Developing and validating a cardiovascular risk score for patients in the community with prior cardiovascular disease

Poppe: CV risk score in prior CVD

[1,2]Katrina K Poppe

[2,3]Rob N Doughty

[1]Sue Wells

[1]Dudley Gentles

[4]Harry Hemingway

[1]Rod Jackson

[1,5]Andrew J Kerr

[1]School of Population Health, University of Auckland, New Zealand

[2]Department of Medicine, University of Auckland, New Zealand

[3]Greenlane Cardiovascular Service, Auckland City Hospital, New Zealand

[4]Farr Institute of Health Informatics Research and Institute of Health Informatics, University College London, United Kingdom

[5]Middlemore Hospital, Counties Manukau District Health Board, Auckland, New Zealand

Corresponding author

Katrina Poppe PhD COPSCT

Epidemiology and Biostatistics, School of Population Health, University of Auckland

Private Bag 92019, Auckland 1142, New Zealand

k.poppe@auckland.ac.nz

ph +64-9-9239803;   Total word count: 3205

1

ABSTRACT

Objective: Patients with atherosclerotic cardiovascular disease (CVD) vary significantly in their risk of future CVD events yet few clinical scores are available to aid assessment of risk. We sought to develop a score for use in primary care that estimates short-term CVD risk in these patients.

Methods: Adults aged <80 years with prior CVD were identified from a New Zealand primary care cohort study (PREDICT), and linked to national mortality, hospitalisation, and dispensing databases. A Cox model with an outcome of myocardial infarction, stroke or CVD death within 2 years was developed. External validation was performed in a cohort from the United Kingdom (UK).

Results: 24,927 patients, 63% men, 63% European, median age 65 years (IQR 58-72 years), experienced 1,480 CVD events within 2 years after a CVD risk assessment. A risk score including ethnicity, co-morbidities, BMI, creatinine, and treatment, in addition to established risk factors used in primary prevention, predicted a median 2-year CVD risk of 5.0% (IQR 3.5-8.3%). A plot of actual against predicted event rates showed very good calibration throughout the risk range. The score performed well in the UK cohort but overestimated risk for those at highest risk, who were predominantly patients defined as having heart failure.

Conclusions: The PREDICT-CVD secondary prevention score uses routine measurements from clinical practice that enable it to be implemented in a primary care setting. The score will facilitate risk communication between primary care practitioners and patients with prior CVD, particularly as a resource to show the benefit of risk factor modification.

Keywords: cardiovascular disease; risk score; secondary prevention; electronic health record

KEY QUESTIONS

What is already known about this subject?

Patients with established cardiovascular disease (CVD) are largely managed in primary care and, in contrast to patients without CVD, there is minimal guidance to the clinician or patient about stratifying risk and thus the benefit of risk factor modification.

What does this study add?

The PREDICT-CVD secondary prevention score shows that not all patients with CVD are at uniformly high risk of experiencing a subsequent CVD event. In addition to established risk factors for primary prevention, the score includes ethnicity, heart failure, the severity of the prior CVD event (MI or stroke), BMI, creatinine, and lipid and blood pressure lowering treatment. The full equation and an example of how it is calculated is included in this manuscript.

How might this impact on clinical practice?

The new score may help in communication with patients with existing CVD, to recognise patient-specific risks of future events and how they may be reduced through therapeutic and behavioural strategies. In combination with primary prevention scores, the new score will enable quantitative stratification of risk across the continuum of primary and secondary CVD prevention in primary care.

INTRODUCTION

Cardiovascular disease (CVD) risk prediction scores are available for use at the time of an acute CVD event.[1, 2] However patients are living longer with established CVD and few scores are available that estimate an individual patient's CVD risk at a time distant to an acute event.[3] International clinical guidelines[4-6] recommend that patients without prior CVD have their CVD risk estimated by risk scores[6-10] and that the intensity of risk factor management should be informed by the predicted absolute CVD risk.[11] In contrast, patients with established CVD are usually considered at "clinically high risk" without any further risk stratification.[4-6] Patients find this message demotivating[12] and primary care practitioners have little to no resource to show the benefit of risk factor modification. Nearly half of CVD events occur in patients with prior CVD[13] and there is evidence of important risk heterogeneity within this group.[13,3] A score to stratify risk in those with known CVD in the primary care setting would complement the established approach of CVD risk assessment for those without CVD, and contribute to a suite of risk scores that are relevant to an individual at any point in their continuum of CVD risk.

The New Zealand (NZ) PREDICT CVD Cohort Study was initiated in 2002 to develop CVD risk scores for a range of populations. PREDICT decision support software includes currently recommended CVD risk scores and is integrated with the electronic patient management systems of over one-third of NZ primary care practitioners. Relevant data from the electronic health record are automatically extracted and additional data are entered during the patient's clinical assessment. Thus the study infrastructure not only collects routine measurements taken directly from clinical practice, but provides a mechanism by which new risk scores can be developed and implemented back into clinical practice.

We sought to develop and validate a CVD risk score for use in primary care that estimates 2-year fatal and non-fatal CVD risk in patients with established CVD. Score development was supported by

anonymised linkage of the PREDICT data to national and regional administrative health datasets. Collaboration with CALIBER investigators enabled a validation cohort from the United Kingdom (UK) to be developed, and for factors affecting the transferability of the score into clinical practice to be identified.

METHODS

Patient cohort

The PREDICT web-based decision support programme has been described previously.[14] When PREDICT is used by a practitioner to estimate CVD risk for a patient, an electronic risk profile is stored both in the patient record and anonymously in a central database. With the permission of health providers, this profile is linked to an encrypted National Health Index number (eNHI) and made available to researchers at the University of Auckland. At the time of these analyses 272,682 people with and without CVD were enrolled in PREDICT.

