medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Cardiovascular risk prediction using metabolomic biomarkers and polygenic risk scores: A cohort study and modelling analyses

Scott C. Ritchie[1–6,*], Xilin Jiang[1,3–7], Lisa Pennells[3,4], Yu Xu[1,3,4], Claire Coffey[1,3,4], Yang Liu[1–4], Praveen Surendran[3,5,8], Savita Karthikeyan[3], Samuel A. Lambert[1,3–6], John Danesh[3–6,9,10], Adam S. Butterworth[3-6,9], Angela Wood[3-6,9,11], Stephen Kaptoge[3,4], Emanuele Di Angelantonio[3-6,9,12], Michael Inouye[1-6,13*]

Affiliations

[1]Cambridge Baker Systems Genomics Initiative, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
[2]Cambridge Baker Systems Genomics Initiative, Baker Heart & Diabetes Institute, Melbourne, Victoria, Australia
[3]British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
[4]Victor Philip Dahdaleh Heart and Lung Research Institute, University of Cambridge, Cambridge, UK
[5]British Heart Foundation Centre of Research Excellence, University of Cambridge, Cambridge, UK
[6]Health Data Research UK Cambridge, Wellcome Genome Campus and University of Cambridge, Cambridge, UK
[7]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, US
[8]Rutherford Fund Fellow, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
[9]National Institute for Health and Care Research Blood and Transplant Research Unit in Donor Health and Behaviour, University of Cambridge, Cambridge, UK
[10]Department of Human Genetics, Wellcome Sanger Institute, Hinxton, UK
[11]Cambridge Centre of Artificial Intelligence in Medicine, University of Cambridge, Cambridge, UK
[12]Health Data Science Research Centre, Human Technopole, Milan, Italy
[13]Department of Clinical Pathology, University of Melbourne, Parkville, Victoria, Australia

*Corresponding authors: sr827@medschl.cam.ac.uk (S.C.R), mi336@cam.ac.uk (M.I.)

Abstract

Metabolomic platforms using nuclear magnetic resonance (NMR) spectroscopy can now rapidly quantify many circulating metabolites which are potential biomarkers of cardiovascular disease (CVD). Here, we analyse ~170,000 UK Biobank participants (5,096 incident CVD cases) without a history of CVD and not on lipid-lowering treatments to evaluate the potential for improving 10-year CVD risk prediction using NMR biomarkers in addition to conventional risk factors and polygenic risk scores (PRSs). Using machine learning, we developed sex-specific NMR scores for coronary heart disease (CHD) and ischaemic stroke, then estimated their incremental improvement of 10-year CVD risk prediction when added to guideline-recommended risk prediction models (i.e., SCORE2) with and without PRSs. The risk discrimination provided by SCORE2 (Harrell's C-index = 0.718) was similarly improved by addition of NMR scores (ΔC-index 0.011; 0.009, 0.014) and PRSs (ΔC-index 0.009; 95% CI: 0.007, 0.012), which offered largely orthogonal information. Addition of both NMR scores and PRSs yielded the largest improvement in C-index over SCORE2, from 0.718 to 0.737 (ΔC-index 0.019; 95% CI: 0.016, 0.022). Concomitant improvements in risk stratification were observed in categorical net reclassification index when using guidelines-recommended risk categorisation, with net case reclassification of 13.04% (95% CI: 11.67%,

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

14.41%) when adding both NMR scores and PRSs to SCORE2. Using population modelling, we estimated that targeted risk-reclassification with NMR scores and PRSs together could increase the number of CVD events prevented per 100,000 screened from 201 to 370 ($\Delta CVD_{prevented}$: 170; 95% CI: 158, 182) while essentially maintaining the number of statins prescribed per CVD event prevented. Overall, we show combining NMR scores and PRSs with SCORE2 moderately enhances prediction of first-onset CVD, and could have substantial population health benefit if applied at scale.

Introduction

Circulating biomarkers play a central role in cardiovascular disease (CVD) risk scores recommended by clinical guidelines to identify high-risk individuals for CVD prevention[1–3]. Total cholesterol and high-density lipoprotein (HDL) cholesterol are routinely measured and used alongside demographic and lifestyle risk factors to assess 10-year risk of CVD using risk scores such as SCORE2[4]. Efforts to improve CVD risk prediction models have considered additional circulating biomarkers[5], such as C-reactive protein (CRP)[6,7], as well as incorporating polygenic risk scores (PRSs) to account for genetic predisposition[8–10]. While PRSs have shown potential to enhance CVD risk screening[11–14], addition of individual CVD biomarkers have thus far shown limited overall incremental benefits[15–17].

High-throughput nuclear magnetic resonance (NMR) spectroscopy has enabled rapid and simultaneous quantification of several biomarkers from a single human blood plasma sample[18,19]. These include cholesterols and other lipids in lipoprotein sub-fractions, fatty acids, ketone bodies, amino acids, glycolysis metabolites and glycoprotein acetyls (GlycA)[20,21]. NMR metabolic biomarker data has been quantified in numerous cohorts over the last decade, helping derive new insights into the genetic determinants, molecular pathogenesis, and epidemiology of CVD[22].

Several studies have investigated the utility of biomarkers combinations from NMR platforms to improve prediction of CVD[23–26]; however, they have focused on multi-disease prediction, used outdated clinical risk prediction scores, and have not investigated improvements relative to clinically relevant guideline-recommended risk thresholds.

Here, we utilize NMR biomarker data in UK Biobank to assess whether NMR biomarkers, in isolation or combination (*i.e.*, NMR scores), can improve 10-year CVD risk prediction when added to the SCORE2 risk model, which is recommended by the European Society of Cardiology (ESC) 2021 guidelines for CVD prevention[3]. We further assess whether incremental improvements in CVD risk prediction are meaningful at ESC 2021 recommended risk thresholds for treatment consideration[3]. In addition we compared the improvement in risk prediction provided by NMR scores to that provided by PRS[11] and also assessed the PRSs and NMR scores combined. Finally, we modelled the potential public health benefits for CVD prevention if applied to the UK primary care population according to the ESC 2021 guidelines for statin initiation. A schematic of the overall study is given in **Figure 1**.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Results

Characteristics of study participants

Of the 502,207 participants enrolled in UK Biobank consenting to electronic health record linkage, 168,517 participants met the inclusion criteria for this study (**Figure S1**), namely: participants who were eligible for 10-year CVD risk assessment according to the ESC 2021 guidelines for CVD prevention[3], had plasma NMR spectroscopy biomarker data available (with <5% missingness), had complete data on risk factors needed to compute SCORE2, were not taking lipid lowering medications, and had data on PRSs (**Methods**). Participants were eligible for 10-year CVD risk assessment at baseline if they were 40 to 69 years of age and were apparently healthy[3]; *i.e.* with no prior history of established atherosclerotic cardiovascular disease, diabetes mellitus, chronic kidney disease, or familial hypercholesterolemia. During the 1,641,935 person-years at risk (median [5th, 95th percentile] follow-up of 10.0 [8.5–10.0] years), 5,096 CVD cases were recorded. Baseline cohort characteristics are detailed in **Table 1**.

Incremental CVD risk discrimination with individual biomarkers

Improvements in CVD risk discrimination were assessed by differences in C-index (ΔC-index) beyond SCORE2 alone for each of the 249 NMR biomarkers (**Table S1**) and 28 clinical chemistry biomarkers (**Table S2**). The ΔC-index was assessed for each biomarker separately using sex-stratified Cox proportional hazards models for 10-year CVD risk with the biomarker as a predictor and the SCORE2 linear predictor as an offset term (**Methods**). The sex-stratified C-index for SCORE2 alone was 0.718 (95% confidence interval [CI]: 0.711, 0.724) in the analysis sample of 168,517 participants (5,096 CVD cases).

Based on a false discovery rate (FDR) adjusted bootstrap P-value < 0.05 (**Methods**) we observed statistically significant improvement in C-index for 35 of the 277 biomarkers (**Table S3A**). Improvements in sex-stratified C-index over SCORE2 (ΔC-index) were modest (**Figure 2A**). The largest ΔC-index observed for any biomarker was with addition of cystatin-C measured by clinical biochemistry assay (**Figure 2A**), with ΔC-index of 0.006 (95% CI: 0.004, 0.008; **Table S3A**). The largest ΔC-index observed for any of the 249 NMR biomarkers was with addition of albumin (**Figure 2A**), with ΔC-index of 0.005 (95% CI: 0.003, 0.006; **Table S3A**). Results were similar when analysing males and females separately, but with reduced power to detect statistically significant differences in ΔC-index (**Table S3B,C**).

NMR biomarker scores

Sex-specific NMR biomarker scores for CHD (4,054 cases; **Table S4A**) and ischaemic stroke (1,280 cases; **Table S4B**) were trained and tested in the 168,517 participants, then later combined (see next section) for 10-year CVD risk prediction (**Figure S2**; **Methods**). NMR biomarker scores were trained for CHD and stroke separately and later combined, as we previously found that combining PRSs for CHD and stroke led to improved prediction of 10-year CVD risk over a single PRS trained for CVD, due to its heterogeneity[11]. NMR scores were also trained in males and females separately to capture sex-specific differences in their concentrations[27] and well-known differences in baseline survival between males and females[3]. NMR scores were trained using elastic-net penalised Cox proportional hazards regression[28,29] in nested cross validation using the 106 non-derived NMR biomarkers as candidate predictors (**Figure S3**; **Methods**). The per-biomarker weights for computing the consensus optimal NMR scores after model training are given in **Table S5**.

All four consensus optimal NMR scores included all 106 non-derived biomarkers (**Figure S3C**, **Table S5**). The biomarker with the strongest contribution to the CHD NMR scores was GlycA, which explained 12.8%

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

of the variance in the CHD NMR score in males, and 11.4% of the variance in the CHD NMR score in females (**Table S5**). The biomarker with the strongest contribution to the ischaemic stroke NMR scores was albumin, which explained 15.7% of the variance in the ischaemic stroke NMR score in males, and 27.1% of the variance in the ischaemic stroke NMR score in females (**Table S5**).

To avoid overestimation of prediction performance in downstream analyses, we used for each NMR score the aggregate of their predicted values across cross-validation test partitions (**Figure S2A**). Sex-specific pairwise correlations between SCORE2, the NMR scores, and PRSs are shown in **Figure S4**. Predicted NMR scores and PRS were statistically significant independent predictors of 10-year CVD risk when fitting sex-stratified Cox proportional hazards regression with SCORE2 as an offset term (**Figure 3A**; **Methods**). Results were similar when fitting models with SCORE2 risk factors as independent predictors variables (**Figure S6**; **Table S6B**; **Methods**). Results were also similar when analysing males and females separately (**Figure S5A, Figure S6, Table S6**).

