Gauging the Performance of a Diagnostic, Prognostic, or Screening Marker. *Am J Epidemiol* 2004;**159**:882–890.

41. Rich-Edwards JW, Klungsoyr K, Wilcox AJ, Skjaerven R. Duration of pregnancy, even at term, predicts long-term risk of coronary heart disease and stroke mortality in women: a population-based study. *Am J Obstet Gynecol* 2015;**213**:518.e1-8.

42. Skjaerven R, Wilcox AJ, Klungsøyr K, Irgens LM, Vikse BE, Vatten LJ, Lie RT. Cardiovascular mortality after pre-eclampsia in one child mothers: prospective, population based cohort study. *BMJ* 2012;**345**:e7677.

43. Riise HKR, Sulo G, Tell GS, Igland J, Nygård O, Vollset SE, Iversen A-C, Austgulen R, Daltveit AK. Incident Coronary Heart Disease After Preeclampsia: Role of Reduced Fetal Growth, Preterm Delivery, and Parity. *J Am Heart Assoc* 2017;**6**.

44. Riise HKR, Sulo G, Tell GS, Igland J, Nygård O, Iversen A-C, Daltveit AK. Association Between Gestational Hypertension and Risk of Cardiovascular Disease Among 617 589 Norwegian Women. *J Am Heart Assoc* 2018;**7**.

45. Cirillo PM, Cohn BA. Pregnancy Complications and Cardiovascular Disease Death: Fifty-Year Follow-Up of the Child Health and Development Studies Pregnancy Cohort. *Circulation* 2015;CIRCULATIONAHA.113.003901.

46. Smith GD, Sterne J, Tynelius P, Lawlor DA, Rasmussen F. Birth weight of offspring and subsequent cardiovascular mortality of the parents. *Epidemiology* 2005;**16**:563–569.

47. Stuart J, Tanz L, Rimm E, Spiegleman D, Missmer S, Mukamal K, Rich-Edwards J. Hypertensive disorders in first pregnancy and maternal cardiovascular disease: mediation by postpartum cardiovascular risk factors. Seattle, Washington; 2017.

48. Conroy RM, Pyörälä K, Fitzgerald AP, Sans S, Menotti A, De Backer G, De Bacquer D, Ducimetière P, Jousilahti P, Keil U, Njølstad I, Oganov RG, Thomsen T, Tunstall-Pedoe H, Tverdal A, Wedel H, Whincup P, Wilhelmsen L, Graham IM. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003;**24**:987–1003.

49. D’Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, Kannel WB. General Cardiovascular Risk Profile for Use in Primary Care: The Framingham Heart Study. *Circulation* 2008;**117**:743–753.

50. Lund I, Lund KE. Lifetime smoking habits among Norwegian men and women born between 1890 and 1994: a cohort analysis using cross-sectional data. *BMJ Open* 2014;**4**:e005539.

51. Decline in fertility. ssb.no. http://www.ssb.no/en/befolkning/statistikker/fodte/aar/2017-03-09 (15 February 2018)

52. Parikh NI, Jeppson RP, Berger JS, Eaton CB, Kroenke CH, LeBlanc ES, Lewis CE, Loucks EB, Parker DR, Rillamas-Sun E, Ryckman KK, Waring ME, Schenken RS, Johnson KC, Edstedt-Bonamy A-K, Allison MA, Howard BV. Reproductive Risk Factors and Coronary Heart Disease in the Women’s Health Initiative Observational Study. *Circulation* 2016;**133**:2149–2158.

53. Lv H, Wu H, Yin J, Qian J, Ge J. Parity and Cardiovascular Disease Mortality: a Dose-Response Meta-Analysis of Cohort Studies. *Sci Rep* 2015;**5**.

54. Stene LC, Eidem I, Vangen S, Joner G, Irgens LM, Moe N. The validity of the diabetes mellitus diagnosis in the Medical Birth Registry of Norway. *Nor Epidemiol* 2007;**17**.

55. Fraser A, Nelson SM, Macdonald-Wallis C, Cherry L, Butler E, Sattar N, Lawlor DA. Associations of Pregnancy Complications With Calculated Cardiovascular Disease Risk and Cardiovascular Risk Factors in Middle AgeClinical Perspective. *Circulation* 2012;**125**:1367–1380.