A derivation cohort was created that included NZ residents aged 18-80 years who received a PREDICT assessment in primary care between January 2006 and April 2012, and where the clinician (GP) had recorded prior CVD as at least one of: angina, myocardial infarction (MI), percutaneous coronary intervention (PCI) or coronary artery bypass grafts (CABG), transient ischaemic attack (TIA), ischaemic stroke, or peripheral vascular disease (PVD). The first PREDICT assessment at which a history of CVD was recorded is the index assessment.

Data considerations

Clinical data from this subset of PREDICT patients were anonymously linked to national hospital discharge and mortality data, pharmaceutical dispensing, and regional laboratory tests via the unique eNHI. Pharmaceutical data were limited to cardiovascular medications dispensed within six months prior to the index assessment. Laboratory data were limited to the most recent values of blood cholesterol fractions recorded up to one year prior or two weeks after the index assessment, or up to five years prior for blood creatinine. A five year timeframe for creatinine was considered realistic among clinically stable out-patients, and preliminary analyses showed 90% of values were measured within 2 years. Thus the study dataset includes clinical and demographic information recorded at the

time of the index assessment, the most recent dispensing and laboratory information, and subsequent hospitalisations and deaths. The end of follow-up was 30 April 2012.

A clinician-defined diagnosis of heart failure (HF) was not available in the PREDICT database. Thus the presence of HF was defined as an ICD-coded hospitalisation for HF at any prior date and/or dispensing of a loop diuretic during the 6 months prior to the index risk assessment. This definition therefore represents a spectrum of HF, including patients who have not been hospitalised. In NZ it is unusual for loop diuretics to be used for indications other than HF. Forty percent of patients with HF as defined met both criteria, 30% were based on prior hospitalisation only, and 31% by loop diuretic use only. Sensitivity analyses showed that the 2-year event rate among patients receiving a loop diuretic was 14%, compared to 18% among patients with a prior HF hospitalisation.

Blood pressure (BP) lowering medication is defined as at least one dispensing of a: beta-blocker, angiotensin-converting enzyme inhibitor (ACEi), angiotensin receptor blocker, calcium channel blocker, or other anti-hypertensive agent in the 6 months prior to the index risk assessment. Anticoagulation is defined as dispensing of warfarin only, as novel oral anticoagulants were not used in NZ during the period of data collection, and low molecular weight heparin is not part of standard general practice in NZ.

The outcome of interest was time to first new CVD event within 2 years of the index risk assessment. A CVD event was defined as MI, ischaemic or haemorrhagic stroke, or CVD death, defined from ICD-10-AM coded national hospital and mortality datasets. Patients having a non-CVD death were censored at the date of death.

Statistical approach

Potential predictors were selected *a priori* based on clinical relevance from those routinely measured in practice. They were: age, sex, ethnicity, family history of premature CVD, diabetes, smoking, prior MI or stroke, HF, body mass index (BMI), systolic BP (SBP), ratio of total to high-density lipoprotein cholesterol (TC:HDL), creatinine, dispensing of a BP lowering medication or a statin. Interactions between SBP and BP lowering medications, TC:HDL and statins, and ethnicity and diabetes, were assessed. Ethnicity was defined according to a national prioritisation protocol in the order: Māori, Pacific, Indian, East Asian (Chinese and other non-Indian Asian), European/other. Over 98% of the latter group were classified as European.

Data on BMI and two laboratory variables were missing for 20% patients (BMI), 22% patients (creatinine up to 5 years prior), and 26% patients (TC:HDL up to 1 year prior). Therefore multiple imputation using chained equations was performed.[15] The imputation process is described in Appendix D.

The 2-year event rate was modelled using multivariable Cox regression. Validity of the proportional hazards assumption was confirmed from visual inspection of Schoenfeld residual plots, and linearity of the relationship between each predictor and the log hazard were assessed via plots of Martingale residuals.[16, 17] Where non-linear relationships affected more than 5% of patients, variables were categorised at clinically relevant thresholds (age, creatinine). BMI and SBP were categorised at pre-determined thresholds to assess known U-shaped relationships between these measures and outcome.[18, 19] Alternative approaches to non-linearity were possible, but as the aim was to produce risk scores to be used in clinical practice, categorisation into clinically relevant groups was the preferred approach.

<u>Risk score</u>

9

Multivariable risk from the Cox model was transformed to absolute risk by estimating the baseline hazard at the mean values of continuous covariates and the reference group of categorical covariates. The prognostic index, or sum of each coefficient multiplied by the measured variable, was centred on the mean prognostic index.[20] Risk was calculated for each imputed dataset and the average obtained.

Performance

Model calibration is represented by plots of the observed event rate (from Kaplan-Meier estimates) against predicted event rate within deciles of predicted risk. Global model fit was assessed with the Cox & Snell R^2 and Nagelkerke's R^2.[21,22] Model discrimination was quantified by Harrell's c-statistic[23] and the Gönen & Heller K-statistic[24], and represented visually by plotting the proportion of events against deciles of the cohort after patients were ranked by increasing predicted risk. The model fit and discrimination statistics are the median and interquartile range (IQR) of each statistic derived from the 25 imputation models.

Validation

External validation was performed by applying the risk scores to a cohort of patients with established CVD in the UK (January 2005-December 2010). In collaboration with the CALIBER research group[25], and with approval from the relevant governance groups, anonymised data from linked primary care and national datasets were accessed through an ISO27001 certified data safe-haven provided by University College London. The definition of CVD matched that used to develop the NZ scores. Appendix D presents a comparison of the derivation and validation cohorts.