Incremental value of NMR biomarker scores and PRSs to 10-year CVD risk prediction

We assessed and compared three models to SCORE2 for 10-year CVD risk prediction in the 168,517 participants: (1) SCORE2 + CHD NMR score + ischaemic stroke NMR score, (2) SCORE2 + CHD PRS + ischaemic stroke PRS, and (3) SCORE2 + NMR scores + PRSs (**Methods**). NMR scores and/or PRSs were combined with SCORE2 as part of the cross-validation procedure described above (**Figure S2**). Per-score weightings for centring adding NMR scores and/or PRSs to SCORE2 in new samples are given in **Table S7**.

Incremental discrimination for 10-year CVD risk for each model was assessed by differences in sex-stratified C-index from SCORE2 alone (**Methods**). We observed statistically significant improvement in ΔC-index for all three models (**Figure 3B**, **Table S8**). The ΔC-index with addition of NMR scores was 0.011 (95% CI: 0.009, 0.014; **Table S8**), almost double the ΔC-index observed for any single biomarker alone (**Table S3A**). This was also similar to the ΔC-index observed with addition of PRSs, which was 0.009 (95% CI: 0.007, 0.012; **Table S8**). Improvement in risk discrimination was greatest when adding both NMR scores and PRSs to SCORE2 (**Figure 3B**), with ΔC-index 0.019 (95% CI: 0.016, 0.022; **Table S8**)—an 8.8% gain in C-index relative to SCORE2 alone—for a total absolute C-index of 0.737. Improvements in ΔC-index were greater in males than in females for all models, and the differences were most pronounced for models incorporating PRSs (**Figure S5B, Table S8**).

Incremental value in risk stratification using ESC 2021 risk thresholds for treatment consideration

Next, we assessed whether incremental improvements in CVD risk prediction were meaningful at clinically relevant risk thresholds. For each model, we calculated absolute 10-year CVD risk using formulae calibrated to the UK population[4] (**Methods**) and stratified participants into categories of low risk, medium risk, and high risk (**Table S9**) using ESC 2021 recommended risk thresholds for treatment consideration[3] (**Methods**). Distributions of predicted absolute risk are compared in **Figure S7**. Improvements in risk stratification over SCORE2 alone were then assessed using categorical net reclassification index (NRI) (**Figure 3C, Table S10**).

Statistically significant improvement in risk stratification over SCORE2 among incident CVD cases was observed for all three alternative models tested (**Figure 3C**, **Table S10A**). Improvements in case classification from NMR scores were more than twice as strong as those from PRSs. We observed a net case reclassification rate of 10.71% (95% CI: 9.33%, 12.08%) with addition of NMR scores, and 4.21%

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

(95% CI: 3.08%, 5.34%; **Table S10A**) with addition of PRSs. Improvements in case classification were strongest with addition of both NMR scores and PRSs, with a net case reclassification rate of 13.04% (95% CI: 11.67%, 14.41%; **Table S10A**). Results were similar in sex-specific analyses (**Figure S5C**, **Table S10B**).

A modest, but statistically significant, inappropriate reclassification for non-cases was also observed for all three alternative models (**Figure 3C**, **Table S10B**). The net reclassification rate for non-cases was −2.51% (95% CI: −2.69%, −2.34%) with addition of NMR scores, −0.58% (95% CI: −0.74%, −0.41%) with addition of PRSs, and −2.90% (95% CI: −3.10%, −2.70%) with addition of both NMR scores and PRSs (**Table S10B**). In sex-specific analyses greater inappropriate reclassification of non-cases was observed in females than in males (**Figure S5C**, **Table S10B**).

Incremental value for CVD prevention with population-wide screening

Next, we estimated the incremental benefits to CVD prevention if applied at scale to the UK population. We simulated a hypothetical population of 100,000 adults 40–69 years of age representative of the general UK population using the age- and sex- structure of the UK population[30] and previously published 10-year CVD incidence rates amongst CVD- and statin- free primary care patients[11] (**Figure S8**, **Methods**). In total, the simulated population comprised 49,156 males (4,391 incident CVD cases) and 50,844 females (2,245 incident CVD cases) (**Table S11**).

To model the benefits of population-wide screening with each model, we stratified the simulated population into the low-, medium-, and high- CVD risk groups based on the proportions allocated to each category in UK Biobank by SCORE2 alone and the three alternative models adding NMR scores and/or PRSs (**Figure S9A**, **Methods**). We modelled statin initiation in the high-risk group, who based on their risk thresholds would be recommended for risk factor treatment by the ESC 2021 guidelines for CVD prevention[3]. The impact of statin initiation was modelled as preventing one in five simulated incident CVD events; assuming a 20% reduction in 10-year CVD risk[31].

Incremental improvements in CVD prevention for each alternative model were assessed by differences from SCORE2 alone in (1) the number of people classified as high risk ($\Delta N_{high\text{-}risk}$); (2) the number of future CVD cases amongst the high-risk group ($\Delta CVD_{high\text{-}risk}$); (3) the number of future CVD events expected to be prevented by initiation of statins in the high-risk group ($\Delta CVD_{prevented}$); (4) the number needed to screen to prevent one CVD event (ΔNNS); and the number of statins prescribed per CVD event prevented (ΔNNT) (**Methods**).

Consistent with the categorical NRI analyses above, we observed statistically significant improvements in CVD prevention with addition of NMR scores and/or PRSs (**Figure 4A**). For all three models we observed a statistically significant increase in the $\Delta N_{high\text{-}risk}$, $\Delta CVD_{high\text{-}risk}$, and $\Delta CVD_{prevented}$, along with a statistically significant decrease in the ΔNNS (**Figure 4A**, **Table S12A**). The number of events prevented increased from 201 per 100,000 screened with SCORE2 alone, to 309 with addition of NMR scores ($\Delta CVD_{prevented}$: 108; 95% CI: 96, 120), to 246 with addition of PRSs ($\Delta CVD_{prevented}$: 108; 95% CI: 35, 56), and to 339 with addition of both NMR scores and PRSs ($\Delta CVD_{prevented}$: 139; 95% CI: 125, 153) (**Table S12A**). Importantly, our modelling indicated no statistically significant change in NNT, the number of statins prescribed per CVD event prevented was constant at 23 (**Table S12A**). Sex-specific analyses were similar (**Figure S10A**, **Table S12A**).

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Incremental value for CVD prevention with targeted screening

Finally, we estimated the incremental benefits to CVD prevention if using NMR scores and/or PRSs for targeted risk-reclassification of those classified as medium-risk with SCORE2, for whom the ESC 2021 guidelines suggest considering, but do not explicitly recommend, risk factor treatment[3]. When applying SCORE2 alone to the simulated population (**Figure S9A, Methods**), there were 36,005 people predicted to be classified to the medium CVD risk category, which included 3,728 incident CVD cases (56% events).

We re-stratified the medium-risk population based on proportions re-stratified into each risk category in UK Biobank by the three alternative models adding NMR scores and/or PRSs (**Figure S9B**, **Methods**). Incremental improvements in CVD prevention from targeted screening with NMR scores and/or PRSs were consistent with and stronger than those observed from population-wide screening (**Figure 4B, Table S12B**). The number of events prevented increased from 201 per 100,000 screened with SCORE2 alone, to 336 with additional targeted screening with NMR scores ($\Delta CVD_{prevented}$: 136; 95% CI: 125, 147), to 277 with additional targeted screening with PRSs ($\Delta CVD_{prevented}$: 77; 95% CI: 68, 86), and to 370 with additional targeted screening with NMR scores and PRSs combined ($\Delta CVD_{prevented}$: 170; 95% CI: 158, 182) (**Table S12B**). A small but not statistically significant increase in the number of statins prescribed per CVD event was observed for all three models (ΔNNT: 1; **Table S12B**). Improvements in CVD prevention were statistically significant in both males and females, with no statistically significant change in the number of statins prescribed per CVD event prevented (**Figure S10**, **Table S12B**).

Discussion

Determining the added value of biomarkers beyond total and HDL cholesterol for 10-year CVD risk prediction is an area of interest for enhancing CVD prevention[3]. Here, we investigated whether 10-year CVD risk prediction in UK Biobank participants eligible for screening could be improved, in comparison to the currently recommended SCORE2[3,4].

We found statistically significant improvements in 10-year CVD risk prediction from 35 of 277 biomarkers quantified either individually by clinical chemistry assays or simultaneously by plasma NMR spectroscopy. Although statistically significant due to the large sample size, the magnitude of these incremental improvements was modest. Combining NMR biomarkers into NMR scores almost doubled the gain in observed predictive performance (ΔC-index) as compared to any single NMR biomarker. NMR biomarker scores and PRSs offered largely orthogonal information and increased SCORE2 C-index to similar degrees.

Apart from the cholesterol component of CVD risk in SCORE2[4], the biomarkers yielding the strongest improvement in 10-year CVD risk prediction were related to inflammation; a well-studied target in CVD prediction and prevention research[32–34]. Cystatin-C is a biomarker of renal function and cardiovascular disease also known to be associated with increased inflammation[35,36]. Albumin was the biomarker yielding the second strongest improvement over SCORE2 and was also the strongest contributor to the ischaemic stroke NMR score. Hypoalbuminemia has been associated with increased risk of stroke in numerous epidemiological studies[37–39] and is also a biomarker of inflammation[40,41]. The strongest contributor to the CHD NMR score was GlycA, an NMR signal quantifying the levels of multiple proteins with key roles in inflammation[21,42] and a stronger biomarker of chronic inflammation than C-reactive protein[43], which has been associated with CVD risk in multiple studies[20,44,45].

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Risk prediction models including NMR scores and/or PRSs also improved risk stratification of future CVD events when using risk thresholds recommended for clinical decision making by the ESC 2021 guidelines for CVD prevention. NMR scores improved net case reclassification to a greater extent than PRSs (10.71% vs 4.21%, respectively); however, when combined, NMR scores and PRSs improved net case reclassification by 13.04%. These results highlight the complementary nature of the information capture by PRSs and NMR scores. While PRSs capture the lifetime risks due to genetics[8–10], NMR scores capture part of the dynamic component of risk conferred by lifestyle and environment[23], which act on that genetic background.