56. Kravdal Ø. Not so Low Fertility in Norway—A Result of Affluence, Liberal Values, Gender-Equality Ideals, and the Welfare State. *Low Fertility, Institutions, and their Policies* Springer, Cham; 2016. p. 13–47.

**Supplemental Table 1.** Description of established CVD risk factor measurement methods by HUNT exam

|  |  |  |
| --- | --- | --- |
| **Risk Factor** | **HUNT2** | **HUNT3** |
| Blood pressure | Staff measured blood pressure three times at 1-minute intervals after the person had come to rest using an automatic oscillometric method (Dinamap, Critikon, Florida), with cuff size adjusted to arm circumference. We defined systolic blood pressure as the mean of the second and third measurements. In HUNT3, 1,914 women were missing a third measurement due to staff shortages. Since the first measurement tends to be too low using Dinamap1, the second measurement was used in these cases. | |
| Cholesterol | Technicians at Levanger Hospital’s Central Laboratory assessed total cholesterol and high-density lipoprotein cholesterol (HDL-C), using the enzymatic colorimetric cholesterol esterase method with reagents from Boehringer Mannheim (Mannheim, Germany) using a Hitachi 911 Autoanalyzer. | Technicians at Levanger Hospital’s Central Laboratory measured total cholesterol using cholesterol esterase methodology and HDL-C using accelerator selective detergent methodology, all using reagents from Abbott (Abbott Ireland, Longford, Ireland; and Abbott Laboratories, Abbott Park, Illinois) using an Architect cSystems ci8200 |

1 Lund-Larsen PG. Blodtrykk målt med kvikksølvmanometer og med Dinamap under feltforhold - en sammenligning [Blood pressure measured with a sphygmomanometer and with Dinamap under field conditions – a comparison] Nor J Epidemiol. 1997;7:235–241.

**Supplemental Text 1.** Description of the cardiovascular event validation

From the electronic patient administrative system at the two primary hospitals in Nord-Trøndelag county, Levanger Hospital and Namsos Hospital (Nord-Trøndelag Hospital Trust), we obtained information on all study participants registered with at least one of the following cardiovascular diagnoses between September 1, 1987 (the first date of electronic recording) and April 24, 2015: ICD-9,402, 404, 410-414, 425, 427.5, 428, 430-438, 440-448; ICD-10, G45, I11, I13, I20-I25, I42, I46, I50, I60-I67, I71, I72, I74. For each identified patient, one of two experienced cardiologists (B.K., H.D.), who were unaware of the pregnancy history of the participants, examined the medical records to determine the first validated occurrence of a range of cardiovascular events, including myocardial infarction, cerebral infarction and intracranial hemorrhage. Myocardial infarction was diagnosed using ESC/ACCF/AHA/WHF criteria1. Cerebral infarction was diagnosed based on typical symptoms and signs combined with radiological evidence from CT or MRI scans. In some cases, the diagnosis of cerebral infarction was based on typical symptoms and signs in absence of radiological evidence, but only if CT or MRI scan had been performed, and no alternative explanation for the clinical presentation was found. Possible cerebrovascular events accompanying intracerebral malignancy, infection or inflammation were not included. Intracranial hemorrhage was classified based on radiological evidence as subarachnoid, intracerebral, subdural or epidural, and hemorrhage secondary to trauma or intracerebral malignancy was excluded. As stroke events in the present study, we included cerebral infarctions as well as intracerebral and subarachnoid hemorrhages.

1. Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR, White HD. Third Universal Definition of Myocardial Infarction. *Circulation* 2012;**126**:2020–2035

**Supplemental Table 2.** ICD diagnosis codes for fatal cardiovascular endpoints

|  |  |  |  |
| --- | --- | --- | --- |
|  | **ICD-8 codes** | **ICD-9 codes** | **ICD-10 codes** |
| Fatal CHD | 410-414 | 410-414 | I20-I25 |
| Fatal stroke | 430-433 | 430, 431, 433-434 | I60, I61, I63 (excluding I63.6), I64 |

