

Published in final edited form as:

JAMA Psychiatry. 2015 February 1; 72(2): 143–151. doi:10.1001/jamapsychiatry.2014.2133.

The PRIMROSE cardiovascular risk prediction models for people with severe mental illness:

Results from the PRIMROSE Research Programme: PRedlction and Management of cardiovascular Risk in peOple with SEvere mental illnesses

David PJ Osborn, PhD^{1,2,‡,*}, Sarah Hardoon, PhD^{1,3,‡}, Rumana Z Omar, PhD⁶, Richard IG Holt, PhD⁴, Michael King, PhD^{1,2}, John Larsen, PhD⁵, Louise Marston, PhD³, Richard W Morris, PhD³, Irwin Nazareth, PhD³, Kate Walters, PhD³, and Irene Petersen, PhD³

¹UCL Division of Psychiatry, UCL, London UK

²Camden and Islington NHS Foundation Trust, London UK

³Research Department of Primary Care and Population Health, UCL, London UK

⁴Human Development and Health Academic Unit, Faculty of Medicine, University of Southampton, Southampton UK

⁵Rethink Mental Illness, London UK

⁶Department of Statistical Science, UCL, London UK

Abstract

Importance—People with Severe Mental Illness (SMI) including schizophrenia and bipolar disorder have excess cardiovascular disease (CVD). Risk prediction models, validated for the general population, may not accurately estimate cardiovascular risk in this group.

Objectives—To develop and validate a risk model exclusive to predicting CVD events in people with SMI, using established cardiovascular risk factors and additional variables.

Design—Prospective cohort and risk score development study.

Setting—UK Primary care

Participants—38,824 people with a diagnosis of SMI (schizophrenia, bipolar disorder or other non-organic psychosis) aged 30-90 years. Median follow-up 5.6 years with 2,324 CVD events (6%).

Main outcomes and measures—Ten year risk of first cardiovascular event (myocardial infarction, angina pectoris, cerebrovascular accidents or major coronary surgery). Predictors included age, gender, height, weight, systolic blood pressure, diabetes, smoking, body mass index (BMI), lipid profile, social deprivation, SMI diagnosis, prescriptions of antidepressant, antipsychotics and reports of heavy alcohol use.

*Corresponding author- Dr Osborn UCL Division of Psychiatry, UCL, Charles Bell House, Riding House Street, London W1W 7EJ. +44 (0)20 7679 9028 (direct) d.osborn@ucl.ac.uk.

‡Joint first authors- equal contributors

Supplemental material: The online only supplement contains additional eMethods, eResults, efigures and e-tables.

Results—We developed two risk models for people with SMI: The PRIMROSE BMI model and a lipid model. These mutually excluded lipids and BMI. From cross-validations, in terms of discrimination, for men, the PRIMROSE lipid model D statistic was 1.92 (1.80-2.03) and C statistic was 0.80 (0.76-0.83) compared to 1.74 (1.54-1.86) and 0.78 (0.75-0.82) for published Framingham risk scores; in women corresponding results were 1.87 (1.76-1.98) and 0.80 (0.76-0.83) for the PRIMROSE lipid model and 1.58 (1.48-1.68) and 0.76 (0.72-0.80) for Framingham. Discrimination statistics for the PRIMROSE BMI model were comparable to those for the PRIMROSE lipid model. Calibration plots suggested that both PRIMROSE models were superior to the Framingham models.

Conclusion and relevance—The PRIMROSE BMI and lipid CVD risk prediction models performed better in SMI than models which only include established CVD risk factors. Further work on their clinical and cost effectiveness is needed to ascertain the best thresholds for offering CVD interventions.

Introduction

It is well established that people with severe mental illnesses (SMI), such as schizophrenia and bipolar disorder, have excess rates of cardiovascular disease (CVD) including myocardial infarctions and strokes¹. The risk of dying from CVD is threefold higher in people with SMI under 50 years and two-fold in those aged 50-75². There has been an increase in clinical and research effort addressing this problem, but we lack knowledge regarding the most effective ways to predict and manage cardiovascular risk in people with SMI. We know that the conventional cardiovascular risk factors including smoking, dyslipidaemia (with high total cholesterol and triglycerides and lower levels of HDL cholesterol), diabetes and obesity and possible raised levels of hypertension are more common in people with SMI^{3,4}, especially those with well-established mental disorders⁵. There may even be a shared genetic predisposition to comorbidities. People with SMI are less likely to exercise and to have unhealthy diets and may receive inferior physical health care⁶. Antipsychotic medications may contribute to cardiovascular risk through increasing weight gain and impacting on glucose and lipid metabolism although mortality studies which explore the role of antipsychotic medication in premature deaths show inconsistent findings.^{2,7,8} For example, the FIN-11 study⁷ reported reduced cardiovascular mortality in people treated with olanzapine and clozapine, but methodological issues in the study have been critiqued in detail, particularly unmeasured confounding⁹.

For the general population, cardiovascular risk is managed by employing CVD risk scores to determine the absolute risk for an individual patient and therefore the likely benefit of prescribing lipid lowering medications and/ or other interventions. The most established scores are the Framingham risk scores¹⁰, currently available as both a BMI model (including BMI but not laboratory results for blood lipids) and a lipid model (including total cholesterol and HDL cholesterol, but not BMI).