Analyses were performed using R v3.0.2[26], including the 'survival', 'Hmisc', and 'mice' packages. The cohort study and research process was approved by the NZ Northern Region Ethics Committee Y

(AKY/03/12/314) with subsequent annual approval by the National Multi-Region Ethics Committee (MEC/07/19/EXP).

RESULTS

Data on 30,343 adults with established CVD were recorded in the PREDICT database between January 2006 and April 2012. After limiting to NZ residents aged <80 years, the remaining 24,927 patients form the derivation cohort (Appendix A Table S1).

Almost two-thirds of the cohort were men and median age was 65 years (IQR 58-72 years). The study population were 17% NZ Māori, 10% Pacific, 6% Indian and 4% East Asian, with the remaining two-thirds were European. One-third had diabetes, one-third had a BMI >30 kg/m^2, and one fifth had HF. Over 80% were receiving a BP lowering medication, 75% a statin, and 9% were on warfarin. A total of 2,240 CVD events occurred, of which 1,480 occurred within 2 years (5.9% mean 2-year event rate). Half of these events occurred within the 20% of the cohort identified as having HF.

For two thirds of the cohort, their prior CVD event was a coded hospital admission. The median time between the most recent hospitalisation and risk assessment in primary care was 1.2 years indicating a cohort with stable CVD (Appendix A Table S2). Dates of previous hospitalisations are not consistently available to GPs at the time of patient assessment so have not been included in the current score.

Multivariable model

The risk of a CVD event within 2 years increased with increasing age, particularly after age 60 (Table 1). Compared to European participants, Pacific and Māori were at higher risk and East Asians were at lower risk. Risk increased if the prior event was an MI or stroke hospitalisation, and was significantly higher among people with heart failure, diabetes, or current smokers. Being underweight was associated with significantly greater risk than those of normal BMI however risk progressively decreased as BMI increased above 30 kg/m^2. A U-shaped relationship between SBP and outcome was seen, with increased risk when SBP <100 mmHg or ≥160 mmHg. Risk increased with increasing

TC:HDL and creatinine. Interactions between SBP and BP lowering medications, TC:HDL and statins, and ethnicity and diabetes, were not statistically significant.

Absolute risk

Using the baseline survival estimate in Table 1, median absolute 2-year risk of a CV event was 5.0% (IQR 3.5-8.3%). Event rate stratified by tertiles of predicted risk is shown in Figure 1. The risk equation is provided in Appendix B.

Model performance

Eighty percent of the cohort had an estimated 2-year risk of <10% and Figure 2a shows very good calibration throughout this range of risk. The model slightly under-predicts risk in the 9th decile of risk (10-18%) however is well calibrated in those at highest risk (>18%). Nagelkerke's R^2 was 5.8% (IQR 5.78, 5.83%), Harrell's c-statistic 0.7236 (IQR 0.723, 0.724), and Gönen & Heller's K-statistic 0.6727 (IQR 0.672-0.673) (Appendix C). Discrimination of the score was also assessed visually after ranking the cohort by increasing estimated risk and dividing into deciles (2493 people per decile with 2490 in decile 10; Figure 3). Over half of the events (52%, n=772) occurred in the 20% of the cohort identified as being at highest risk, showing good discrimination.

External validation was performed by applying the NZ equation to the UK dataset of 32,756 patients (Appendix E). Among the UK cohort, 1517 (4.7%) CVD events occurred within 2 years and median estimated risk was 5.2% (IQR 3.7-8.2%). The calibration plot showed the NZ equation generally overestimated risk in the UK cohort. Overestimation was minor for the 60% of the cohort with an event rate of <5%, increased slightly for the 20% of the cohort with an event rate of 5-7%, then clearly overestimated risk in the highest 20% of the cohort (Figure 2b).

DISCUSSION

We present a new algorithm to improve CVD risk stratification in primary care for patients with established CVD. This is the first contemporary score that enables quantitative risk estimation for all patients with CVD, at a time distant to their CVD event. The risk of a future CVD event varied widely, with the median 2-year risk of CV death, MI or stroke ranging from 2% in the lowest risk decile to 34% in the highest risk decile. Over half of all CVD events occurred in the twenty percent of the cohort with the highest risk. When combined with primary prevention scores, this new score will enable quantitative stratification of risk across the continuum of primary and secondary CVD prevention.

External validation was performed by applying the score to a cohort of patients in a different country, allowing us to assess its geographic transportability.[27] The median risks estimated for UK and NZ patients were very similar (UK 5.2%, NZ 5.0%) although the NZ score tended to overestimate risk in the UK cohort. Compared to the NZ cohort, the UK cohort had a very different ethnic profile (no Māori or Pacific Islander), lower rates of prior MI or stroke (25% vs 47%) and diabetes (29% vs 36%), and greater BP lowering (91% vs 82%) and statin use (87% vs 75%) (Appendix D). Risk overestimation was greatest among the 20% of patients at highest risk, which are almost entirely those with HF. Heart failure was defined in the same way in both datasets however the event rate in this group of patients was significantly lower in the UK than in NZ (8.5% vs 14.6%). The majority (70%) of HF patients in the NZ cohort were defined via a prior hospitalisation for HF. In comparison, only 31% of patients with HF in the UK cohort had a prior hospitalisation for HF. Thus differences in clinical practice between the two countries, such as the use of loop diuretics and coding of HF, may have contributed to the difference in score performance. Further analysis, including characterising the patients in whom the scores did not predict risk well, is beyond the scope of the current paper and will be explored in subsequent work.