When modelling the potential benefits for CVD prevention in the wider UK population eligible for 10-year CVD risk screening, we found adding NMR scores and/or PRSs to SCORE2 significantly increased the those would be recommended for statin initiation (following the ESC 2021 guidelines for risk factor treatment for CVD prevention[3]) and who would subsequently experience a CVD event. Importantly, the number of statins prescribed per CVD event prevented would stay constant. We estimated that adding NMR scores to SCORE2 would increase the number of CVD events prevented from 201 to 309 (per 100,000); adding PRSs to SCORE2 would increase the number of CVD events prevented to 246; and adding both NMR scores and PRSs to SCORE2 would increase the number of CVD events prevented to 339.

To increase its efficiency, we also modelled the potential benefits of targeted follow-up screening in those at medium risk, for whom the ESC 2021 guidelines suggest considering, but do not explicitly recommend, risk factor treatment[3]. We estimated that, per 100,000 screened, that targeted risk-reclassification with NMR scores would increase the number of CVD events prevented to 336; targeted risk-reclassification with PRSs would increase the number of CVD events prevented to 277; and targeted risk-reclassification with both NMR scores and PRSs would increase the number of CVD events prevented to 370. We also estimated that this targeted follow-up screening would essentially maintain the number statins prescribed per CVD event prevented.

This study represents the largest population health assessment of metabolomic and genomic biomarkers for CVD to date. While our findings suggest that there are potential gains for CVD risk prediction and prevention, there are obvious challenges for validating clinical utility and potential implementation. Commercial providers of NMR biomarkers and PRSs exist, yet fidelity, scale, and cost frequently mean that real world benefits are less than those estimated in prospective cohort studies. Nevertheless, our results indicate that current technologies that can scale to populations (e.g. NMR metabolomics and genomics) have the capacity to moderately improve CVD risk prediction. Our results from clinical biochemistry assays also indicate substantial potential benefit for CVD risk prediction from proteins, which with few exceptions are not measurable by NMR spectroscopy[22]. Initial studies of plasma proteomics scores in UK Biobank have also shown promise for enhancing CVD risk prediction[46], but larger sample sizes are needed to investigate the potential for improving risk stratification at clinically relevant decision-making thresholds. For the ultimate goal of primordial prevention, further studies are also needed to investigate the potential for circulating biomarkers for CVD risk prediction in younger adults[47].

This study has several limitations. Ascertainment bias in UK Biobank means the analysis cohort is healthier than the general UK population[48]. The distribution of risk factors and biomarkers is likely wider in the general population eligible for 10-year CVD risk prediction, and thus the incremental improvement of NMR scores and PRSs may be higher than estimated in this study. On the other hand, the incremental

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

benefits in CVD prevention are potentially overestimated when modelling the population eligible for 10-year CVD risk prediction with SCORE2 because no data were available on the age- and sex- structure of the CVD- and statin- free population, so our population was modelled using the age- and sex- structure of the wider UK population including those ineligible for screening. Nevertheless, despite these potential sources of bias, we believe our findings are robust as we have designed our analyses such that they do not differentially impact the different models for 10-year CVD risk prediction being compared. Finally, our study comprised middle-age adults of almost entirely (>95%) European ancestries in the UK Biobank. Our observations may not generalise to other countries, healthcare systems, or other ancestry groups within the UK. Similarly, while all training and test sets were distinct, the NMR scores were all generated within UK Biobank; therefore, their (relative) performances may differ in other populations. Further studies are needed to evaluate the efficacy and cost-effectiveness of NMR scores and PRSs for improving 10-year CVD risk prediction in these settings.

In conclusion, our results indicate that incorporating scores of NMR metabolomic biomarkers into 10-year CVD risk prediction could enhance prediction of first-onset CVD. We further add to the growing body of evidence that PRSs can be used to enhance CVD risk prediction over conventional risk factors[10,11] and show that improvements in 10-year CVD risk prediction from PRSs are orthogonal to, and can be combined with, NMR scores. Applied at scale, integrating NMR scores alongside PRSs with SCORE2 may have moderate population health benefit.

Methods

Study cohort

UK Biobank is a cohort comprised of ~500,000 participants 35–75 years of age with written informed consent for health related research[48,49]. Participants were members of the UK population recruited through primary care lists whom accepted invitation to attend one of 22 assessment centres across the UK between 2006 and 2010[48].

In this study we analysed a subset of 168,517 participants who at baseline assessment (1) consented for electronic health record linkage, (2) were eligible for 10-year CVD risk prediction with SCORE2[3,4], (3) were not prescribed statins or other lipid lowering medications, and (4) had completed information on risk factors required for SCORE2 computation, (5) had NMR biomarker data (with <5% missingness), (6) had imputed genotypes, and (7) are were eligible for joint analysis with PRSs. **Figure S1** shows the sample exclusions at each step.

Sample exclusion criteria

Eligibility for 10-year CVD risk prediction with SCORE2 was determined following the ESC 2021 guidelines for CVD prevention in clinical practice[3]. UK Biobank participants were included if at baseline assessment they were (1) 40 years of age or older and less than 70 years of age, (2) did not have established atherosclerotic cardiovascular disease (ASCVD), (3) did not have diabetes mellitus, (4) did not have chronic kidney disease, and (5) did not have familial hypercholesterolemia.

Disease history was determined using a combination of self-reported medical history (UK Biobank fields #6150, #4728, #20002, and #20004), prescription medications (fields #6153, #6177, and #20003), and

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

retrospective linkage to hospital episode statistics (fields #41259, #41234, and #41149). Prevalent ASCVD included acute myocardial infarction, acute coronary syndromes, transient ischaemic attack, peripheral arterial disease, and history of revascularization procedures. A full list of International Classification of Diseases (ICD) codes and self-report codes used to define ASCVD are given in **Table S13**. Prevalent diabetes mellitus was determined using the Eastwood *et al.* algorithms[50], and included participants with probable or possible type 1 or type 2 diabetes. Prevalent chronic kidney disease was determined using the UK Biobank algorithmically defined outcome for end-stage renal disease (field #42026).

Familial hypercholesterolemia (FH) was determined using the Dutch Lipid Clinical Network (DLCN) diagnostic criteria[51] as described in the ESC 2021 guidelines for CVD prevention in clinical practice[3]. Participants were excluded where they had low density lipoprotein (LDL) cholesterol quantified by clinical biochemistry assay ≥ 8.5 mmol/L (8 points on the DLCN diagnostic score indicative of probable FH) (fields #30780 and #30786). All participants with possible clinical history used for the DLCN diagnostic criteria were excluded due to prevalent ASCVD. Physical examination and family history data relevant to the DLCN diagnostic criteria were not collected at UK Biobank assessment. Functional mutations in the *LDLR*, *APOB*, and *PCSK9* genes were not assessed in this study.

Participants already prescribed statins or other lipid lowering medications for CVD prevention were excluded as the primary purpose of 10-year CVD risk prediction with SCORE2 is to identify high-risk individuals for statin initiation in apparently healthy adults[3]. These participants were identified as those who at baseline assessment self-reported cholesterol lowering medication on the touchscreen questionnaire on health and medical history (field #6153 for women and field #6177 for men) and/or had been prescribed any of 18 lipid lowering prescription medications for CVD prevention (field #20003). The list of qualifying prescription medications (**Table S14**) was determined by cross-referencing the list of medications present in UK Biobank with the British National Formulary[52] chapter 2.12: Lipid-regulating drugs with the restriction that the drug indications must include prevention of cardiovascular diseases.

Participants were also excluded where it was not possible to predict 10-year CVD risk with SCORE2 due to missing quantitative risk factor information. Quantitative risk factors with missing data included systolic blood pressure (SBP) (missing data at baseline for all instances of fields #93 and #4080) and total cholesterol and high-density lipoprotein (HDL) cholesterol as measured by clinical biochemistry assays (fields #30690 and #30760 respectively).

Participants with >5% missing NMR biomarker data were excluded, as after removal of technical variation, this excess missing data primarily arose due to removal of outlier plates of non-biological origin (89% of missing values).

Finally, participants were excluded if they were used as part of the training for the PRSs for CHD (PGS000018)[8] and ischaemic stroke (PGS000039)[9] analysed in this study.

Electronic health records

UK biobank participants were linked by UK Biobank to hospital inpatient admissions records (fields #41259, #41234, and #41149) for hospitals in England, Wales, and Scotland and to national death registry records (fields #40000, #40001, and #40002). All incident hospital events or death records were coded with ICD-10 codes (or OPCS-4 codes for surgical procedures). Hospital and death records follow-up was available up to 6th March 2018 for events occurring in hospitals in Wales, and to 2021 (beyond 10 years of follow-up) for events occurring in hospitals in England and Scotland.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Retrospective follow-up in hospital records was available from 27th July 1993 for events occurring in hospitals in England, 2nd December 1980 for events occurring in hospitals in Scotland, and 18th April 1991 for events occurring in hospitals in Wales, with median of 15.75 years retrospective follow-up (maximum 29.76 years). Retrospective hospital events were coded with a combination of ICD-10 and ICD-9 codes (or OPCS-4 and OPCS-3 codes for surgical procedures).

Participants with withdrawn consent for electronic health record linkage were identified from field #190 for sample exclusion.

NMR biomarker data quantification and quality control

NMR metabolite biomarker data was quantified in ~275,000 randomly selected participants as previously described[18,53]. Briefly, NMR spectroscopy (Nightingale Health Plc.) was used to measure the absolute concentrations of 168 biomarkers and 81 biomarker ratios from non-fasting plasma samples (UK Biobank aliquot 3). Details on the identity of the 249 NMR metabolite biomarkers are provided in **Table S1**.

Technical variation was subsequently removed using a modified version of our previously described pipeline[27] that has been updated to reflect our exploration the additional ~150,000 participants measured since the pipeline development. The updated pipeline is available in version 2.2.1 of the ukbnmr R package. Briefly, technical variation removed included (1) time between sample preparation and sample measurement, (2) systematic differences in biomarker concentrations in each shipping batch due to sample position on the 96-well shipment plate, (3) measurement drift over time, (4) inter-spectrometer differences, and (5) shipment plates with systematically extreme concentrations of non-biological origin. For further details see https://github.com/sritchie73/ukbnmr.