Clinical guidelines, such as the UK NICE guidelines on schizophrenia, recommend more intensive screening for cardiovascular risk in people with SMI^{1,11}. However, we do not know how well the modifiable cardiovascular risk factors, included in models such as

Framingham, predict CVD risk in people with SMI. There is reason to believe that existing models may not accurately determine the high level of risk conferred by having a long term severe mental illness. First, the established scores such as the Framingham risk scores were developed excluding people with SMI and since then have not been tested in this population and secondly they do not consider SMI specific exposures, such as antipsychotic medication. No previous studies have assessed the performance of CVD risk prediction models in people with SMI.

Using data from a large UK primary care database, The Health Improvement Network (THIN)¹², we aimed to develop and validate cardiovascular risk prediction models specific to people with SMI- the PRIMROSE models. These new models, from the PRIMROSE study, included traditional cardiovascular risk factors and additional SMI-specific variables. We compared the performance of these new risk prediction models against existing published Framingham scores in people with SMI since these are widely used internationally and their coefficients are readily available to allow comparison.

This work formed part of a programme of research, PRIMROSE funded by the UK National Institute for Health Research. www.ucl.ac.uk/primrose

Method

Study design

A prospective study involving the development and validation of a ten year risk score for predicting newly recorded cardiovascular events in people with SMI.

Setting

We used The Health Improvement Network (THIN)¹² United Kingdom primary care database which includes data from routine clinical practice. Primary care physicians and staff use a hierarchical system of Read codes to enter information in THIN such as symptoms and diagnoses, during clinical appointments and administration¹³. This creates a longitudinal record for each patient. At the time we developed the PRIMROSE risk models, THIN included almost 10 million patients with geographical coverage broadly representative of the UK population¹⁴. Approximately 98% of the population is registered with a GP in the UK¹³. THIN data are subject to a range of quality assurance procedures¹⁶ and have been successfully used in wide ranging studies on cardiovascular diseases^{17, 18} including cardiovascular risk score validation work in the general population¹⁹. Primary care data are a particularly suitable source for assessing cardiovascular risk in people with SMI in the UK since most people with SMI are registered with a general practitioner whom they see frequently and most of the required data for risk scores (such as laboratory and blood pressure measurements) are available due to policy initiatives which incentivise annual cardiovascular screening²⁰. SMI diagnoses have been validated in UK general practice²¹.

Participants

We included individuals aged between 30 and 90 years with a diagnostic entry in their primary care electronic health records for a severe mental illness at any time during their

follow-up period. We defined SMI as 1) schizophrenia and schizoaffective disorder 2) Bipolar affective disorder and 3) Other non-organic psychoses. We created lists of the diagnostic codes used by GPs or administrators. The codes are usually based on assessments by mental health specialists. We extracted data between 1995 and December 2010.

Main Outcome

Newly recorded fatal and non-fatal cardiovascular events defined as a diagnostic record for: Myocardial infarction, angina pectoris, coronary heart disease, major coronary surgery and revascularisation, cerebrovascular accident (CVA) and transient ischaemic attacks (TIA).

Analysis

We developed two PRIMROSE risk models, the BMI model and Lipid model. The online-only supplement (see eMethods) contains detailed description of the follow-up period, the variables considered in our analysis, the development of the PRIMROSE risk models and their tenfold internal cross-validation. It also describes the imputation of missing data and our sample size calculation.

In summary, we performed cox regression with backwards elimination to derive the PRIMROSE models. The variables considered in the models are listed in tables 1-3. We compared the performance of different models by calculating the D^{22} statistic and C^{23} index for discrimination, with calibration plots and by assessing the numbers of people classified as “high risk” of CVD over ten years (>20%) who went on to have a CVD event.

After comparing the PRIMROSE risk models against published Framingham models, we performed two supplementary comparisons. First we wanted to ascertain whether our results could be explained by differences between North American and UK source populations. Thus we re-estimated the Framingham model to the UK general population and compared the PRIMROSE risk models against this. Secondly we wanted to examine if there is anything gained using the PRIMROSE models over and above a simple re-estimate of the coefficients for the Framingham in the SMI population, thus we compared the PRIMROSE model to a model which only included the variables in Framingham risk score, but with their coefficients re-estimated within our SMI cohort.

Results

We identified 38,824 individuals who met eligibility criteria (see flowchart, efigure 1), in 430 general practices. There was a median follow-up period of 5.6 years (interquartile range 2.5-9.2 years), which compares favorably with other European risk score studies in primary care databases¹⁹. 8020 people (21%) had 10 years follow-up or more. There were 18,417 men (47%) in the cohort, the mean age was 49.5 years and more patients lived in areas with greater social deprivation (table 1). Approximately a third of patients had a diagnosis of schizophrenia, with marginally fewer having diagnoses of bipolar disorder and psychosis not otherwise specified. In terms of data completeness, during the period over which the longitudinal imputation model was applied (see eMethods for details of imputation), 96% of people with SMI had a record for smoking status 89% for Systolic Blood Pressure, 79% for

weight, 66% for height, 52 % for total cholesterol, and HDL cholesterol measurements for 43% . The remaining variables were all complete. The imputed data are presented in etable 1.

There were 2,324 newly recorded CVD events during follow-up, corresponding to a crude incidence rate of 9.72 (9.33-10.1) per 1000 person years. CVD incidence increased with age and was higher in men within all age categories (etable 2). The most common events were ischaemic or unspecified strokes (n=778; 33.5% of the total events), myocardial infarctions (n=414; 17.8%), transient ischaemic attacks (349; 15.0%), angina (n=325; 14.0%), coronary heart disease unspecified (n=304; 13.1%), unstable angina (n=65; 2.8%) and haemorrhagic stroke (n=46; 2.0%).