Within secondary prevention models, a number of factors have been consistently important in their association with a subsequent event. In a Framingham model[28], TC:HDL and diabetes were significant predictors, with SBP and smoking also significant among women. The REACH registry[29] of >45,000 patients with CVD from 29 countries was the basis of a model with similar eligibility and outcome definitions to our study, and similarly found HF, diabetes, smoking, a prior ischaemic event, and BMI<20 kg/m^2 to be statistically significant. Models from patients with CAD [3, 30, 31] also found similar significant factors, although HF was not significant in the EUROPA trial when predicting CV death, MI or cardiac arrest. The CALIBER group included 31 variables in models for stable CAD patients to estimate risk of all-cause mortality and of non-fatal MI or coronary death.[3] In contrast to other studies, we did not see a gender difference in risk and instead suggest a levelling effect once CVD has been established. This may be influenced by the greater representation of men in other cohorts, comprising up to 85% of subjects in clinical trials.

Clinical implications

The risk estimate guides clinicians and patients in their discussions of individualised care planning. Instead of patients with prior CVD simply being told they are at "clinically high risk", risk stratification informs more individualised care, and aids a discussion of factors amenable to risk reduction with lifestyle and medication changes. Most of the patients in this cohort were already on secondary prevention therapy and the observed risk is therefore an on-treatment risk, although this does not imply optimal dosage or therapy combination. High risk can be mitigated by ensuring patients are not only receiving standard secondary prevention therapies, but that dosages of medications are optimised for the individual, other risk factors (such as smoking) are readdressed, and management of coexisting conditions such as diabetes are optimised. The score will also facilitate risk communication, aiding the healthcare professional/patient interaction and thus optimising implementation of risk reduction therapies. Clinical trials stratifying patients on the basis of risk are

needed to quantify the benefit of additional therapies. A relatively low risk should not be used as a justification for not intensifying management or for withdrawing treatment.

Limitations

Additional variables that may inform CVD risk, such as NT-proBNP, hsCRP and cognitive status, are not routinely measured in all patients and so are not available for inclusion in risk model development using general practice records and administrative data. Similarly, an index of socioeconomic disadvantage was included in preliminary analyses, in which increasing deprivation was associated with a modest increase in risk (adjusted HR = 1.07 per quintile of deprivation score). However as socioeconomic indices are country-specific, and may not be routinely available in primary care, we elected to remove this variable from the risk score.

The risk horizon of 2 years is shorter than the more familiar 5 or 10 years associated with primary prevention scores, which is important to bear in mind to avoid underestimation of risk in this patient group. Future risk scores will include extending the risk horizon to 5 years (when there is sufficient duration of follow up for such models to be developed).

It is difficult to identify an appropriate population for external validation. The NZ data used to develop the risk scores came from an explicit standardised CVD risk assessment undertaken in primary care and augmented by national dispensing and regional laboratory data. In contrast, the UK data used for validation were from incidental primary care visits not necessarily related to a CVD risk assessment. Completeness and timing of the measurements were therefore variable and, as discussed above, the NZ and UK cohorts have different risk profiles and event rates, particularly among HF patients.

Conclusions

16

A new score has been developed to improve CVD risk stratification for patients with prior CVD in the community. The score can identify the highest risk patients who may benefit from more intensive risk management, and when made available alongside models for predicting risk in those without prior CVD, will enable risk stratification over the life-course of cardiovascular risk and disease.

Disclosures

Outside of this study SW has received a research grant from Roche Diagnostics Ltd.

REFERENCES

1.   Antman EM, Cohen M, Bernink P, et al. The TIMI Risk Score for Unstable Angina/Non-ST Elevation MI. A method for prognostication and therapeutic decision making. *J Am Med Assoc*. 2000;284:835.

2.   Pieper KS, Gore JM, FitzGerald G, et al and for the Global Registry of Acute Coronary Events (GRACE) Investigators. Validity of a risk-prediction tool for hospital mortality: The Global Registry of Acute Coronary Events. *Am Heart J*. 2009;157:1097-1105.

3.   Rapsomaniki E, Shah A, Perel P, et al. Prognostic models for stable coronary artery disease based on electronic health record cohort of 102 023 patients. *Eur Heart J*. 2014;35:844-852.

4.   New Zealand Guidelines Group. *New Zealand Cardiovascular Guidelines Handbook: A summary resource for primary care practitioners.* 2nd ed. Wellington: Ministry of Health; 2009.

5.   Perk J, De Buyzere M, Gohlke H, et al. European guidelines on cardiovascular disease prevention in clinical practice (version 2012). *Eur Heart J*. 2012;33:1635-1701.

6.   Goff D, Lloyd-Jones D, 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;63:2935-59.

7.   Anderson K, Odell P, Wilson P, et al. Cardiovascular disease risk profiles. *Am Heart J*. 1991;121:293-298.

8.   Hippisley-Cox J, Coupland C, Vinogradova Y, et al. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ 2002;324:1570-1576*. 2007:bmj.39261.471806.55.

9.   Woodward M, Brindle P, Tunstall-Pedoe H and for the SIGN group on risk estimation. Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). *Heart*. 2007;93:172-176.

10.   Conroy RM, Pyorala K, Fitzgerald AP, et al and on behalf of the SCORE project group. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J*. 2003;24:987-1003.

11.   Jackson R, Lawes C, Bennett D, et al. Treatment with drugs to lower blood pressure and blood cholesterol based on an individual's absolute cardiovascular risk. *Lancet*. 2005;365:434-41.

12.   Broadbent E, Leggat A, McLachlan , et al. Providing cardiovascular risk management information to acute coronary syndrome patients: a randomized trial. *Br J Health Psych*. 2013;18:83-96.

13.   Kerr AJ, Broad J, Wells S, et al. Should the first priority in cardiovascular risk management be those with prior cardiovascular disease? *Heart*. 2009;95:125-129.

14.	Wells S, Riddell T, Kerr A, et al. Cohort Profile: The PREDICT Cardiovascular Disease Cohort in New Zealand Primary Care (PREDICT CVD-19). *Int J Epidemiol*. 2015.