Clinical biochemistry assay quantification and quality control

Targeted blood biochemistry assays were quantified in all ~500,000 participants as previously described[54]. Briefly, absolute concentrations of 30 circulating biomarkers were quantified from serum (29 biomarkers) or red blood cell samples (glycated haemoglobin; HbA1c) using 24 analysis methods across six analytical platforms from five manufacturers (AU5800, Beckman Coulter; AU5800, Randox; LIASON XL, DiaSorin Ltd.; VARIANT II Turbo, Bio-Rad; and ADVIA 1800, Siemens). Missing data arising due to biomarker concentrations being above or below limits of reportability or detection were replaced with the largest or smallest non-missing value of that biomarker respectively with a small offset of 0.0001 units. Further details on the 30 biomarkers can be found in **Table S2**.

Genotyping, imputation, and polygenic risk scores

UK Biobank participants were genotyped on UK BiLEVE arrays and UK Biobank Axiom arrays and imputed to the 1000 genomes, UK10K, and Haplotype Reference Consortium panels[55] using human genome build GRCh37[49]. Here, we converted the data to PLINK2 probabilistic dosages[56] for analyses.

PRSs for CHD and ischaemic stroke were obtained from the polygenic score (PGS) catalog[57] (accessions PGS000018 and PGS000039 respectively) and computed in UK Biobank participants using the PLINK2 software[56] (PLINK v2.00a3LM AVX2 Intel [2 Mar 2021]) linear scoring function applied to the probabilistic allele dosages. These PRSs were chosen for this study as they are the most predictive PRSs for CHD and ischaemic stroke that do not include GWAS summary statistics derived from UK Biobank participants in their model development.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Cardiovascular risk factors

Age and sex at baseline assessment were obtained from UK Biobank fields #21003 and #31 respectively.

Systolic blood pressure (SBP) was measured using either an automated digital device (OMRON) (field: #4080) and/or by manual sphygmomanometer (field: #93). In both cases, two measurements were taken several moments apart, and the average was taken to obtain a single representative measure of SBP.

Smoking status was defined as "current" or "other" based on self-reported current smoking (field #20116). Missing data ("don't know", "prefer not to answer") were set to "other".

Incident cardiovascular disease

Incident CVD events were defined following the definition used by the SCORE2 working group and ESC Cardiovascular risk collaboration[4] to include fatal hypertensive disease (ICD-10 codes I10–I16), fatal ischaemic heart disease (ICD-10 codes I20–I25), fatal arrhythmias or heart failure (ICD-10 codes I46–I52, excluding I51.4), fatal cerebrovascular disease (ICD-10 codes I60–I69, excluding I60, I62, I67.1, I68.2, and I67.1), fatal atherosclerosis or abdominal aortic aneurysm (ICD-10 codes I70–I73), sudden death and death within 24 hours of symptom onset (ICD-10 codes R96.0 and R96.1), non-fatal myocardial infarction (ICD-10 I21–I23), and non-fatal stroke (ICD-10 codes I60–I69, excluding I60, I62, I67.1, I68.2, and I67.1).

Follow-up time for each participant was restricted to a maximum of 10 years, defined as the difference in years between baseline assessment and the earliest of the following records: (1) the date of the first CVD event, (2) the date of death, (3) the date lost to follow-up (fields #190 and #191; e.g. participant reported to NHS or UK Biobank as having left the UK), (4) the maximum follow-up date in hospital records from Wales—6th March 2018—for participants located in Wales at baseline assessment or inferred to have moved to Wales since baseline assessment (based on presence of hospital records from Wales before 2018 and none in England or Scotland after, or change in location for subsequent UK Biobank assessments), or (5) date of baseline assessment plus ten years. Cohort characteristics reported in **Table 1** include details on the number of people with non-CVD related mortality or otherwise lost to follow-up prior to 10 years.

SCORE2

Sex-specific per-participant values for SCORE2 (linear predictors) were computed from age, smoking status, SBP, total cholesterol, and HDL cholesterol using formulae published by the SCORE2 working group and ESC Cardiovascular risk collaboration[4]. Each risk factor was transformed as described in Supplementary methods Table 2 of the SCORE2 publication[4]:

$$age = (age - 60)/5$$

$$smoking = \begin{cases} 1 \; if \text{ "}current\text{"} \\ 0 \; if \text{ "other"} \end{cases}$$

$$SBP = (SBP - 120)/20$$

$$tchol = (Total\, cholesterol - 6)/1$$

$$HDL = (HDL\, cholesterol - 1.3)/0.5$$

Then each transformed risk factor and risk factor × age- interaction was multiplied by the log hazard ratio obtained in the SCORE2 sensitivity analysis excluding UK Biobank participants from the log hazard ratio estimation (obtained from Supplementary Table 8 of the SCORE2 publication[4] under the "Excluding UK Biobank" heading):

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

$$SCORE2_{male} = \log(1.50)\ age + \log(1.77)\ smoking + \log(1.33)\ SBP + \log(1.13)\ tchol + \log(0.80)\ HDL + \log(0.92)(age \times smoking) + \log(0.98)\ (age \times SBP) + \log(0.98)(age \times tchol) + \log(1.04)\ (age \times HDL)$$

$$SCORE2_{female} = \log(1.64)\ age + \log(2.09)\ smoking + \log(1.39)\ SBP + \log(1.11)\ tchol + \log(0.81)\ HDL + \log(0.89)(age \times smoking) + \log(0.97)\ (age \times SBP) + \log(0.98)(age \times tchol) + \log(1.06)\ (age \times HDL)$$

These log hazard ratios were used to prevent overestimation of the efficacy of SCORE2 for 10-year CVD risk prediction in this study (**Figure S11**), which could result in underestimation of any potential improvements in risk discrimination from addition of biomarkers or PRS.

The sex-stratified C-index for SCORE2 in the 168,517 study participants was computed directly from this SCORE2 linear predictor using the concordance function in the survival R package version 3.3-1. The 95% confidence interval was computed from the standard error, which was computed by the survival R package using the infinitesimal jackknife method.

Incremental value in 10-year CVD risk prediction for individual biomarkers

Incremental improvement in 10-year CVD risk prediction for individual biomarkers beyond SCORE2 alone was assessed using differences in C-index from SCORE2 alone (ΔC-index) (**Figure 2**, **Table S3**). Incremental improvement in 10-year CVD risk was assessed for the 249 NMR biomarkers (**Table S1**) and 28 of the 30 clinical biochemistry assay biomarkers (**Table S2**): clinical biochemistry assays for HDL cholesterol and total cholesterol were not assessed here as they were used to compute SCORE2 linear predictor (see above).

For each biomarker, we fit a sex-stratified Cox proportional hazards regression for 10-year CVD risk with the biomarker as an independent variable and SCORE2 as an offset term in the 168,517 study participants (5,096 incident CVD cases) (**Figure 2**, **Table S3A**). SCORE2 was treated as an offset term, rather than an independent variable, as we sought to develop scores that added biomarkers to the existing SCORE2 weights. Cox proportional hazards regressions were fit using the coxph function in the survival R package version 3.3-1. The hazard ratio, absolute C-index, and its standard error were obtained directly from the returned result for each SCORE2 + biomarker model (**Table S3A**). The standard error was computed by the survival R package using the infinitesimal jackknife method.

The ΔC-index was subsequently computed for each SCORE2 + biomarker model by subtracting the sex-stratified C-index for the SCORE2 linear predictor as described above (**Figure 2**, **Table S3A**). A bootstrap procedure with 1,000 bootstraps was used to estimate the standard error for the ΔC-index. Bootstrap resampling was performed using methods appropriate for right-censored data[58]; using the censboot function in the boot R package version 1.3-28. 95% confidence intervals and two-sided P-values were computed from the bootstrap standard error using the first order normal approximation method. P-values were corrected for multiple testing across the 277 tested biomarkers using Benjamini-Hochberg false-discovery rate (FDR) correction.

In addition to sex-stratified analysis, incremental improvements in C-index beyond SCORE2 alone were similarly assessed in sex-specific analysis (**Table S3B–C**); performed separately in the 72,441 male participants (3,208 incident CVD cases) (**Table S3B**) and in the 96,076 female participants (1,888 incident CVD cases) (**Table S3C**).

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

NMR biomarker score training

Four NMR biomarker scores were developed: (1) an NMR score in males for predicting 10-year risk of coronary heart disease (CHD), (2) an NMR score in females for predicting 10-year risk of CHD, (3) an NMR score in males for predicting 10-year risk of ischaemic stroke, and (4) an NMR score in females for predicting 10-year risk of ischaemic stroke.

Incident CHD was defined as fatal or non-fatal myocardial infarction (ICD-10 codes I21–I24 or I25.2) or by presence of major coronary surgery (ICD-10 codes Z95.1 or OPSC-4 codes K40–K46, K49, K50.1, K75). Incident ischaemic stroke was determined using the respective UK Biobank algorithmically defined outcome (field #42008). Follow-up time was defined for each endpoint separately; *i.e.*, a non-fatal ischaemic stroke event prior to CHD was not counted as a competing risk when defining follow-up for incident CHD and *vice versa*. Case numbers and follow-up characteristics are reported in **Table S4A** for CHD and **Table S4B** for ischaemic stroke.

Each of the four NMR biomarker scores were developed using elastic-net penalised Cox proportional hazards regression[28,29] in data from the 168,517 participants. A nested cross-validation procedure was used, comprising a 5-fold outer layer and 10-fold inner layer (**Figure S2**). For each iteration at the outer layer, NMR biomarker scores were trained in 4/5^{ths} of the data, then predicted in the withheld 1/5th of the data (test fold) (**Figure S2A**). Model training in each iteration used 10-fold cross-validation (inner-layer) for hyperparameter tuning of the elasticnet model (**Figure S2B**). Random allocation of participants to test folds was performed using the caret R package version 6.0-92 to balance the number of males and females, and incident cases within males and females, across test folds. Test folds at the inner cross-validation layer were balanced by incident CHD or ischaemic stroke cases respectively when training NMR scores for the respective endpoint. Test-folds at the outer cross-validation layer were balanced by incident CVD cases as the outer cross-validation layer was also used for the training procedure used to combine NMR scores for CHD and stroke described in the next section.

For each of the five iterations of training, sex- and endpoint- specific NMR scores were trained using elastic-net penalised Cox proportional hazards regression using the glmnet R package version 4.1-6 (**Figure S2A**). Each NMR biomarker score was trained with the 106 non-derived NMR biomarkers with SCORE2 as an offset term. Biomarkers were classified as non-derived where they could not be computed by summing or dividing two or more other biomarkers[27] (**Table S1**).