The associations between age, sex, deprivation and established cardiovascular risk factors were all in the expected direction for known CVD risk factors (table 2). After adjusting for age and sex, CVD was positively associated with higher total cholesterol, lower HDL cholesterol, increasing weight, deprivation, blood pressure and age as well as smoking and male sex (table 2). SMI diagnosis, antidepressant use and a history of heavy alcohol consumption were also predictive of CVD. There was no strong evidence of non-linear associations between the continuous variables and CVD hazard, with the exception of age in years, for which a log transformation improved linearity.

Development of PRIMROSE models

For the PRIMROSE BMI model all variables from table 2 were retained in the model after backwards elimination with the exception of prescription of lithium at baseline (Table 3). For the PRIMROSE lipid model, receipt of both lithium and first generation antipsychotics were eliminated from the model, however, baseline receipt of both second generation antipsychotics and anti-depressants were retained as predictors of 10 year CVD risk (Table 3). The complete formulae for the new PRIMROSE models, for application in practice, are available in the eResults section.

Discrimination and calibration

Both PRIMROSE risk scores (BMI and lipid) performed well compared to the published Cox Framingham BMI and lipid models in the SMI cohort (table 4, Figure 1 and etables 3-4). In terms of their discrimination statistics the D statistics were better for the PRIMROSE models in both men and women (higher D scores indicate better discrimination and an increase of 0.1 point has been defined as important²³). C statistics were broadly similar (table 4, etables 3-4). The PRIMROSE lipid model discrimination was not substantially better than the PRIMROSE BMI model discrimination. The calibration plots suggested that the PRIMROSE lipid and BMI models predicted more accurately than the published Cox Framingham models, with greater agreement between the predicted and observed risks for CVD (figure 1). The published Cox Framingham models tended to over-predict CVD risk in this population, particularly among men.

Risk classification

When “high risk” for CVD was defined using the conventional threshold of 20% risk over 10 years, the PRIMROSE models were better at predicting the percentage of people with

SMI who would go onto have a CVD event (Table 4). For men categorized as high risk (>20% risk), the proportions developing CVD within 10 years were 19.2% for the PRIMROSE lipid model and 18.6% for the PRIMROSE BMI model, both performing better than the corresponding published Framingham CVD risk models⁸ (13.3% lipid and 12.3% BMI). For men categorized as low risk (less than 20%), slightly more developed a CVD event when the PRIMROSE model was applied, compared to the published Cox Framingham model (3.3% with PRIMROSE lipid model, 2.0% with published Cox Framingham lipid model: table 4). For women, the numbers correctly classified at the 20% threshold were similar with both PRIMROSE and published Cox-Framingham lipid models.

Supplementary and confirmatory analyses

The calibration of the *Cox-Framingham model, re-estimated to the UK general population*, in the SMI cohort was good for men, but somewhat poorer for women, with a degree of under-prediction (efigure 2). The discrimination of this re-estimated model was better than the published Framingham models. However, the PRIMROSE model still appeared somewhat superior to both (etables 3-5). Therefore, the superiority of the PRIMROSE model was not simply explained by the differences between US and UK populations.

The Cox-Framingham model, re-estimated to the SMI cohort, had good calibration (efigure2) and discrimination (etables 3-5) for people with SMI. However, the PRIMROSE lipid model (containing the additional SMI variables) still showed better discrimination. This suggests that PRIMROSE model performs better for people with SMI than models which only contain the standard variables usually found in risk tools such as the Framingham, including Systolic Blood Pressure, smoking, BMI and diabetes, even when these models are re-estimated in people with SMI.

Discussion

This is the first study to develop and assess the performance of CVD risk prediction tools in people with SMI. We derived two new CVD risk prediction models specific to people with SMI, the PRIMROSE *lipid* model and the PRIMROSE *BMI* model. These new PRIMROSE models included additional variables for psychiatric diagnosis, psychotropic medication at baseline and harmful use of alcohol, anti-depressants and social deprivation score. Both models performed better than the available published Framingham CVD risk models in predicting newly recorded CVD events in people with SMI, with better discrimination and calibration. The published Framingham models appeared to over-predict CVD risk in this population, especially among men.

This (over prediction) is likely to reflect differences between historical North American men and contemporary UK males. This phenomenon has been observed when European cardiovascular scores have been validated in non-SMI general practice populations, where the Framingham model over-predicts CVD events by 32% in UK men¹⁹ This over prediction has been seen at all levels of CVD risk including men with an elevated risk equivalent to that seen in SMI¹⁹

The PRIMROSE BMI model performed almost as well as the PRIMROSE lipid model supporting use of the BMI model as an alternative to the lipid model in situations where blood test results are not available. This might be important for people with SMI who are reluctant to provide blood samples or where blood results are not readily unavailable. While data regarding alcohol might be inaccurately reported to GPs, we only included recognized severe alcohol problems in the model. The other variables in the model are available to GPs including the medications they prescribe and the Townsend index for the individual's postcode.

Re-estimating the parameters of the published Framingham models to the UK general population and within the SMI cohort improved the discrimination and calibration of the Framingham models in men. However, the PRIMROSE models remained superior. This indicates that the additional variables in PRIMROSE, such as diagnoses and medication at baseline, are important to include in the prediction of CVD risk in SMI. This offers improved prediction over and above accounting for differences between i) the US and UK populations and ii) the general UK population and the population with SMI (assessed by evaluating the model re-estimated to the SMI cohort). Among women, the Framingham model re-estimated within the SMI cohort also improved discrimination and calibration. However, the Framingham model re-estimated to the UK general population appeared to have poorer calibration, with a degree of under-prediction. This may be explained by women with SMI having a particularly high CVD risk relative to women in the UK general population (greater than the differences among men with and without SMI) and, as such, application of the general population model for women to the SMI women relies on a degree of extrapolation.