15.	van Buuren S and Oudshoorn C. Multivariate Imputation by Chained Equations: MICE V1.0 User's manual. . *TNO Report PG/VGZ/00038*. 2000.

16.	Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrica*. 1982;69:239-241.

17.	Therneau T, Grambsch P and Fleming T. Martingale-based residuals for survival models. *Biometrica*. 1990;77:147-160.

18.	Leonardi-Bee J, Bath PM, Phillips SJ, et al. Blood pressure and clinical outcomes in the International Stroke Trial. *Stroke*. 2002;33:1315-1320.

19.	Flack JM, Neaton J, Grimm Jr R, et al and for the Multiple Risk Factor Intervention Trial Research Group. Blood pressure and mortality among men with prior myocardial infarction. *Circulation*. 1995;92:2437-2445.

20.	Royston P and Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC Medical Research Methodology*. 2013;13.

21.	Cox D and Snell E. *The analysis of binary data.* Second ed. London: Chapman & Hall; 1989.

22.	Nagelkerke N. A note on a general definition of the coefficient of determination. *Biometrika*. 1991;78:691-692.

23.	Harrell FE, Califf RM, Pryor DB, et al. Evaluating the yield of medical tests. *J Am Med Assoc*. 1982;247:2543-2546.

24.	Gonen M and Heller G. Concordance probability and discriminatory power in proportional hazards regression. *Biometrika*. 2005;92:965-70.

25.	Denaxas SC, George J, Herrett E, et al. Data Resource Profile: Cardiovascular disease research using linked bespoke studies and electronic health records (CALIBER). *Int J Epidemiol*. 2012;41:1625-1638.

26.	R: A language and environment for statistical computing. [computer program]. Version 3.0.2. Vienna, Austria: R Foundation for Statistical Computing; 2013.

27.	Collins GS, Reitsma J, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Br Med J*. 2014;350:g7594.

28.	D'Agostino RB, Russell MW, Huse DM, et al. Primary and subsequent coronary risk appraisal: New results from the Framingham study. *Am Heart J*. 2000;139:272-281.

29.	Bhatt DL, Eagle KA, Ohman EM, et al and for the REACH Registry Investigators. Comparative determinants of 4-year cardiovascular event rates in stable outpatients at risk of or with atherothrombosis. *J Am Med Assoc*. 2010;304:1350-1357.

30.   Cui J, Forbes A, Kirby A, et al. Laboratory and non-laboratory-based risk prediction models for secondary prevention of cardiovascular disease: the LIPID study. *Eur J Cardiovasc Prev Rehabil*. 2009;16:660.

31.   Battes L, Barendse R, Steyerberg E, et al. Development and validation of a cardiovascular risk assessment model in patients with established coronary artery disease. *Am J Cardiol*. 2013;112:27-33.

Figure legend

Figure 1   **Event rate stratified by tertiles of predicted risk**

The absolute risk bounds for each tertile are: 1.0-3.7%, 3.7-6.9%, 6.9-75.7%. Dashed lines represent the 95% confidence interval.

Figure 2   **Model calibration in the development and external validation cohorts**

Calibration of the risk equation in a) the development cohort (New Zealand), and b) the external cohort (United Kingdom). For both plots, calibration is shown as the actual survival from Kaplan-Meier analysis against estimated risk (in deciles). Dashed line = perfect calibration.

Figure 3   **Discrimination plot showing the percentage of events that occurred per decile of the development cohort, ranked by estimated risk.**

The absolute risk bounds for each decile are: 1.0-2.6%, 2.6-3.2%, 3.2-3.8%, 3.8-4.3%, 4.3-5.0%, 5.0-5.9%, 5.9-7.3%, 7.3-10.1%, 10.1-16.0%, 16.0-75.7%.

Table 1   **Multivariable models of time to subsequent cardiovascular event within 2 years**

Variable	Levels	Adjusted Hazard Ratio (95% confidence interval)
Patient factors		
Male		1.04 (0.93, 1.16)
Age, years	50-59	1.05 (0.83, 1.34)
	60-69	**1.27 (1.01, 1.61)**
	70-79	**1.69 (1.33, 2.15)**
Ethnicity	NZ Māori	**1.26 (1.10, 1.45)**
	Pacific	**1.51 (1.29, 1.77)**
	Indian	1.18 (0.95, 1.48)
	East Asian	**0.60 (0.40, 0.90)**
Medical history		
Prior MI or stroke		**1.52 (1.37-1.69)**
Heart failure		**2.85 (2.53-3.22)**
Diabetes		**1.36 (1.22, 1.53)**
Current smoker		**1.27 (1.10-1.47)**
Clinical factors		
Body mass index, kg/m^2	<18.5	**2.26 (1.32, 3.88)**
	18.5-25	1.10 (0.92, 1.30)
	30-35	**0.85 (0.73, 0.99)**
	35-40	**0.77 (0.63, 0.93)**
	≥40	**0.68 (0.53, 0.86)**
Systolic BP, mmHg	<100	**1.57 (1.08, 2.28)**
	120-140	0.99 (0.85, 1.14)
	140-160	1.10 (0.94, 1.30)
	≥160	**1.39 (1.12, 1.74)**
Laboratory values		
Total:HDL cholesterol		**1.11 (1.07, 1.17)**
Creatinine, μmol/L	100-150	**1.34 (1.17, 1.53)**
	≥150	**1.96 (1.63, 2.36)**
Medications		
BP lowering		1.14 (0.96, 1.36)
Statin		1.10 (0.96, 1.25)

| Baseline survival | 0.94235 |

NZ=New Zealand, BP=blood pressure, MI=myocardial infarction, HDL=high-density lipoprotein.