Prior to NMR biomarker score training, missing data in the 106 NMR biomarkers were imputed as glmnet could not handle missing data. Missing data were imputed a single time in the 168,517 participants using the impute R package version 1.70.0 with the K-nearest neighbours algorithm[59]. K was set to 20 based on the number of principal components that cumulatively explained >95% of the variation in the 106 non-derived biomarkers. Prior to imputation, >92% of participants had no missing NMR biomarker concentrations, >6% had only one biomarker missing, and the remaining <2% had 2–5 of 106 biomarkers missing. In total 0.1% of biomarker concentrations were imputed. Per-biomarker missingness rates are given in **Table S1**. Biomarker concentrations were standardised in males and females separately so that sex-specific coefficients fit by glmnet were comparable across biomarkers.

For each of the five training iterations, the optimal elastic-net fit was determined with a grid search of α and λ where the range of α was [0 (ridge regression), 0.1, 0.25, 0.5, 0.75, 0.9, and 1 (lasso regression)] and a sequence of 100 λ values was automatically determined by the glmnet function for each α (**Figure S3A**).

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

β coefficients estimated for each biomarker for the optimal fit in each training data are shown in **Figure S3C**. The proportion of variance in the NMR score explained by each biomarker was calculated as the ratio of its β^2 to the sum of β^2 across all biomarkers.

To obtain a single representative score (*i.e.*, for computing NMR scores in new samples), coefficients were averaged across the five training iterations (**Figures S3C**, **Table S5**). Proportions of variance explained in the NMR score by each biomarker were also averaged across the five training iterations (**Table S5**).

To avoid overfitting in our downstream analyses of individual NMR scores (**Figure 3A**, **Figure S4**, **Figure S5A**, **Figure S6**, **Table S6**), the aggregate of the predicted NMR scores in the five withheld test folds were used (**Figure S2A**).

Independent associations of NMR scores, PRS, and risk factors with 10-year CVD risk

Two sets of multivariable Cox proportional hazards regressions were fit in the 168,517 participants to assess independent associations between predicted NMR scores, PRSs, and SCORE2 risk factors with 10-year CVD risk prediction (**Figure 3A**, **Figure S5A**, **Table S6**): (1) Multivariable Cox proportional hazards regression was fit with the CHD NMR score, stroke NMR score, CHD PRS, and stroke PRS as predictor variables and SCORE2 as an offset term (**Figure 3A**, **Figure S5A**, **Table S6A**); and (2) multivariable Cox proportional hazards regression was fit additionally with the individual SCORE2 risk factors as predictor variables instead of using SCORE2 as an offset term (**Figure S6**, **Table S6B**). Both multivariable models were fit in sex-stratified analysis and sex-specific (fit in the male and female participants separately) analyses (**Figure 3A, Figure S5A, Figure S6, Table S6**).

Combining NMR biomarker scores and/or PRSs with SCORE2

NMR scores and PRSs were combined with SCORE2 into three new scores for predicting 10-year CVD risk: (1) SCORE2 + NMR scores, (2) SCORE2 + PRSs, and (3) SCORE2 + NMR scores + PRSs.

Sex-specific combined linear predictors were created by adding zero-centred linear predictors for the CHD NMR score, stroke NMR score, CHD PRS, and stroke PRS to the SCORE2 linear predictor after multiplying each score by scaling factors weighting their relative contributions to 10-year CVD prediction over SCORE2 alone. Sex- and model- specific scaling factors were estimated using Cox proportional hazards regression fit for 10-year CVD risk with the relevant respective scores as independent predictor variables and SCORE2 as an offset term.

To avoid overfitting, offsets for score centring and scaling factors were calculated as part of the cross-validation procedure described above (**Figure S2C**). For downstream analyses (**Figure 3B–C, Figure 4, Figure S5B–C, Figure S9, Table S8–S10, Table S12**), the aggregate of the predicted values for the combined linear predictors across the five withheld test folds were used (**Figure S2A**). To obtain a single set of representative model coefficients for computing the combined scores in new samples, model coefficients were average across the five training iterations (**Table S7**).

Incremental value in 10-year CVD risk prediction for NMR biomarker scores and PRSs

Incremental improvement in 10-year CVD risk prediction beyond SCORE2 alone for each of the three combined scores (above) were assessed using differences in sex-stratified C-index in the 168,517 study participants (5,096 incident CVD cases) (**Figure 3B**, **Table S8**).

Sex-stratified C-indices and their infinitesimal jackknife standard errors were computed directly from the linear predictors for each of the three models using the concordancefit function in the survival package. The

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

sex-stratified ΔC-index was computed as the difference between the C-index for SCORE2 and the C-index for each model. 95% confidence intervals and P-values were calculated from the bootstrap standard error after estimating the ΔC-index in a bootstrap procedure with 1,000 bootstraps. Bootstrap resampling was performed using methods appropriate for right-censored data[58]; using the censboot function in the boot R package version 1.3-28.

In addition to sex-stratified analysis, incremental improvements in C-index beyond SCORE2 alone were similarly assessed in sex-specific analysis (**Figure S5B**, **Table S8**); performed separately in the 72,441 male participants (3,208 incident CVD cases) and in the 96,076 female participants (1,888 incident CVD cases).

Computation of absolute risks

Linear predictors for SCORE2, SCORE2 + NMR scores, SCORE2 + PRSs, and SCORE2 + NMR scores + PRSs were converted into predictions of absolute 10-year CVD risk using formulae developed by the SCORE2 working group and ESC cardiovascular risk consortium[4]:

First, uncalibrated absolute risks were calculated using sex-specific estimates of baseline survival, which they calculated as the median baseline survival across 44 cohorts (including UK Biobank)[4]:

$$Uncalibrated\ 10year\ risk_{Males} = 1 - 0.9605^{\exp(linear\ predictor)}$$

$$Uncalibrated\ 10year\ risk_{Females} = 1 - 0.9776^{\exp(linear\ predictor)}$$

Second, these were converted into absolute 10-year risks calibrated to the UK population using their scaling factors for the low-risk European region (which included the UK) applied in their risk recalibration formula[4]:

$$Calibrated\ 10year\ risk_{Males} = 1 - exp\left(-exp\left(-0.5699 + 0.7476 \times \ln\left(-ln(1 - Uncalibrated\ 10year\ risk)\right)\right)\right)$$

$$Calibrated\ 10year\ risk_{Females} = 1 - exp\left(-exp\left(-0.7380 + 0.7019 \times \ln\left(-ln(1 - Uncalibrated\ 10year\ risk)\right)\right)\right)$$

Distributions of absolute risk for each model are shown in **Figure S8**.

Incremental value in stratification at risk thresholds used for clinical decision making

Categorical net reclassification improvement (NRI) analysis[60,61] was used to assess the incremental value of NMR scores and/or PRSs over SCORE2 for stratifying individuals based on risk thresholds used for clinical decision making.

Participants were stratified into categories of low risk, moderate risk, and high risk based on their absolute 10-year CVD risk predicted by each model using risk thresholds recommended by the ESC 2021 guidelines for CVD prevention[3] (**Table S9**). Participants <50 years of age were allocated to the low-risk group if their absolute 10-year CVD risk was < 2.5%, to the medium risk group if their absolute 10-year CVD risk was < 7.5% risk, and to the high-risk group if their absolute 10-year CVD risk was ≥ 7.5%. Participants 50 years or older were allocated to the low-risk group if their absolute 10-year CVD risk was < 5%, to the medium risk group if their absolute 10-year CVD risk was < 10% risk, and to the high-risk group if their absolute 10-year CVD risk was ≥ 10%.

The three models adding NMR scores and/or PRSs to SCORE2 were assessed in comparison to SCORE2 alone in categorical NRI analysis using the nricens R package version 1.6. Categorical NRI analysis was

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

used to assess (1) the % of incident CVD cases correctly reclassified from a lower risk group into higher risk group, and (2) the % of non-cases correctly reclassified from a higher risk group into a lower risk group. Categorical NRI analysis was performed separately in all participants (**Figure 3C**), male participants (**Figure S5C**), and female participants (**Figure S5C**). Bootstrap resampling of the categorical NRI analysis was performed using the nricens R package, and 95% confidence intervals and P-values were subsequently calculated from the bootstrap standard error (**Figure 3C**, **Figure S5C**, **Table S10**).

Population simulation

For downstream analyses we simulated a hypothetical population of approximately 100,000 individuals with age- and sex- structure and expected 10-year CVD incidence rates representative of the UK population eligible for 10-year CVD risk screening with SCORE2 (**Figure S8, Table S11**).

Sex-specific CVD incidence rates expected in each five-year age-group (**Figure S8**) were obtained from Table A in the S2 Text in Sun *et al* 2021[11]. These CVD incidence rates were obtained by Sun *et al.* from a random sample of 2.1 million people amongst 11.3 million CVD- and statin- free primary care patients 35–74 years of age who, between 2004–2017, attended any of 674 general practices opting into data linkage to the UK Clinical Practice Research Datalink (CPRD)[62]. The published per-1000-year CVD incidence rates in the five-years ahead (Table A in the S2 Text in Sun *et al* 2021[11]; **Figure S8**) were converted into percentages of males and females expected to have an incident CVD event within 10 years (**Figure S8**) as 1−exp(rate/1000). Numbers of CVD- and statin- free individuals in each sex- and age- group in CRPD were not published[4,11]. Therefore, we derived the sizes of each age- and sex- group using the age- and sex-structure of the general UK population in the mid-2020 UK population estimates published by the Office of National Statistics[30].

The simulated population was derived by multiplying by 100,000 the relative sizes of each age-group and sex among those 40–69 in the mid-2020 UK population estimate, then multiplying by the percentages expected to have incident CVD within 10 years derived from CPRD (**Figure S8**, **Table S11**).

Incremental value for CVD prevention with population-wide screening

To assess the relative benefits for CVD prevention of each model if applied at scale to the UK population (**Figure 4A, Table S12A**) we stratified the simulated population into the low-, medium-, and high-risk groups based on the proportions allocated to each category in UK Biobank by SCORE2 alone (**Figure S9A**), SCORE2 + NMR scores, SCORE2 + PRSs, and SCORE2 + NMR scores + PRSs. We modelled statin initiation in the high-risk group, who based on their risk thresholds would be recommended for risk factor treatment by the ESC 2021 guidelines for CVD prevention[3]. The impact of statin initiation was modelled as preventing one in five simulated incident CVD events assuming a 20% reduction in 10-year CVD risk[31].