Overall, the new PRIMROSE models offer the most consistent performance for people with SMI.

The PRIMROSE models advance our understanding of the best ways to assess and manage the increased cardiovascular risk in people with SMI. Previously we have not known if cardiovascular events are predicted by the same risk factors in people with SMI, including smoking, diabetes, hypertension, BMI and dyslipidaemia. Our results suggest that these traditional risk factors do mediate the association between SMI and high rates of CVD events. However, a model including specific SMI diagnosis, baseline prescription of antipsychotics and antidepressants, as well as harmful use of alcohol, offers improved prediction suggesting that these variables are associated with the increased risk in CVD independently of traditional CVD risk factors.

Caution is required as this does not necessarily imply a causal relationship between these factors and future CVD events. Rather, they are markers at baseline and the reason for their inclusion may be related to characteristics of individuals prescribed these medications in the UK including issues which might directly influence the decision to prescribe of these drugs. It would be incorrect to interpret our study results in the same way as epidemiological studies. Our study design was a risk score development study where we developed and validated the best risk score model for people with SMI. Variables may be retained in the final model despite the fact that their coefficients have 95% confidence intervals which cross

unity, with corresponding tests of significance outside the conventional 5% level. Furthermore, risk score development studies do not aim adjust for important confounders- they seek to derive the best prediction model.

Strengths and limitations

Our study benefited from large numbers of people with SMI who are representative of the UK SMI population²⁴ with longitudinal data regarding a range of CVD risk factors and long enough follow-up to provide sufficient data to develop and validate a CVD risk prediction model which included additional variables not traditionally included in such models. THIN data offer a longitudinal record of health variables in generalisable samples of patients in the UK. Finally we carefully assessed whether the superior performance of the PRIMROSE models could be attributed to the fact that existing CVD models have been developed in different populations. But our supplementary analyses supported the conclusion that the PRIMROSE models were better than the published Cox Framingham models even when the parameters were re-estimated for the general UK population and people with SMI.

Limitations included the fact that routine clinical data may be less complete in terms of predictor variables than cohorts designed for research. For instance levels of detected diabetes and hypertension may have been underestimated if people had not received screening. However it is reassuring that the proportion with diabetes in the PRIMROSE sample is intermediate between published meta-analyses of diabetes risk in people with treated and untreated psychoses. While UK primary care databases have been used for CVD risk score development studies, we acknowledge that the outcome of CVD is based on GP diagnostic codes rather than review of the whole medical record to establish CVD diagnoses. However GP records of coronary heart disease has been validated by medical records review in THIN²⁵.

One advantage of routine clinical data is that they are reflective of the data available to primary care and the setting in which the risk scores will be predominantly used clinically. Further, we also used multiple imputation techniques to impute the missing data which utilize the entire patient record, taking into account the temporal patterns of the records rather than relying solely on baseline measurement. We accept that imputation does not necessarily guarantee that problems with incomplete data are eliminated. Like many risk models, we only accounted for baseline variables. For many time-varying factors, exposure status may change during the follow-up period. This is particularly true for variables such as antipsychotic exposure (first and second generation) and BMI. However, using baseline variables reflects the real life clinical information available to a physician and a participant when they need information in order to make decisions on the likely risk of a CVD event for an individual over the next ten years.

In terms of generalisability, the PRIMROSE population contained a diverse range of people with SMI including bipolar disorder, schizophrenia and other non-organic psychoses of varying ages in primary care. Therefore baseline exposure to medication varied and a considerable number did not receive two consecutive antipsychotic prescriptions during their baseline 6 month period. However in line with other SMI studies², half the sample were smokers and almost one in ten had a record of severe alcohol problems. Although most

people with SMI in the UK are registered with a general practitioner, including those living in community supported accommodation, our study might have missed the very small proportion who are long stay hospital inpatients.

We did not compare our results directly with UK Q-Risk score since we did not have access to the parameters, but we did find that the PRIMROSE models were better than more simple Cox-Framingham models, re-estimated to the general UK population. This suggests that the PRIMROSE models perform better than the Q-Risk score in people with SMI, but this is an area worth further exploration.

Variables such as atrial fibrillation or renal disease are included in other risk scores¹⁷. However, the number of individuals in this cohort with these conditions was too small to provide reliable predictions. Sufficient data were not available to examine individual antipsychotic drugs and specific interactions for subgroups of the SMI populations. Furthermore, we were unable to include data regarding ethnicity since historically this variable has not been recorded systematically in primary care. However since the models already perform well, adding further variables might make them unnecessarily complicated without improving their performance.

Implications and Conclusions—The results suggest that the newly developed SMI-specific PRIMROSE CVD risk prediction models offer improved prediction of CVD in people with SMI, over and above existing CVD risk scores. The new models could therefore be a valuable tool in the prevention and management of CVD in people with SMI. We would recommend analysis of the cost effectiveness of the new models when used for making treatment decisions in routine clinical practice, and analyses to identify the optimum threshold for risk modification. Traditionally, the optimal threshold for risk modification in the general population has been set to 20%, and this is the risk threshold we assessed in this study. However, we do not know if this is the optimal risk threshold for risk modification in people with SMI. Furthermore, since we conducted this study there have been changes to international CVD risk management recommendations, precipitating wide ranging debate. The UK National Institute for Health and Care Excellence suggested commencing statins at a 10% threshold²⁶ while the American College of Cardiology/American Heart Association recommended intervening at 7.5% CVD risk²⁷. This reflects a general focus on using CVD risk scores to determine the threshold for statin initiation and moving away from “treating to target” with lipid levels. Our study did not aim to determine the best threshold for commencing statins, but we need to perform this work in populations with SMI to assess the effectiveness of various interventions in reducing CVD risk. Furthermore the use of single thresholds to make treatment decisions has been criticized as over-simplistic and neglecting to include patient preference for interventions, as well as the true risk/benefit for an individual when data are imperfect and derived at the population level.