Referent for: sex=women; age=18-49 years; ethnicity=European/other; BMI=25-30 kg/m^2; systolic BP=100-120 mmHg; creatinine=<100 umol/L.

Values in bold represent statistical significance (p<0.05).

Figure 1   **Event rate stratified by tertiles of predicted risk**
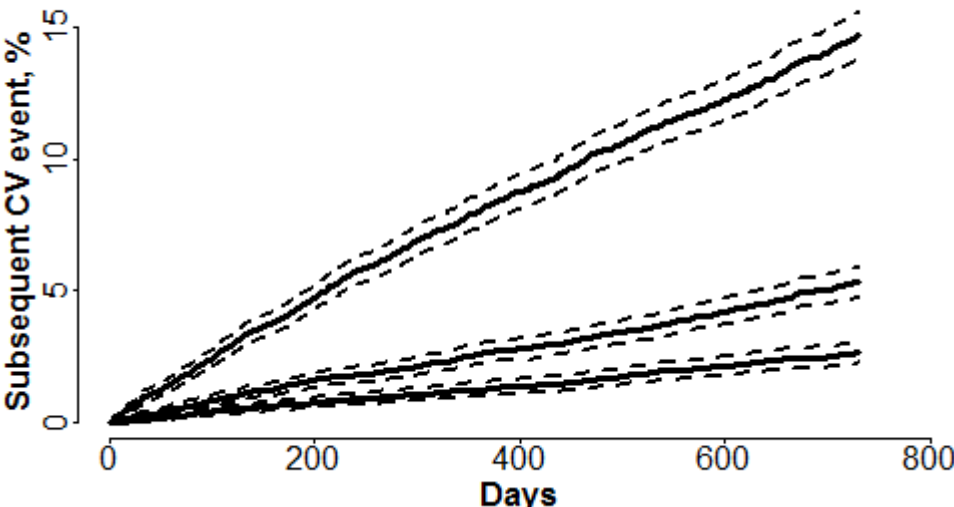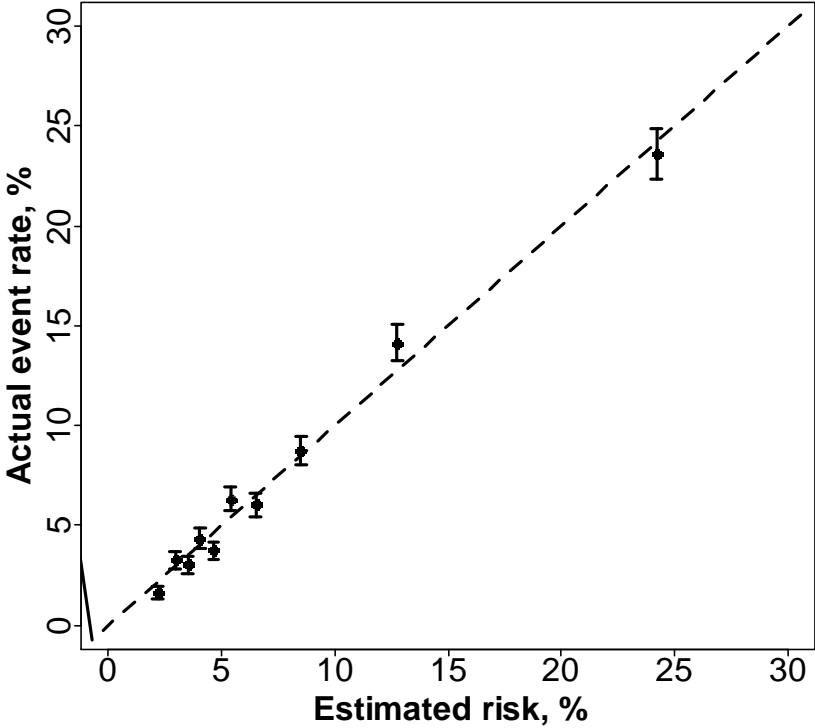
Figure 2   Model calibration in the development and external validation cohorts
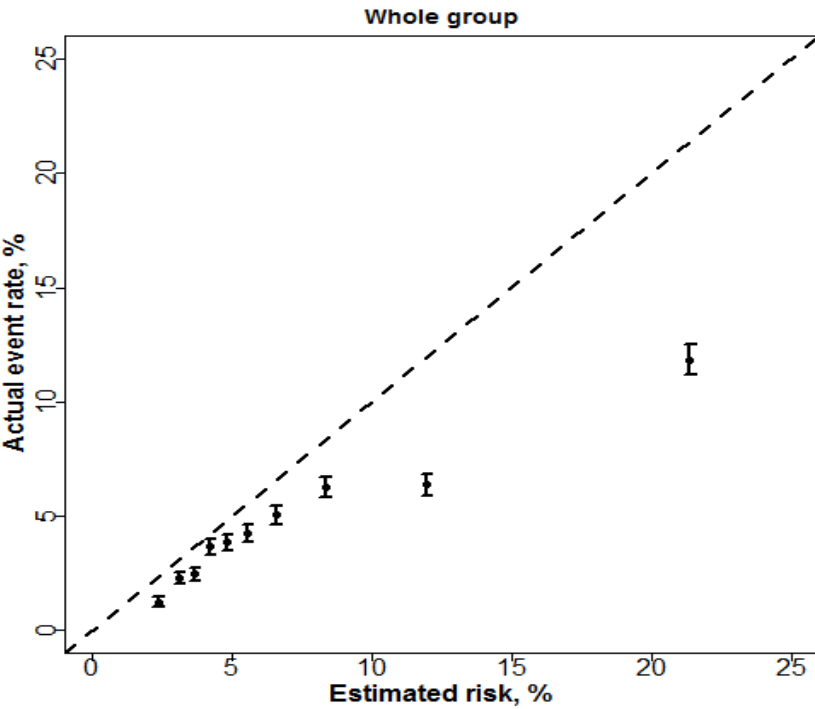
a)

b)