Benefits for CVD prevention were quantified for each model using five statistics: (1) the number of people classified as high-risk ($N_{high\text{-}risk}$), (2) the number of incident CVD events amongst those classified as high-risk ($CVD_{high\text{-}risk}$), (3) the expected number of events prevented by statin initiation in the high-risk group ($CVD_{prevented}$), (4) the number of people needed to screen to prevent one CVD event (NNS; calculated as $N/CVD_{prevented}$), and (5) the number of statins prescribed per CVD event prevented (NNT; calculated as $CVD_{high\text{-}risk}/CVD_{prevented}$). Incremental benefits for CVD prevention were assessed by differences in these five statistics from SCORE2 alone ($\Delta N_{high\text{-}risk}$, $\Delta CVD_{high\text{-}risk}$, $\Delta CVD_{prevented}$, ΔNNS, and ΔNNT).

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

95% confidence intervals and P-values were calculated from the bootstrap standard error estimated for each statistic in a bootstrap procedure with 1,000 bootstraps. Bootstrap resampling was performed using methods appropriate for right-censored data[58]; using the censboot function in the boot R package version 1.3-28.

Incremental value for CVD prevention with targeted screening

Incremental benefits to CVD prevention were further modelled when using SCORE2 + NMR scores, SCORE2 + PRSs, or SCORE2 + NMR scores + PRSs for targeted risk-reclassification of those classified as medium risk by SCORE2 alone (**Figure 4B, Table S12B**), for whom the ESC 2021 guidelines suggest considering, but do not explicitly recommend, risk factor treatment[3]. For targeted-screening, the SCORE2-classified medium-risk subset of the simulated population was re-stratified into the low-, medium, and high- CVD risk groups based on proportions re-stratified into each category in UK Biobank by each alternative model (**Figure S9B**, **Methods**). After targeted screening, the high-risk group comprised those classified as high-risk either by SCORE2 alone or the alternative model incorporating NMR scores and/or PRSs. Benefits for CVD preventing were quantified as described above.

Ethics statement

This study was approved under UK Biobank Projects 30418 and ethics approval was obtained from the North West Multi-Center Research Ethics Committee. UK Biobank participants provided written informed consent for health related research[48].

Data Availability

All data described are available through UK Biobank subject to approval from the UK Biobank access committee. See https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access for further details.

Code Availability

Code underlying this paper are available at https://github.com/sritchie73/UKB_NMR_CVD_prediction/. This repository and specific release for this paper are permanently archived by Zenodo at https://doi.org/10.5281/zenodo.10059234.

Acknowledgements

The authors are grateful to UK Biobank for access to data to undertake this study (Projects #30418).

Nightingale Health Plc is acknowledged for early access to the UK Biobank NMR biomarker data and discussions regarding sources of experimental variation.

This work was performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (www.csd3.cam.ac.uk), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/P020259/1), and DiRAC funding from the Science and Technology Facilities Council (www.dirac.ac.uk).

This work was supported by core funding from the: Cambridge BHF Centre of Research Excellence (RE/18/1/34212) and BHF Chair Award (CH/12/2/29428).

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

S.C.R. and S.K. were funded by a British Heart Foundation (BHF) Programme Grant (RG/18/13/33946). S.C.R. was also funded by the National Institute for Health and Care Research (NIHR) Cambridge BRC (BRC-1215-20014; NIHR203312) [*]. X.J. was funded by British Heart Foundation (CH/12/2/29428) and Wellcome Trust (227566/Z/23/Z). L.P. and P.S. were supported by a Rutherford Fund Fellowship from the Medical Research Council grant MR/S003746/1. Y.X. and M.I. were supported by the UK Economic and Social Research Council (ES/T013192/1). S.A.L. was supported by a Canadian Institutes of Health Research postdoctoral fellowship (MFE-171279). E.D.A. holds a NIHR Senior Investigator Award. J.D. holds a BHF Professorship and a NIHR Senior Investigator Award. M.I. is supported by the Munz Chair of Cardiovascular Prediction and Prevention and the NIHR Cambridge Biomedical Research Centre (NIHR203312).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

*The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Competing Interests

During the course of this project P.S. became a full-time employee of GSK Plc. All significant contributions to this study were made prior to this role and GSK Plc had no input to the study. J.D. serves on scientific advisory boards for AstraZeneca, Novartis, and UK Biobank, and has received multiple grants from academic, charitable and industry sources outside of the submitted work. A.S.B. reports institutional grants from AstraZeneca, Bayer, Biogen, BioMarin, Bioverativ, Novartis, Regeneron and Sanofi. The remaining authors declare no competing interests.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

References

1. Rabar, S., Harker, M., O'Flynn, N., Wierzbicki, A. S. & Guideline Development Group. Lipid modification and cardiovascular risk assessment for the primary and secondary prevention of cardiovascular disease: summary of updated NICE guidance. *BMJ* **349**, g4356 (2014).
2. Arnett, D. K. *et al.* 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J. Am. Coll. Cardiol.* **74**, e177–e232 (2019).
3. Visseren, F. L. J. *et al.* 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. *Eur. Heart J.* **42**, 3227–3337 (2021).
4. SCORE2 working group and ESC Cardiovascular risk collaboration. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. *Eur. Heart J.* **42**, 2439–2454 (2021).
5. Niiranen, T. J. & Vasan, R. S. Epidemiology of cardiovascular disease: recent novel outlooks on risk factors and clinical approaches. *Expert Rev. Cardiovasc. Ther.* **14**, 855–869 (2016).
6. Mora, S. & Ridker, P. M. Justification for the Use of Statins in Primary Prevention: an Intervention Trial Evaluating Rosuvastatin (JUPITER)—can C-reactive protein be used to target statin therapy in primary prevention? *Am. J. Cardiol.* **97**, 33A-41A (2006).
7. Ridker, P. M. *et al.* Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *N. Engl. J. Med.* (2017) doi:10.1056/NEJMoa1707914.
8. Inouye, M. *et al.* Genomic Risk Prediction of Coronary Artery Disease in 480,000 Adults: Implications for Primary Prevention. *J. Am. Coll. Cardiol.* **72**, 1883–1893 (2018).
9. Abraham, G. *et al.* Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke. *Nat. Commun.* **10**, 5819 (2019).
10. Polygenic Risk Score Task Force of the International Common Disease Alliance. Responsible use of polygenic risk scores in the clinic: potential benefits, risks and gaps. *Nat. Med.* **27**, 1876–1884 (2021).

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

11. Sun, L. *et al.* Polygenic risk scores in cardiovascular risk prediction: A cohort study and modelling analyses. *PLoS Med.* **18**, e1003498 (2021).
12. Patel, A. P. *et al.* A multi-ancestry polygenic risk score improves risk prediction for coronary artery disease. *Nat. Med.* **29**, 1793–1803 (2023).
13. O'Sullivan, J. W., Ashley, E. A. & Elliott, P. M. Polygenic risk scores for the prediction of cardiometabolic disease. *Eur. Heart J.* **44**, 89–99 (2023).
14. O'Sullivan, J. W. *et al.* Polygenic Risk Scores for Cardiovascular Disease: A Scientific Statement From the American Heart Association. *Circulation* **146**, e93–e118 (2022).
15. The Emerging Risk Factors Collaboration. C-Reactive Protein, Fibrinogen, and Cardiovascular Disease Prediction. *N. Engl. J. Med.* **367**, 1310–1320 (2012).
16. Emerging Risk Factors Collaboration *et al.* Lipid-related markers and cardiovascular disease prediction. *JAMA* **307**, 2499–2506 (2012).
17. Lin, J. S. *et al.* Nontraditional Risk Factors in Cardiovascular Disease Risk Assessment: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA* **320**, 281–297 (2018).
18. Würtz, P. *et al.* Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technology. *Am. J. Epidemiol.* 1–13 (2017).
19. Ala-Korpela, M., Zhao, S., Järvelin, M.-R., Mäkinen, V.-P. & Ohukainen, P. Apt interpretation of comprehensive lipoprotein data in large-scale epidemiology – disclosure of fundamental structural and metabolic relationships. *medRxiv* (2021) doi:10.1101/2021.03.08.21253123.
20. Connelly, M. A., Otvos, J. D., Shalaurova, I., Playford, M. P. & Mehta, N. N. GlycA, a novel biomarker of systemic inflammation and cardiovascular disease risk. *J. Transl. Med.* **15**, 219 (2017).
21. Ritchie, S. C. *et al.* Elevated serum alpha-1 antitrypsin is a major component of GlycA-associated risk for future morbidity and mortality. *PLoS One* **14**, e0223692 (2019).

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

22. Soininen, P., Kangas, A. J., Würtz, P., Suna, T. & Ala-Korpela, M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ. Cardiovasc. Genet.* **8**, 192–206 (2015).

23. Nightingale Health Biobank Collaborative Group *et al.* Metabolomic and genomic prediction of common diseases in 477,706 participants in three national biobanks. *medRxiv* 2023.06.09.23291213 (2023) doi:10.1101/2023.06.09.23291213.

24. Buergel, T. *et al.* Metabolomic profiles predict individual multidisease outcomes. *Nat. Med.* **28**, 2309–2320 (2022).

25. Jin, D., Trichia, E., Islam, N., Lewington, S. & Lacey, B. Predictive value of metabolic profiling in cardiovascular risk scores: analysis of 75 000 adults in UK Biobank. *J. Epidemiol. Community Health* (2023) doi:10.1136/jech-2023-220801.

26. Fu, Z. *et al.* Association between NMR metabolomic signatures of healthy lifestyle and incident coronary artery disease. *Eur. J. Prev. Cardiol.* **30**, 243–253 (2023).

27. Ritchie, S. C. *et al.* Quality control and removal of technical variation of NMR metabolic biomarker data in ~120,000 UK Biobank participants. *Scientific Data* **10**, 1–15 (2023).

28. Simon, N., Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *J. Stat. Softw.* **39**, 1–13 (2011).

29. Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **33**, 1–22 (2010).

30. Park, N. Population estimates for the UK, England and Wales, Scotland and Northern Ireland - Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2020 (2021).

31. Collins, R. *et al.* Interpretation of the evidence for the efficacy and safety of statin therapy. *Lancet* **388**, 2532–2561 (2016).