In fact the PRIMROSE models, like other UK derived risk models, would lead to fewer people being treated with statins in the UK even if a conventional 20% threshold were employed. This is because US published CVD risk models like the cox-Framingham tend to over-predict CVD in the UK population. While the PRIMROSE models outperformed the Framingham models in the SMI population, it should be noted that the discrimination

statistics for the published Framingham models were still good. Therefore for the time being, clinicians should utilise existing CVD risk prediction tools in people with SMI, but preferably ones which are calibrated or re-estimated to their local general populations rather than the published Framingham equations from North American populations which may over-predict CVD risk in the other settings, irrespective of SMI status.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

All authors state they have no conflicts of interest relating to this manuscript.

SH and DO had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis

Funding: This paper summarises independent research funded by the National Institute for Health Research (NIHR) under its National Institute for Health Research's Programme Grants for Applied Research (PGfAR) Programme (Grant Reference Number RP-PG-0609-10156). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The funders had no role in study design, data collection, management and analysis, decision to publish, or preparation, review or approval of the manuscript

References

1. De Hert M, Dekker JM, Wood D, et al. Cardiovascular disease and diabetes in people with severe mental illness position statement from the European Psychiatric Association (EPA), supported by the European Association for the Study of Diabetes (EASD) and the European Society of Cardiology (ESC). *European Psychiatry*. 2009; 24:412–24. [PubMed: 19682863]
2. Osborn DPJ, Levy G, Nazareth I, Petersen I, Islam A, King M. Relative risk of cardiovascular and cancer mortality in people with severe mental illness from the United Kingdom's General Practice Research Database. *Archives of General Psychiatry*. 2007; 64:242–249. [PubMed: 17283292]
3. Osborn DPJ, Wright CA, Levy G, King MB, Deo R, Nazareth I. Relative risk of diabetes, dyslipidaemia, hypertension and the metabolic syndrome in people with severe mental illnesses. Systematic review and metaanalysis. *BMC Psychiatry*. 2008; 8:843.
4. Lieberman JA, Scott Stroup T, McEvoy JP, et al. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *New England Journal of Medicine*. 2005; 353(12):1209–1223. [PubMed: 16172203]
5. Mitchell AJ, Vancampfort D, A. et al. Is the prevalence of metabolic syndrome and metabolic abnormalities increased in early schizophrenia? A comparative meta-analysis of first episode, untreated and treated patients *Schizophrenia Bulletin*. 2013; 39:295–305.
6. De Hert M, Correll CU, Bobes J, et al. Physical illness in patients with severe mental disorders. I. Prevalence, impact of medications and disparities in health care. *World Psychiatry*. 2011; 10:52–77. [PubMed: 21379357]
7. Tiihonen J, Lönnqvist J, Wahlbeck K, et al. 11-year follow-up of mortality in patients with schizophrenia: a population-based cohort study (FIN11 study). *Lancet*. 2009; 374:620–27. [PubMed: 19595447]
8. Crump C, Winkleby MA, Sundquist K, et al. Comorbidities and mortality in persons with schizophrenia: a Swedish national cohort study. *Am J Psychiatry*. 2013; 170:324–33. [PubMed: 23318474]
9. De Hert M, Correll C, Cohen D. Do antipsychotic medications reduce or increase mortality in schizophrenia? A critical appraisal of the FIN-11 Study. *Schizophrenia Research*. 2010; 117(1):68–74. [PubMed: 20060684]