Figure 3   Discrimination plot comparing the proportion of the cohort, ranked from lowest to highest predicted risk , to the percentage of events that actually occurred in that proportion of the cohort
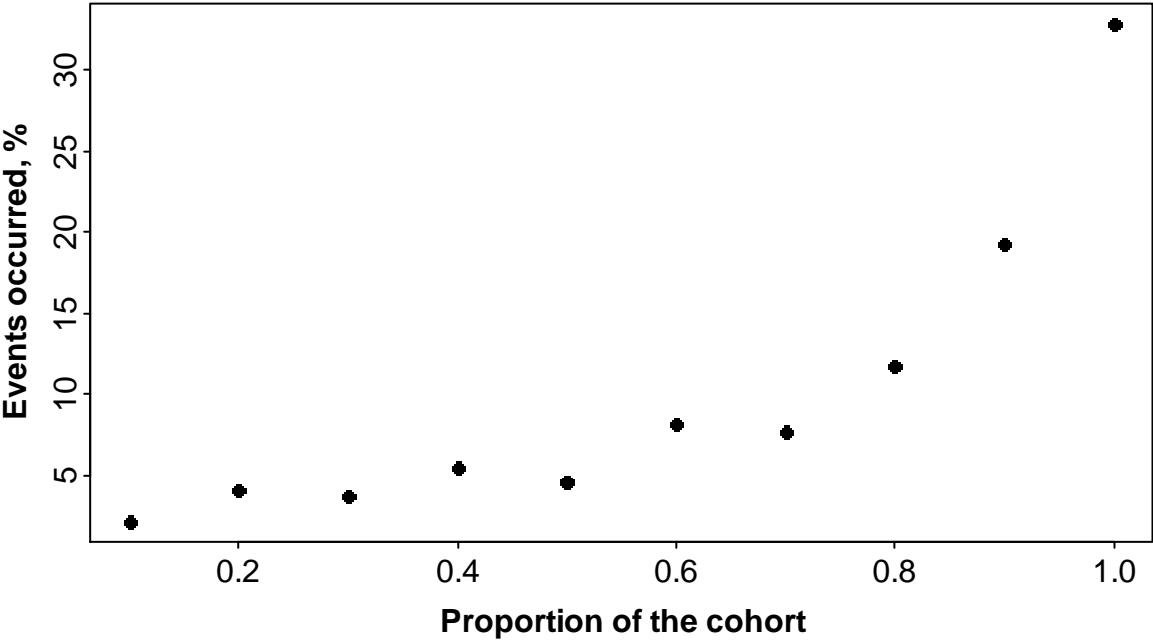
SUPPLEMENTAL MATERIAL

Developing and validating a cardiovascular risk score for patients in the community with prior cardiovascular disease

Katrina K Poppe

Rob N Doughty

Sue Wells

Dudley Gentles

Harry Hemingway

Rod Jackson

Andrew J Kerr

Corresponding author

Katrina Poppe PhD COPSCT

Epidemiology and Biostatistics, School of Population Health, University of Auckland

Private Bag 92019, Auckland 1142, New Zealand

k.poppe@auckland.ac.nz

ph +64-9-9239803

APPENDIX A   DERIVATION COHORT

Table S1   Participant characteristics, n=24,927

Men	15,814 (63)
Age, years	65 (58, 72)
≥ 70 years	8,201 (33)
European/other*	15,665 (63)
NZ Māori	4,197 (17)
Pacific	2,592 (10)
Indian	1,503 (6)
East Asian	970 (4)
Family history of premature CVD	5,141 (21)
Medical history	
Prior MI or stroke	11,695 (47)
Heart failure	5,100 (20)
Diabetes	8,930 (36)
Current smoker	3,678 (15)
Clinical measurements	
Body mass index (BMI), kg/m^2	29 (26, 33)
BMI < 18.5 kg/m^2	163 (0.7)
BMI ≥30 kg/m^2	10,992 (44)
Systolic BP (SBP), mmHg	130 (121, 141)
SBP < 100 mmHg	253 (1.0)
SBP ≥160 mmHg	1,383 (5.5)
TC:HDL	3.6 (3.0, 4.4)
Creatinine, μmol/L	85 (73, 99)
Medications	
BP lowering	20,513 (82)
Statin	18,739 (75)
Anticoagulation	2,339 (9)
Follow-up	
Total follow-up, years	2.0 (0.9, 3.3)
CV deaths	764 (3.1)
CV events	2,240 (9.0)

By 2 years	
CV deaths	423 (1.7)
CV events	1,480 (5.9)

Values are n (%) or median (inter-quartile range). *<2% were "other"

NZ=New Zealand, BP=blood pressure, MI=myocardial infarction, TC:HDL=ratio of total to high-density lipoprotein, CV=cardiovascular, CV event = MI, ischaemic or haemorrhagic stroke, or CV death

Table S2   **Time between hospitalisation and index risk assessment in primary care**

Timeframe	Number of patients	%
< 30 days	1277	7.8
30 days – 6 months	3584	22.0
6 months – 12 months	2677	16.4
12 – 18 months	1706	10.5
18-24 months	1244	7.6
≥ 24 months	5795	35.6
TOTAL	**16283**	