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

32. Libby, P., Ridker, P. M. & Maseri, A. Inflammation and atherosclerosis. *Circulation* **105**, 1135–1143 (2002).

33. IL6R Genetics Consortium Emerging Risk Factors Collaboration *et al.* Interleukin-6 receptor pathways in coronary heart disease: a collaborative meta-analysis of 82 studies. *Lancet* **379**, 1205–1213 (2012).

34. Ridker, P. M. & Rane, M. Interleukin-6 Signaling and Anti-Interleukin-6 Therapeutics in Cardiovascular Disease. *Circ. Res*. **128**, 1728–1746 (2021).

35. Taglieri, N., Koenig, W. & Kaski, J. C. Cystatin C and cardiovascular risk. *Clin. Chem.* **55**, 1932–1943 (2009).

36. Knight, E. L. *et al.* Factors influencing serum cystatin C levels other than renal function and the impact on renal function measurement. *Kidney Int.* **65**, 1416–1421 (2004).

37. Chien, S.-C., Chen, C.-Y., Lin, C.-F. & Yeh, H.-I. Critical appraisal of the role of serum albumin in cardiovascular disease. *Biomark Res* **5**, 31 (2017).

38. Gillum, R. F., Ingram, D. D. & Makuc, D. M. Relation between serum albumin concentration and stroke incidence and death: the NHANES I Epidemiologic Follow-up Study. *Am. J. Epidemiol.* **140**, 876–888 (1994).

39. Ronit, A. *et al.* Plasma Albumin and Incident Cardiovascular Disease: Results From the CGPS and an Updated Meta-Analysis. *Arterioscler. Thromb. Vasc. Biol.* **40**, 473–482 (2020).

40. Soeters, P. B., Wolfe, R. R. & Shenkin, A. Hypoalbuminemia: Pathogenesis and Clinical Significance. *JPEN J. Parenter. Enteral Nutr*. **43**, 181–193 (2019).

41. Gabay, C. & Kushner, I. Acute-phase proteins and other systemic responses to inflammation. *N. Engl. J. Med*. **340**, 448–454 (1999).

42. Otvos, J. D. *et al.* GlycA: A Composite Nuclear Magnetic Resonance Biomarker of Systemic Inflammation. *Clin. Chem.* **61**, 714–723 (2015).

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

43. Ritchie, S. C. *et al.* The Biomarker GlycA Is Associated with Chronic Inflammation and Predicts Long-Term Risk of Severe Infection. *Cell Systems* **1**, 293–301 (2015).

44. Akinkuolie, A. O., Buring, J. E., Ridker, P. M. & Mora, S. A novel protein glycan biomarker and future cardiovascular disease events. *J. Am. Heart Assoc.* **3**, e001221 (2014).

45. Duprez, D. A. *et al.* GlycA, a novel NMR biomarker of inflammation, and incident cardiovascular events: the Multi-Ethnic Study of Atherosclerosis. *J. Am. Coll. Cardiol.* **65**, (2015).

46. Gadd, D. A. *et al.* Blood protein levels predict leading incident diseases and mortality in UK Biobank. *medRxiv* 2023.05.01.23288879 (2023) doi:10.1101/2023.05.01.23288879.

47. Braunwald, E. Will primordial prevention change cardiology? *Eur. Heart J.* **44**, 3307–3308 (2023).

48. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).

49. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).

50. Eastwood, S. V. *et al.* Algorithms for the Capture and Adjudication of Prevalent and Incident Diabetes in UK Biobank. *PLoS One* **11**, e0162388 (2016).

51. Nordestgaard, B. G. *et al.* Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: consensus statement of the European Atherosclerosis Society. *Eur. Heart J.* **34**, 3478–90a (2013).

52. Joint Formulary Committee. *British National Formulary*. (London: BMJ Group and Pharmaceutical Press).

53. Julkunen, H. *et al.* Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. *Nat. Commun.* **14**, 604 (2023).

54. Allen, N. E. *et al.* Approaches to minimising the epidemiological impact of sources of systematic and random variation that may affect biochemistry assay data in UK Biobank. *Wellcome Open Research* **5**, (2020).

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

55. Loh, P.-R. *et al*. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet*. **48**, 1443–1448 (2016).

56. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).

57. Lambert, S. A. *et al.* The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation. *Nat. Genet.* **53**, 420–425 (2021).

58. Efron, B. Censored Data and the Bootstrap. *J. Am. Stat. Assoc.* **76**, 312–319 (1981).

59. Troyanskaya, O. *et al*. Missing value estimation methods for DNA microarrays. *Bioinformatics* **17**, 520–525 (2001).

60. Pencina, M. J., D'Agostino, R. B., Sr & Steyerberg, E. W. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat. Med*. **30**, 11–21 (2011).

61. Uno, H., Tian, L., Cai, T., Kohane, I. S. & Wei, L. J. A unified inference procedure for a class of measures to assess improvement in risk prediction systems with survival data. *Stat. Med.* **32**, 2430–2442 (2013).

62. Herrett, E. *et al.* Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int. J. Epidemiol*. **44**, 827–836 (2015).

63. Ritchie S. C. Cardiovascular risk prediction using metabolomic biomarkers and polygenic risk scores: A cohort study and modelling analyses. *Zenodo*. https://doi.org/10.5281/zenodo.10059234 (2023).

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Figures

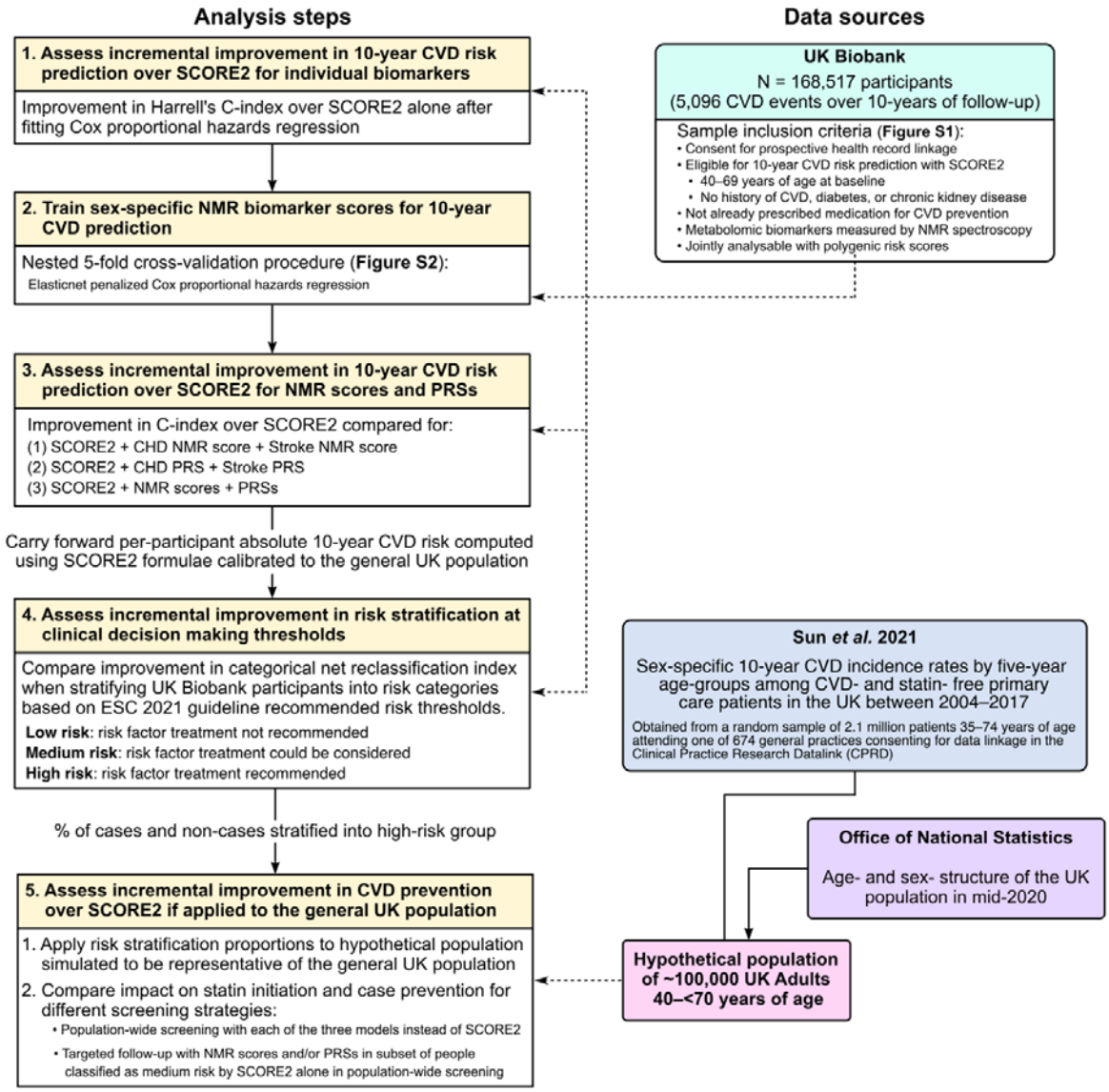

Figure 1: Study design.