10. D'Agostino RB Sr, Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation*. 2008; 117(6):743–53. [PubMed: 18212285]
11. National Institute for Health and Clinical Excellence. NICE clinical guideline 82. National Institute for Health and Clinical Excellence; London: 2009. Schizophrenia: core interventions in the treatment and management of schizophrenia in adults in primary and secondary care. <http://www.nice.org.uk/CG082>
12. [Accessed 18.01.2014] The Health Improvement Network website. www.THIN-UK.com
13. Dave S, Petersen I. Creating medical and drug code lists to identify cases in primary care databases. *Pharmacoepidem Drug Safe*. 2009; 18:704–707.
14. Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. *Inform Prim Care*. 2011; 19:251–255. [PubMed: 22828580]
15. Lis Y, Mann RD. The VAMP Research multi-purpose database in the UK. *J Clin Epidemiol*. 1995; 48:431–43. [PubMed: 7897464]
16. Horsfall L, Walters K, Petersen I. Identifying periods of acceptable computer usage in primary care research databases. *Pharmacoepidem Drug Saf*. 2013; 22(1):64–69.
17. Fardet L, Petersen I, Nazareth I. Risk of cardiovascular events in people prescribed glucocorticoids with iatrogenic Cushing's syndrome: cohort study. *BMJ*. 2012; 345:e4928. [PubMed: 22846415]
18. Horsfall LJ, Nazareth I, Petersen I. Cardiovascular events as a function of serum bilirubin levels in a large, statin-treated cohort. *Circulation*. 2012; 126:2556–2564. [PubMed: 23110860]
19. Collins GS, Altman DG. An independent external validation and evaluation of QRISK cardiovascular risk prediction: a prospective open cohort study. *BMJ*. 2009; 339:b2584. [PubMed: 19584409]
20. [accessed 18.01.2014] Quality and outcomes framework guidance for GMS contract 2011/12. http://www.nhsemployers.org/SiteCollectionDocuments/QOFguidanceGMScontract_2011_12_FL%2013042011.pdf
21. Nazareth I, King M, Haines A, Rangel L, Myers S. Accuracy of diagnosis of psychosis on general practice computer system. *BMJ*. 1993; 307(6895):32–4. [PubMed: 8343670]
22. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med*. 2011; 30(10):1105–17. [PubMed: 21484848]
23. Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. *Stat Med*. 2004; 23:723–748. [PubMed: 14981672]
24. Hardoon S, Hayes JF, Blackburn R, Petersen I, Walters K, Nazareth I, Osborn DPJ. Recording of Severe Mental Illness in United Kingdom Primary Care, 2000-2010. *PLoS One*. 2013 DOI: 10.1371/journal.pone.0082365.
25. Garcia Rodriguez LA, Tacconelli S, Patrignani P. Role of dose potency in the prediction of risk of myocardial infarction associated with nonsteroidal anti-inflammatory drugs in the general population. *J Am Coll Cardiol*. Nov 11; 2008 52(20):1628–36. [PubMed: 18992652]
26. Majeed A. Statins for primary prevention of cardiovascular disease. *BMJ*. 2014; 348:g3491. [PubMed: 24868092]
27. Hayward RA. Moneyball, Gambling, and the New Cholesterol Guidelines. *Circ Outcomes*. 114.000876.