CHD=coronary heart disease

**Supplemental Text 2.** Additional description of statistical methods

a) Cluster-robust variance estimates

In the presence of clustered data, as in this manuscript where observations were clustered within women, naively estimated confidence intervals can be falsely narrow. To adjust for the correlation within women, we used cluster-robust estimates of the variance-covariance matrix. Robust variance estimation can be referred to by many different names, including sandwich or Huber-White1,2 estimation. In brief, the robust variance estimator can be thought of as a sandwich where the “bread” is the conventional estimator of variance and, in the context of survival analysis, the “meat” is typically an empirical variance estimator calculated using score residuals. The cluster-robust variance estimator is an extension of the robust variance estimator where the score residuals used to compute the “meat” of the sandwich are aggregated over identical ID values (in this case, aggregated across observations for the same woman). We implemented this method in SAS using the COVSANDWICH(AGGREGATE) option in PROC PHREG3.

b) Multiple imputation

Multiple imputation methods overview

To impute pregnancy complications for pregnancies either 1) missing from the registry or 2) with incomplete information in the registry, we used multiple imputation by fully conditional specification (FCS)4. FCS is also sometimes referred to as multiple imputation by chained equations (MICE or ICE) or sequential regression multiple imputation. In short, the FCS algorithm uses the following steps5:

1. Fill in random values for missing variables by sampling with replacement from the observed values
2. Starting with one variable (X1 for example), build a prediction model for the observed values of the variable based on all of the other variables
3. Impute missing values of X1 based on this prediction model, filling in the randomly chosen values with these imputed values
4. Repeat for all variables until all of the random values are filled in with imputed values

We used PROC MI in SAS to run the FCS algorithm. We chose to impute 20 times which means these steps were repeated 20 times resulting in 20 datasets with different imputed values. At 20 imputations, the efficiency in estimating beta coefficients is similar to the efficiency we would have seen had we used an infinite number of iterations5. In addition, we found little change in estimates when moving from 10 imputations to 20, suggesting that the increased computing time to produce additional imputations would have little to no effect on results.

In all Fine and Gray competing risk models reported in this paper that included pregnancy complications, we ran the model on each of the 20 datasets. Beta coefficients were calculated as the average across the 20 datasets and standard errors were corrected using Rubin’s rules6. We used PROC MIANALYZE in SAS to calculate corrected standard errors, confidence intervals, and p-values, including confidence intervals reported in Table 2 and the Wald test for model fit reported in the results section. All other prediction metrics (for calibration, discrimination, and reclassification) are based only on point estimates and not prediction model standard errors so they did not require any corrections for multiple imputation.

Imputation outcome variables:

We imputed pregnancy complications separately by birth order given the known differences in prevalence of pregnancy complications, particularly preeclampsia, by birth order. We imputed complications only for the first three pregnancies because information about later pregnancies was more complete. In addition, there were few observed pregnancy complications in fourth or later pregnancy (for example, only 56 observed preterm deliveries in fourth or later pregnancies) and imputation is less accurate with fewer observations. In total, we imputed 12 variables (i.e. 4 individual pregnancy complications across a woman’s first three pregnancies). After imputing the pregnancy complication status for each pregnancy separately, we aggregated across all pregnancies (observed and imputed) to indicate whether the women had any history of the complication of interest.

Imputation predictor variables

We included the following predictors in our imputation models: 1) maternal birth year, 2) whether the mother had a CVD event during follow-up, 3) whether the mother died from a non-CVD cause during follow-up, 4) all covariates included in the prediction model (age, age-squared, systolic blood pressure, serum total cholesterol, daily smoking, interaction terms for systolic blood pressure and age and daily smoking and age, antihypertensive use, low HDL-cholesterol, and family history of premature MI), and 5) pregnancy complications in other births. We assumed that the data were missing at random (MAR)6 after accounting for these variables. In other words, that the probability we had complete information on pregnancy complications did not depend on whether the woman had a pregnancy complication, after controlling for these variables. We think this is a reasonable assumption, particularly after controlling for birth year since the major reason for missing data was that births occurred prior to the start of the registry in 1967 which was more likely to happen for older women.

Imputation prediction method

We used the discriminant function method to classify women as having had a pregnancy complication or not (for example, preeclampsia in first pregnancy yes or no) for each imputation. We chose the discriminant function a priori since it has less issues with convergence than logistic regression; however, in sensitivity analyses, we found similar results using logistic regression.