APPENDIX B   RISK SCORE

	Coefficient	EXAMPLE	
		Patient A	Variable * coefficient
Male	0.03587575	Female	0
Age 50-59 years	0. 05342182		0
Age 60-69 years	0. 24286832	63 years old	0.24286832
Age 70-79 years	0.52462088		0
East Asian	- 0.50879097		0
Indian	0.16911057	Indian	0.16911057
Māori	0.23015796		0
Pacific	0.41315010		0
Diabetes	0.31073697	Diabetes	0.31073697
Current smoker	0.23873539	Non-smoker	0
Prior MI or stroke	0.42372976	Prior MI	0.42372976
Heart failure	1.07493660	With heart failure	1.07493660
BMI < 18.5	0.81581102		0
BMI 18.5-25	0.09252985		0
BMI 30-35	- 0.16091766		0
BMI 35-40	- 0.26400754	BMI = 36 kg/m^2	-0.26400754
BMI ≥ 40	- 0.38904755		0
SBP < 100	0.44988432		0
SBP 120-140	- 0.01460494		0
SBP 140-160	0.09897459	SBP = 142 mmHg	0.09897459
SBP ≥160	0.33063441		0
TC:HDL, per unit	0.10795178	TC:HDL = 4.2	0.4533975
Creatinine 100-150	0.29119313	Creatinine = 80 µmol/L	0
Creatinine≥150	0.67363029		0
Statin	0.09119116	Taking a statin	0.09119116
BP lowering	0.13462192	Taking ACEi & b-blocker	0.13462192
		PI = Sum =	2.73556
Mean PI	1.577184		
Baseline survival	0.94235		

PI = prognostic index

<u>PREDICT secondary prevention risk score</u>

2-year risk of CVD event = 1 − Baseline survival$^{\exp(PI - \text{mean PI})}$

$$= 1 - 0.94235^{\exp(PI - 1.577184)}$$

Thus Patient A's 2-year risk of a subsequent CVD event

$$= 1 - 0.94235^{\exp(2.73556 - 1.577184)}$$

$$= 0.1723014$$

$$= 17.2\%$$

APPENDIX C   MODEL PERFORMANCE

		Performance metrics
Apparent performance*		
Model fit	Cox & Snell R^2, %	4.00 (3.98, 4.02)
	Nagelkerke R^2, %	5.81 (5.78, 5.83)
Discrimination	Harrell c-statistic	0.7236 (0.7228, 0.7244)
	Gönen & Heller	0.6727 (0.6725, 0.6734)
Calibration	Observed v expected plots	
Internal validation**		
Model fit	Cox & Snell R^2, %	4.09 (4.07, 4.12)
	Nagelkerke R^2, %	5.95 (5.90, 5.98)
Discrimination	Harrell c-statistic	0.7258 (0.7249, 0.7264)
	Gönen & Heller	0.6754 (0.6749, 0.6759)
External validation		
Calibration	Observed v expected plots	

*Values are median (IQR) from 25 imputation models; **Values are median (IQR) from all bootstrap and imputation

APPENDIX D   MULTIPLE IMPUTATION OF THE DERIVATION (PREDICT) DATA

In the PREDICT cohort, the proportion missing from each variable were: BMI 20%, TC:HDL 26%, creatinine 22%. It was reasonable to assume that data were missing at random therefore multiple imputation using chained equations was performed.

All predictors in the risk model were in the imputation model, as well as glucose, urate, WCC, and Hb. The event indicator and the Nelson-Aalen estimator were also included. There were variable patterns of missingness among patients with and without diabetes thus imputation was performed separately in those with and without diabetes then combined for each iteration of the substantive model.   For each of 25 imputations, 10 random draws from the distribution defined by the imputation equations were made for each of the variables with missing data.

APPENDIX E   EXTERNAL VALIDATION

Table S3   Comparison of participant characteristics in derivation and external validation cohorts

	New Zealand (PREDICT)	United Kingdom (CALIBER)*	United Kingdom (CALIBER final)**
n	24,927	80,425	32,756
Patient factors			
Male	63%	60%	61%
Age, years	65 (58, 72)	69 (62, 74)	70 (62, 75)
European	63%	82%	97%
NZ Māori	17%	Not available	Not available
Pacific	10%	Not available	Not available
Indian	6%	2%	3%
East Asian	4%	0.1%	0.1%
Family history of prem. CVD	21%	Not available	0%
Medical history			
Prior MI or stroke	47%	24%	25%
Heart failure	20%	18%	20%
Diabetes	36%	24%	29%
Current smoker	15%	23%	22%
Clinical factors			
Body mass index, kg/m^2	29 (26, 33)	28 (25, 33)	29 (25, 33)
Systolic BP, mmHg	130 (121, 141)	139 (125, 149)	139 (126, 150)
Laboratory values			
TC:HDL	3.6 (3.0, 4.4)	3.5 (2.9, 4.3)	3.5 (2.9, 4.3)
Glucose, mmol/L	5.7 (5.1, 7.2)	5.8 (5.0, 8.1)	5.8 (5.0, 8.1)
Creatinine, µmol/L	85 (73, 99)	94 (82, 109)	94 (81, 109)
Urate, mmol/L	0.36 (0.30, 0.43)	0.35 (0.28, 0.43)	0.35 (0.29, 0.43)
White cell count, mmol/L	7.1 (6.0, 8.7)	7.1 (6.0, 8.6)	7.1 (5.9, 8.7)
Haemoglobin, g/L	141 (129, 150)	140 (131, 150)	138 (128, 149)
Medications			
BP lowering	82%	84%	91%
Statin	75%	78%	87%
Warfarin	9%	13%	14%
By 2 years			
CV events	5.9%	5.7%	4.7%

*After exclusions for inconsistent dates or measurements outside the nominal start date of 1 January 2006;**Only includes ethnic groups included in the NZ equations. Family history of premature CVD was not available in the UK data so this was set to 'none' for all subjects.