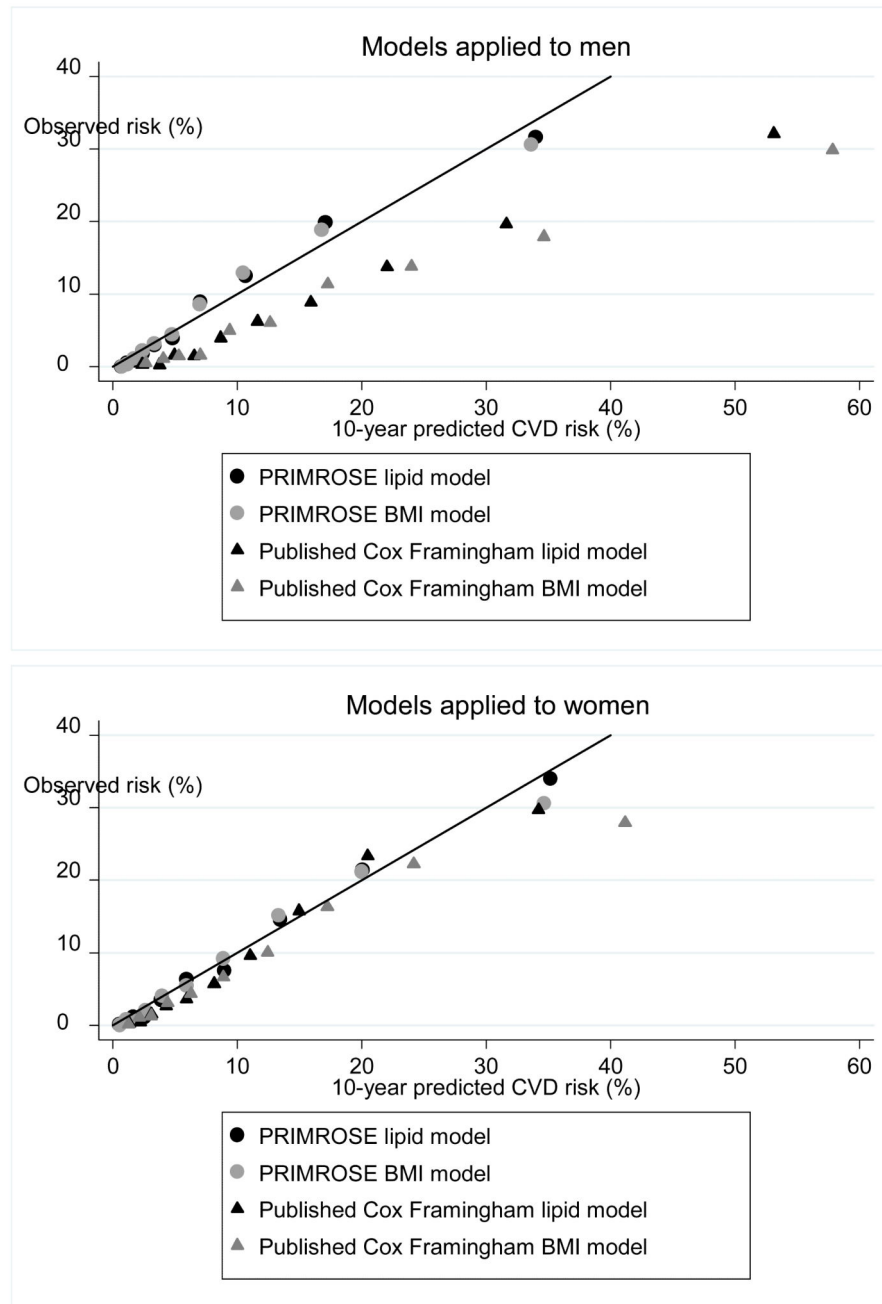


Figure 1.
calibration plots: PRIMROSE Models and published Framingham models

Table 1
Characteristics of eligible patients with SMI Data before imputation

	Men (n=18417, 47%)			Women (n=20407)			All (n=38824)		
	Obs	Mean	sd	Obs	Mean	sd	Obs	Mean	sd
Baseline Age, years	18,417	46.2	14.0	20,407	52.4	16.4	38,824	49.5	15.6
Baseline Systolic blood pressure, mmHg	5,586	132.5	17.6	8,977	131.2	20.3	14,563	131.7	19.3
Baseline Weight, kg	3,478	84.8	18.0	5,105	73.5	18.2	8,583	78.1	19.0
Baseline Height, m	11,927	1.76	0.08	13,515	1.61	0.07	25,442	1.68	0.10
Baseline BMI, kg/m²	2,921	27.74	5.48	4,200	28.37	6.75	7,121	28.11	6.27
Baseline Total cholesterol, mmol/L	1,737	5.37	1.13	1,971	5.62	1.12	3,708	5.51	1.13
Baseline HDL cholesterol, mmol/L	1,127	1.22	0.37	1,199	1.46	0.44	2,326	1.34	0.42
	Obs	Freq	%	Obs	Freq	%	Obs	Freq	%
Predominant smoking history	17,573			19,592			37,165		
Never		5,030	29		9,219	47		14,249	38
Ex		2,317	13		2,376	12		4,693	13
Current		10,226	58		7,997	41		18,223	49
History of diabetes at baseline	18,417	597	3	20,407	759	4	38,824	1,356	3
Use of anti-hypertensive drugs at baseline	18,417	1,543	8	20,407	2,505	12	38,824	4,048	10
Townsend score of social deprivation (quintiles)	18,417			20,407			38,824		
1 (=least deprived)		2,433	13		3,588	18		6,021	16
2		2,849	15		3,750	18		6,599	17
3		3,688	20		4,279	21		7,967	21
4		4,512	25		4,740	23		9,252	24
5 (=most deprived)		4,935	27		4,050	20		8,985	23
Severe mental illness diagnosis	18,417			20,407			38,824		
Schizophrenia		7,606	41		5,626	28		13,232	34
Bipolar disorder		4,005	22		6,093	30		10,098	26
Other psychosis		4,967	27		6,238	31		11,205	29
Unknown - on Severe mental illness register		1,839	10		2,450	12		4,289	11
Use of anti-depressants at baseline	18,417	5,679	31	20,407	8,439	41	38,824	14,118	36
Use of second generation antipsychotics at baseline	18,417	4,184	23	20,407	3,760	18	38,824	7,944	20
Use of first generation antipsychotics at baseline	18,417	3,864	21	20,407	4,769	23	38,824	8,633	22
Use of lithium at baseline	18,417	1,626	9	20,407	2,499	12	38,824	4,125	11
History of very heavy drinking/ alcohol problem at baseline	18,417	2,513	14	20,407	1,111	5	38,824	3,624	9

Table 2
Associations of predictors with new onset CVD after imputation. Results from Cox regression, including age and gender

	HR	95% CI	p
Gender (female v male)	0.69	0.64 to 0.75	<0.001
Age, years (per log year increase)	43.06	35.95 to 51.57	<0.001
Total cholesterol, mmol/L (per 1 unit increase)	1.13	1.05 to 1.22	0.002
HDL cholesterol, mmol/L (per 1 unit increase)	0.41	0.23 to 0.73	0.007
SBP, mmHg (per 10 unit increase)	1.10	1.06 to 1.13	<0.001
SBP if no baseline anti-hypertensive use, mmHg (per 10 unit increase)	1.10	1.06 to 1.15	<0.001
SBP if baseline anti-hypertensive use, mmHg (per 10 unit increase)	1.00	0.95 to 1.05	0.9
Weight, kg (per 10 unit increase)	1.02	0.98 to 1.05	0.4
Height, m (per 10 cm increase)	0.86	0.80 to 0.93	<0.001
History of Diabetes (yes v no)	1.75	1.49 to 2.05	<0.001
Predominant smoking history			<0.001
Never	1		
Ex	1.10	0.98 to 1.23	
Current	1.47	1.34 to 1.61	
Calendar year at baseline (per 1 year increase)	0.94	0.92 to 0.95	<0.001
Townsend quintile of deprivation			<0.001
1 (= least deprived)	1		
2	1.15	1.01 to 1.32	
3	1.26	1.10 to 1.44	
4	1.34	1.16 to 1.55	
5 (= most deprived)	1.45	1.24 to 1.69	
Use of anti-depressants at baseline (yes v no)	1.27	1.16 to 1.39	<0.001
History of heavy drinking (yes v no)	1.52	1.29 to 1.80	<0.001
Severe mental illness diagnosis			0.003
Schizophrenia	1		
Bipolar disorder	1.08	0.97 to 1.21	
Other psychosis	1.20	1.08 to 1.34	
Unknown - on severe mental illness register	1.01	0.87 to 1.17	
Use of second generation antipsychotics at baseline	0.96	0.83 to 1.09	0.5
Use of first generation antipsychotics at baseline	1.18	1.07 to 1.31	0.001
Use of lithium at baseline (yes v no)	1.08	0.96 to 1.21	0.2