Comparison of women with and without missing pregnancy information

Women with incomplete pregnancy complication history (i.e. women with at least one pregnancy that was either missing from the registry or who had incomplete information in the registry) were more likely to be born earlier and, thus, older at the time of the HUNT examinations. After age-standardizing, women with incomplete history had somewhat higher levels of blood pressure, serum total cholesterol, and smoking. These variables were all included in the imputation prediction models.

|  |  |  |
| --- | --- | --- |
|  | **Complete pregnancy complication history (women=11,802 observationsa=16,889)** | **Incomplete pregnancy complication history (women=6,429 observationsa=9,655)** |
| Maternal birth year, median (IQR) | 1952 (1948-1956) | 1943 (1938-1948) |
| Age at HUNT exam in years, median (IQR) | 49 (45-55) | 57 (51-64) |
| *Age-standardized b risk factors from NORRISK 2 Model* |  |  |
| Systolic blood pressure in mmHg, median (IQR) | 127 (116-140) | 129 (117-144) |
| Serum total cholesterol in mmol/L, median (IQR) | 5.7 (5.1-6.5) | 5.9 (5.1-6.7) |
| Current daily smoking | 30% | 32% |
| Current anti-hypertensive use | 13% | 13% |
| Low HDL-Cc | 37% | 37% |
| Family history of premature MId | 17% | 17% |

a Women who participated in only HUNT2 (n=3,603) or HUNT3 (n=6,315) contributed one observation while women who participated in both HUNT2 and HUNT3 (n=8,313) contributed two.

b Risk factors standardized to the age distribution of the study population

c Low HDL-C: <1.3 mmol/L

d First degree family member suffered MI before the age of 60 years

IQR = Interquartile range; HDL-C = high density lipoprotein cholesterol; MI = myocardial infarction

c) Prediction performance measures

Overview

This table summarizes the prediction performance measures used in this analysis.

|  |  |  |
| --- | --- | --- |
| **Measure type** | **Measure assesses…** | **Specific measure used** |
| Calibration | Equivalence between observed CVD risk and model-predicted CVD risk | Greenwood-Nam-D’Agostino test |
| Discrimination | Model ability to distinguish between CVD cases and non-cases | C-index and C-index difference |
| Reclassification | Improvement in CVD risk classification of the new model over the established model | Net reclassification improvement (overall and for events and non-events separately)  Integrated discrimination improvement |

Greenwood-D’Agostino-Nam test

The Greenwood-Nam-D’Agostino (GND) test7 is an extension of the Hosmer-Lemeshow test to situations with censored survival data. It is used to compare observed and expected (i.e. predicted) categories of CVD risk, typically categorized into deciles (G=10 in the formula below). The formula for the test statistic is:

Where:

is the Kaplan-Meier failure probability in the gth decile at time t

is the mean predicted probability of failure for subjects in the gth decile at time t

is the Greenwood estimator of the variance of the probability of failure, where and are the number of failures and number at risk, respectively, at time .

C-index

The concordance-statistic (C-statistic) is a frequently used measure of discrimination in logistic regression which reflects the area under the ROC curve (AUC). This measure is not appropriate for survival analysis, but the C-index8 is an extension of this same concept. We can interpret the C-index as the probability of concordance (agreement) between the predicted and observed survival times of any two subjects. For example, if subject A had a CVD event before subject B and subject A had a shorter predicted survival time, this would be counted as a concordant pair. If subject B had a shorter predicted survival time, this would be a discordant pair. Only pairs where at least one subject had an event are considered in the formula. The C-index is the percentage of all possible pairs who were concordant. Another way to express this is:

Where

and denote observed survival time for subjects *i* and *j*

and denote predicted survival time for subjects *i* and *j*

Across all pairs of subjects *i* and *j* where

Net reclassification improvement

The net reclassification improvement (NRI) quantifies the improvement of a new model over an established model. In order to calculate the NRI, predicted probabilities must be divided into pre-established categories; for example, in this paper, we divided the 10-year risk of CVD into low (<5%), intermediate (5-<10%), and high (≥10%) risk groups. Improvements (i.e. correct reclassification) occur when subjects that had an event move “up” in their risk category under the new model and subjects that didn’t have an event move “down” in risk category. The basic formula is:

To extend the NRI to survival analysis9, this formula can be re-written using Bayes theorem to:

, , and can all be estimated using the Kaplan-Meier approach while and are straightforward to estimate in the dataset.

We also calculated the NRI separately by event status, which are just components of the overall NRI formula broken out:

Integrated discrimination improvement

The integrated discrimination improvement (IDI) is a continuous measure of reclassification.10 Improvements occur when the mean predicted probabilities (here, 10-year risk of CVD) increase under the new model for subjects that had an event and decrease for subjects that did not have an event. The IDI can be directly estimated from survival models.

Where, for examples, is the predicted probability of the event under the new model among subjects who had an event.

References

1. Huber PJ. The behavior of maximum likelihood estimates under nonstandard conditions. The Regents of the University of California; 1967.
2. White H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica* 1980;**48**:817–838.
3. Gharibvand L, Liu L. Analysis of Survival Data with Clustered Events. SAS Global Forum 2009. Paper 237-2009. https://support.sas.com/resources/papers/proceedings09/237-2009.pdf (10 September 2018).
4. Buuren S van. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res* 2007;**16**:219–242.
5. Carpenter J, Kenward M. Multiple Imputation and its Application. 1st ed. United Kingdom: John Wiley & Sons; 2013.
6. Rubin D. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons; 1987.
7. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. *Statist Med* 2015;**34**:1659–1680.
8. Pencina MJ, D’Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med* 2004;**23**:2109–2123.
9. Pencina MJ, D’Agostino RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 2011;**30**:11–21.
10. Pencina MJ, D’Agostino RB, D’Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;**27**:157–172; discussion 207-212.

**Supplemental Table 3.** Comparison of NORRISK 2 model as reported in Selmer et al.1 to version used in this paper

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Covariatesa** | **NORRISK 2 Model as reported in Selmer et al. for women (N=35,267, endpoints=2,459)** | | | **Version of NORRISK 2 used in this paper  (N=18,231, endpoints=965)** | | |
| **Beta** | **95% CI** | | **Beta** | **95% CI** | |
| Age (per 1 year) | 0.130 | 0.111 | 0.150 | 0.090 | 0.061 | 0.120 |
| Age squared(per 1 year) | -0.001 | -0.001 | -0.0002 | -0.000004 | -0.0007 | 0.0007 |
| Systolic blood pressure (per 10 mmHg) | 0.252 | 0.201 | 0.304 | 0.176 | 0.104 | 0.247 |
| Serum total cholesterol (per 1 mmol/L) | 0.072 | 0.033 | 0.111 | 0.113 | 0.058 | 0.169 |
| Daily smoking (yes/no) | 1.268 | 1.053 | 1.483 | 1.168 | 0.862 | 1.474 |
| Systolic blood pressure (per 10 mmHg) x age | -0.005 | -0.007 | -0.003 | -0.002 | -0.005 | 0.002 |
| Daily smoking x age | -0.025 | -0.033 | -0.016 | -0.019 | -0.034 | -0.004 |
| Anti-hypertensives (yes/no) | 0.192 | 0.091 | 0.293 | 0.347 | 0.190 | 0.504 |
| Low HDL-cholesterolb (yes/no) | 0.324 | 0.239 | 0.408 | 0.499 | 0.372 | 0.627 |
| Family history of premature MIc: One family member | 0.254 | 0.142 | 0.366 |  |  |  |
| Family history of premature MIc: At least two family members | 0.549 | 0.321 | 0.777 |
| Family history of premature MIc (yes/no) |  |  |  | 0.422 | 0.272 | 0.573 |
|  |  |  |  |  |  |  |
| Λ0(5)d | 0.0010 |  |  | 0.0010 |  |  |
| Λ0(8)d | 0.0017 |  |  | 0.0018 |  |  |
| Λ0(10)d | 0.0023 |  |  | 0.0027 |  |  |

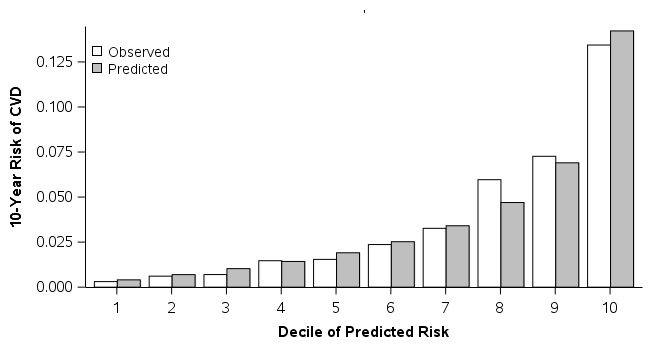
a Age is centered at 40. Systolic blood pressure is centered at 120. Serum total cholesterol is centered at 4