HR: Hazard ratio, SBP: systolic blood pressure, HDL: high density lipoprotein

Table 3
Final PRIMROSE models coefficients, after backwards eliminations

	PRIMROSE Lipid model				PRIMROSE BMI model			
	β Coefficient	HR	95% CI	p	β Coefficient	HR	95% CI	p
Gender (female v male)	-0.1795	0.84	0.72 to 0.96	0.02	-0.49376	0.61	0.53 to 0.7	<0.001
Age, years (per log year increase)	3.78124	43.87	34.61 to 55.61	<0.001	3.50943	33.43	26.91 to 41.52	<0.001
SBP if no baseline anti-hypertensive use, mmHg (per 10 unit increase)	0.07651	1.08	1.04 to 1.12	<0.001	0.0893	1.09	1.05 to 1.14	<0.001
SBP if baseline anti-hypertensive use, mmHg (per 10 unit increase)	-0.00316	1	0.95 to 1.05	0.9	0.0005	1	0.95 to 1.05	>0.9
Total cholesterol, mmol/L (per unit increase)	0.11763	1.12	1.03 to 1.22	0.009				
HDL cholesterol, mmol/L (per unit increase)	-0.8183	0.44	0.24 to 0.8	0.01				
Weight, kg (per 10 unit increase)					0.0068	1.01	0.97 to 1.05	0.7
Height, cm (per 10 unit increase)					-0.12413	0.88	0.81 to 0.96	0.004
History of Diabetes (yes v no)	0.37734	1.46	1.2 to 1.77	<0.001	0.44971	1.57	1.33 to 1.85	<0.001
Predominant smoking history				<0.001				<0.001
Never	0	1			0	1		
Ex	0.01639	1.02	0.9 to 1.15		0.0738	1.08	0.96 to 1.21	
Current	0.29659	1.35	1.2 to 1.51		0.38081	1.46	1.32 to 1.62	
Calendar year at baseline (per 1 year increase)	-0.07043	0.93	0.92 to 0.95	<0.001	-0.07524	0.93	0.91 to 0.94	<0.001
Use of anti-depressants at baseline (yes v no)	0.2104	1.23	1.12 to 1.36	<0.001	0.21846	1.24	1.13 to 1.37	<0.001
History of heavy drinking (yes v no)	0.41392	1.51	1.25 to 1.83	<0.001	0.30721	1.36	1.15 to 1.61	<0.001
Townsend quintile of deprivation				0.06				0.01
1 (= least deprived)	0	1			0	1		
2	0.10963	1.12	0.96 to 1.29		0.10919	1.12	0.97 to 1.28	
3	0.16388	1.18	1.03 to 1.35		0.18412	1.2	1.06 to 1.37	
4	0.1828	1.2	1.03 to 1.4		0.20238	1.22	1.06 to 1.41	
5 (= most deprived)	0.22126	1.25	1.06 to 1.46		0.24762	1.28	1.1 to 1.49	
Severe mental illness diagnosis				0.001				0.002
Schizophrenia	0	1			0	1		
Bipolar disorder	0.11177	1.12	0.99 to 1.26		0.0978	1.1	0.98 to 1.24	
Other psychosis	0.21004	1.23	1.1 to 1.39		0.19063	1.21	1.08 to 1.35	

	PRIMROSE Lipid model				PRIMROSE BMI model			
	β Coefficient	HR	95% CI	p	β Coefficient	HR	95% CI	p
Unknown - on severe mental illness register	0.01526	1.02	0.86 to 1.19		-0.01138	0.99	0.85 to 1.15	
Use of second generation antipsychotics at baseline (yes v no)	0.12121	1.13	0.96 to 1.32	0.1	0.17662	1.19	1.02 to 1.4	0.03
Use of first generation antipsychotics at baseline (yes v no)					0.1205	1.13	1.02 to 1.25	0.02
So(10)	0.968011				0.951285			

So(10) is the predicted baseline survival at ten years for the model

Table 4
Performance of PRIMROSE and Published Cox Framingham models in predicting CVD events in the SMI cohort.

	PRIMROSE BMI	PRIMROSE lipid	Cox Framingham published BMI	Cox Framingham published lipid
Men				
C statistic	0.78 (0.74-0.83)	0.80 (0.76-0.83)	0.76 (0.71 - 0.80)	0.78 (0.75 - 0.82)
D statistic	1.84 (1.73 - 1.96)	1.92 (1.80 - 2.03)	1.65 (1.54 - 1.76)	1.74 (1.63 - 1.86)
N (%) of high risk patients developing CVD	390/2096 (18.6)	410/2137 (19.2)	679/5528 (12.3)	666/5028 (13.3)
N (%) of low risk patients developing CVD	550/16321 (3.4)	529/16280 (3.3)	260/12889 (2.0)	273/13389 (2.0)
Women				
C statistic	0.78 (0.74 - 0.83)	0.80 (0.76 - 0.83)	0.76 (0.72 - 0.80)	0.77 (0.73 - 0.81)
D statistic	1.80 (1.70 - 1.90)	1.87 (1.76 - 1.98)	1.52 (1.43 - 1.61)	1.58 (1.48 - 1.68)
N (%) of high risk patients developing CVD	531/2989 (17.8)	570/2991 (19.1)	618/4082 (15.2)	526/3173 (16.6)
N (%) of low risk patients developing CVD	641/17418 (3.7)	602/17416 (3.5)	554/16325 (3.4)	646/17234 (3.8)

From 10-fold internal cross-validation

High risk patients defined as those scoring above 20% on the risk score

C and D statistics are the median (6th) estimates with bootstrapped 95% confidence intervals from our ten validation datasets.

Estimates by validation test sets are presented in tables 3-4