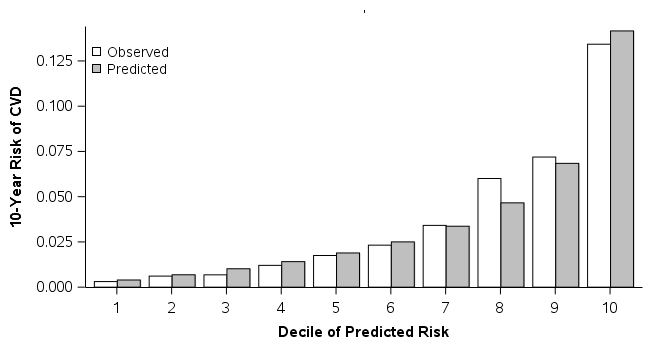
b Low HDL-cholesterol: < 1.3 mmol/L

c First degree family member suffered MI before the age of 60 years

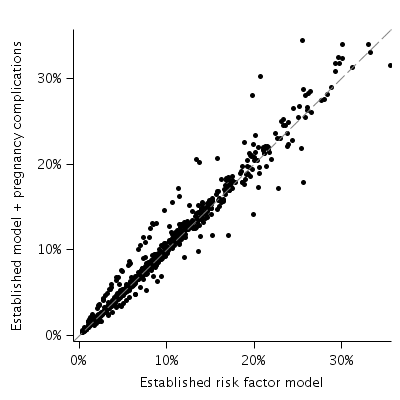
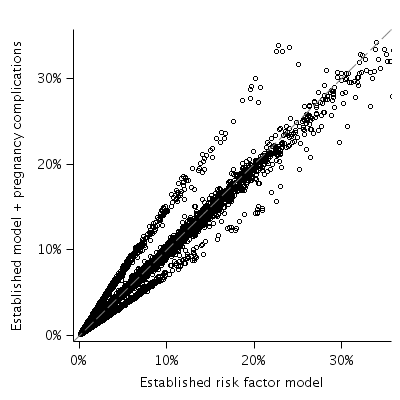
d Λ0(5), Λ0(8) and Λ0(10) are baseline cumulative sub distribution hazards at 5, 8 and 10 years for age 40, systolic blood pressure 120 mmHg, serum total cholesterol 4 mmol/L, no smoking, HDL not low, no family history and not current user of antihypertensives.

1 Selmer R, Igland J, Ariansen I, Tverdal A, Njølstad I, Furu K, Tell GS, Klemsdal TO. NORRISK 2: A Norwegian risk model for acute cerebral stroke and myocardial infarction. Eur J Prev Cardiol 2017;24:773–782.

**A** 

**B** ****

**Supplemental Figure 1.** Calibration plots for A) the established risk factor model and B) the established risk factor model with addition of pregnancy complications

 **A** **B**

Improved performance

Improved performance

**Supplemental Figure 2.** Reclassification plots of predicted probability before and after inclusion of pregnancy complications among A) observations where no CVD event occurred (n= 25,579), here observations below the dashed line indicate improved model performance and B) observations where a CVD event occurred (n=965), here observations above the dashed line indicate improved model performance. The cluster of women on or near the dashed line represent women without pregnancy complications, whose predicted probability remained similar in both models. The cluster of women well above the line represent women with preeclampsia, whose predicted probability increased and women well below the line represent women with gestational hypertension, whose predicted probability decreased after adding pregnancy complications to the model.

**Supplemental Table 4.** Reclassification of CVD risk category after including preeclampsia history a

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Established CVD risk factor model + preeclampsia history | | | | | | | |
|  | Low risk  (0-<5%) | | Intermediate risk  (5-<10%) | | High risk  (≥10%) | | Total | |
| Established CVD risk factor model | N | % | N | % | N | % | N | % |
| *Observations with incident CVD event* |  |  |  |  |  |  |  |  |
| 0-<5% | 388 | 97% | 17 | 6% | 0 | 0% | 405 | 42% |
| 5-<10% | 10 | 3% | 263 | 92% | 15 | 5% | 288 | 30% |
| ≥10% | 0 | 0% | 6 | 2% | 266 | 95% | 272 | 28% |
| **Total** | 398 |  | 286 |  | 281 |  | 965 |  |
|  |  |  |  |  |  |  |  |  |
| *Observations with no incident CVD event* |  |  |  |  |  |  |  |  |
| 0-<5% | 19,857 | 99% | 175 | 5% | 0 | 0% | 20,032 | 78% |
| 5-<10% | 223 | 1% | 3,520 | 93% | 85 | 5% | 3,828 | 15% |
| ≥10% | 0 | 0% | 103 | 3% | 1,616 | 95% | 1,719 | 7% |
| **Total** | 20,080 |  | 3,798 |  | 1,701 |  | 25,579 |  |

a Shaded areas represent improvements in reclassification after the addition of preeclampsia history

Net reclassification improvement (NRI) = 0.02 (95% CI: -0.005, 0.04), p-value=0.11

NRI for events = 0.01 (95% CI: -0.002, 0.04), p-value=0.17

NRI for non-events = 0.002 (95% CI: 0.0006, 0.004), p-value=0.004

Integrated discrimination improvement (IDI) = 0.0001 (95% CI: -0.0005, 0.0008), p-value=0.83

**Supplemental Table 5.** Results from sensitivity analyses

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Main findings | | | Sensitivity analyses | | | | | |
|  | Established risk factor model + pregnancy complication history a | | | Established risk factor model + pregnancy complication history including recurrent pregnancy complications a,b | | | Established risk factor model + pregnancy complication history including interactions between pregnancy complications a,c | | |
|  | Estimate | 95% CI | p-value | Estimate | 95% CI | p-value | Estimate | 95% CI | p-value |
| *Model fit* |  |  |  |  |  |  |  |  |  |
| Wald test |  |  | 0.04 |  |  | 0.06 |  |  | 0.11 |
| *Calibration* |  |  |  |  |  |  |  |  |  |
| GND test d |  |  | 0.26 |  |  | 0.27 |  |  | 0.16 |
| *Discrimination* |  |  |  |  |  |  |  |  |  |
| C-index | 0.793 | (0.776, 0.807) | <0.001 | 0.794 | (0.777, 0.808) | <0.001 | 0.794 | (0.778, 0.809) | <0.001 |
| C-index difference e | 0.004 | (0.002, 0.006) | <0.001 | 0.004 | (0.002, 0.007) | 0.001 | 0.005 | (0.003, 0.007) | <0.001 |
| *Reclassification* e |  |  |  |  |  |  |  |  |  |
| NRI | 0.02 | (0.0002, 0.05) | 0.04 | 0.02 | (0.0002, 0.04) | 0.04 | 0.03 | (0.01, 0.05) | 0.001 |
| NRI events | 0.02 | (-0.002, 0.04) | 0.08 | 0.02 | (-0.004, 0.04) | 0.10 | 0.03 | (0.009, 0.04) | 0.01 |
| NRI non-events | 0.004 | (0.002, 0.006) | <0.001 | 0.004 | (0.002, 0.006) | <0.001 | 0.004 | (0.002, 0.006) | <0.001 |
| IDI | -0.0002 | (-0.001, 0.0007) | 0.65 | <0.0001 | (-0.0001, 0.0001) | 0.99 | -0.005 | (-0.001, 0.0004) | 0.29 |

a Pregnancy complications include preeclampsia, gestational hypertension, preterm delivery, and small for gestational age delivery

b Models include an additional term for each pregnancy complication indicating a history of two or more pregnancies with complication

c Models include interaction terms between each pair of pregnancy complication history variables

d Greenwood-Nam-D’Agostino test of calibration for censored survival data. Higher p-values indicate better model calibration

e Compared to the established risk factor model

**Supplemental Table 6.** Comparing established risk factor model with and without pregnancy complication historya after adjusting for optimismb

|  |  |  |
| --- | --- | --- |
| Metric | Original estimate | Estimate adjusted for optimism |
| *Discrimination* |  |  |
| C-index difference | 0.004 | 0.003 |
| *Reclassification* |  |  |
| NRI | 0.02 | 0.02 |
| NRI events | 0.02 | 0.02 |
| NRI non-events | 0.004 | 0.004 |
| IDI | -0.0002 | -0.0002 |

a Pregnancy complications include preeclampsia, gestational hypertension, preterm delivery, and small for gestational age delivery

b Optimism is a measure of the degree of model overfitting