For details on sample inclusion and exclusion criteria when defining study eligibility for UK Biobank participants see **Figure S1**. A schematic of the NMR score training procedure is provided in **Figure S2**.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

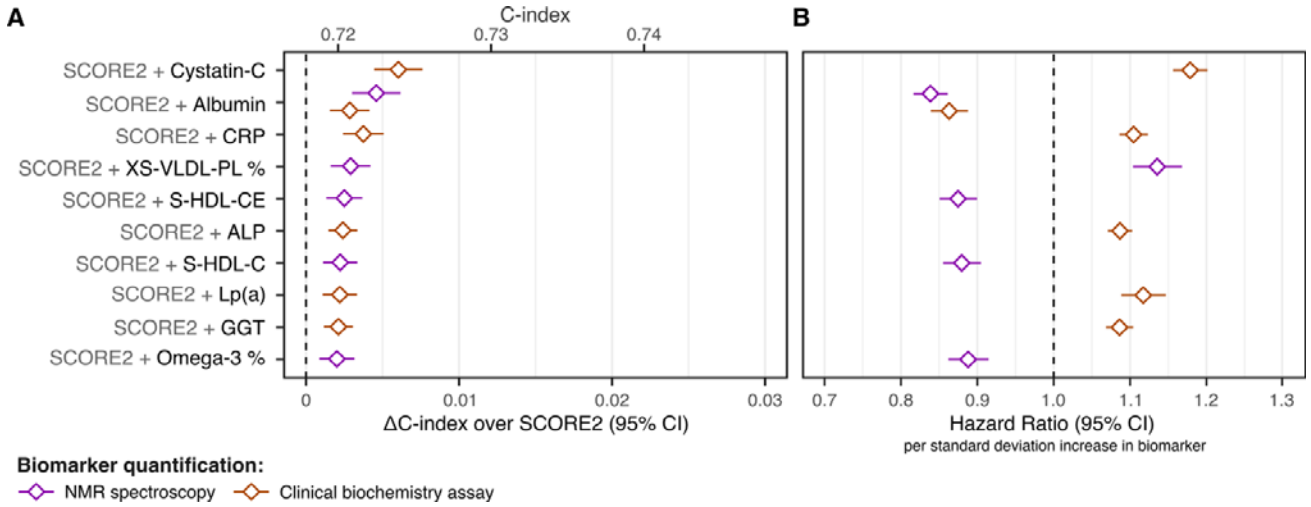

Figure 2: Incremental 10-year CVD risk discrimination over SCORE2 for individual biomarkers
A) Change in C-index (ΔC-index) relative to SCORE2 for the top 10 biomarkers by ΔC-index. The ΔC-index were assessed for each biomarker and quantification method separately (277 tests total; **Table S3**) in sex-stratified Cox proportional hazards regression in the subset of the 168,517 study participants with non-missing biomarker concentrations for each test. Among the top 10 biomarker shown, only albumin concentrations were quantified by both clinical biochemistry assays (orange) and high-throughput NMR spectroscopy (purple). 95% confidence intervals for the ΔC-index were estimated via a bootstrap sampling procedure with 1000 bootstraps (**Methods**). **B)** Hazard ratios per standard deviation increase in the respective biomarker concentration in the sex-stratified Cox proportional hazards model fit with SCORE2 as an offset term.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

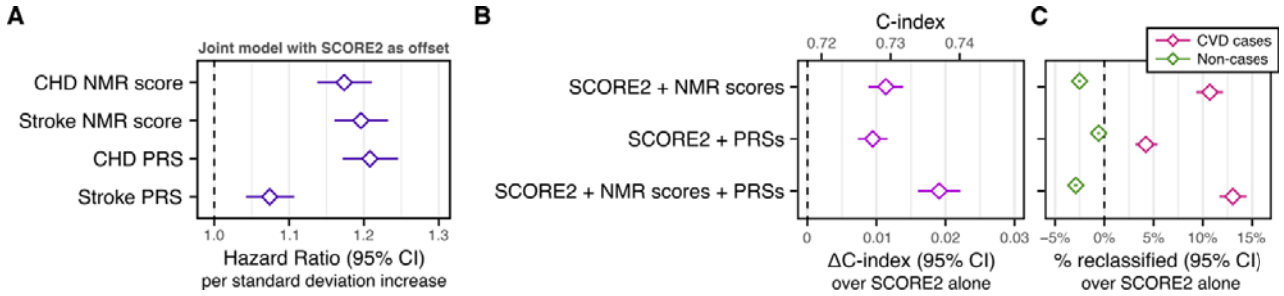

Figure 3: Incremental 10-year CVD risk discrimination over SCORE2 for NMR scores and PRSs
A) Independent associations between NMR scores and PRSs with 10-year CVD risk. Sex-stratified hazard ratios for each NMR score and PRS were fit in multivariable Cox proportional hazards regression with SCORE2 as an offset term in the 168,517 study participants (5,096 incident CVD cases). Hazard ratios and 95% confidence intervals are detailed in **Table S6A**. **B)** Incremental improvement in 10-year CVD risk discrimination over SCORE2 alone (ΔC-index) for models combining NMR scores and PRSs with SCORE2 (**Table S7**; **Methods**). 95% confidence intervals for the ΔC-index were estimated via a bootstrap sampling procedure with 1000 bootstraps (**Methods**). C-indices, ΔC-indices, and 95% confidence intervals are detailed in **Table S8**. **C)** Categorical net reclassification improvement (NRI) index relative to SCORE2 when stratifying the 168,517 study participants into low-, medium-, and high-risk categories based on their absolute 10-year CVD risk predicted by each model. Risk categories were defined using a risk-thresholds recommended by the European Society of Cardiology (ESC) 2021 guidelines for clinical decision making on initiating risk factor treatment for CVD prevention (**Methods**). Numbers allocated to each risk category by each model are detailed in **Table S9**. % reclassified: net % of cases that were correctly reclassified into a higher risk category (pink) or net % of non-cases that were correctly reclassified into a lower risk category (green) when comparing the given model to SCORE2. 95% confidence intervals were estimated via a bootstrap sampling procedure with 1000 bootstraps (**Methods**). Categorical NRI details are provided in **Table S10**.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

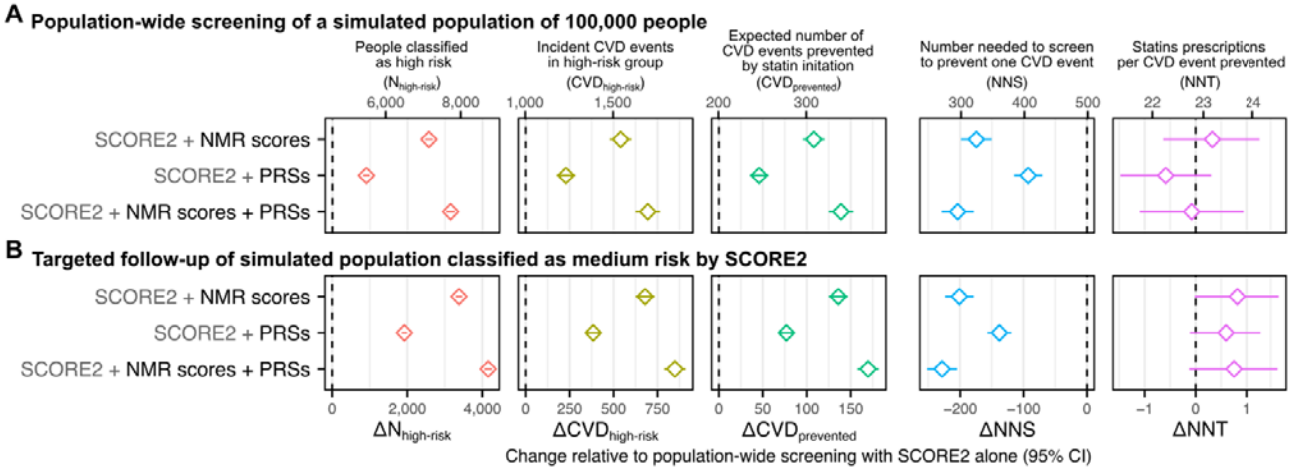

Figure 4: Incremental value for CVD prevention in simulated population of 100,000 UK adults
A) Incremental value of NMR scores and/or PRSs for CVD prevention over SCORE2 alone in population-wide screening of a simulate population of 100,000 UK adults (6,636 incident CVD cases) (**Methods**). **B)** Incremental value of NMR scores and/or PRSs for CVD prevention when used for targeted risk-reclassification of the SCORE2-classified medium-risk individuals using models with NMR scores and/or PRSs (**Methods**). **A–B)** From left to right each plot shows, relative to population-wide screening with SCORE2 alone, the change in: (1) the number of people classified as high risk ($\Delta N_{high-risk}$), who based on their risk thresholds would be recommended for risk factor treatment by the ESC 2021 guidelines for CVD prevention[3]; (2) the number of expected incident CVD events amongst those classified as high-risk ($\Delta CVD_{high-risk}$); (3) the expected number of events prevented with initiation of statins in those classified as high risk ($CVD_{prevented}$), assuming a 20% reduction in 10-year CVD risk[31]; (4) the number of people needed to screen to prevent one CVD event (ΔNNS); and (5) the number of statins prescribed per CVD event prevented (ΔNNT). 95% confidence intervals were estimated via a bootstrap sampling procedure with 1000 bootstraps (**Methods**). Point estimates and 95% confidence intervals are detailed in **Table S12**.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.31.23297859; this version posted September 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Tables

Table 1: Cohort characteristics

SBP: systolic blood pressure. HDL: High density lipoprotein. LDL: Low density lipoprotein. sd: standard deviations. IQR: interquartile range.

Baseline characteristics	**Males**	**Females**	**Total**
Number of participants	72,441	96,076	168,517
Age (years), mean (SD)	55 (8.2)	55 (7.9)	55 (8.0)
Cardiovascular risk factors			
Current smokers, n (%)	8,867 (12.2%)	8,232 (8.6%)	17,099 (10.1%)
SBP (mmHg), mean (SD)	140 (17)	134 (19)	137 (19)
Total cholesterol (mmol/L), mean (SD)	5.8 (1.0)	6.0 (1.1)	5.9 (1.1)
HDL cholesterol (mmol/L), mean (SD)	1.3 (0.3)	1.6 (0.4)	1.5 (0.4)
SCORE2 10-year CVD risk (%), mean (SD)	5.6% (3.0%)	3.3% (2.2%)	4.3% (2.8%)
Incident CVD within 10 years of follow-up			
Events, n (%)	3,208 (4.4%)	1,888 (2.0%)	5,096 (3.0%)
Years to first event, median (IQR)	6.1 (3.5, 8.3)	6.7 (4.1, 8.6)	6.3 (3.8, 8.4)
Fatal events, n (%)	661 (20.6%)	283 (15.0%)	944 (18.5%)
Event is primary, n (% of events)	2,442 (76.1%)	1,252 (66.3%)	3,694 (72.5%)
Non-CVD mortality within 10 years of follow-up			
Events, n (%)	2,307 (3.2%)	2,233 (2.3%)	4,540 (2.7%)
Years to fatal event, median (IQR)	6 (3.6, 8.3)	6.1 (3.7, 8.1)	6.1 (3.7, 8.2)
Lost to follow-up before 10 years			
Events, n (%)	1,208 (1.7%)	1,637 (1.7%)	2,845 (1.7%)
Maximum follow-up in Welsh hospital records, n (% lost to follow-up)	1,044 (86.4%)	1,415 (86.4%)	2,459 (86.4%)
Years until end of Welsh hospital record linkage, median (IQR)	9.8 (8.6, 9.9)	9.8 (9.3, 9.9)	9.8 (8.8, 9.9)
Other reason (e.g., left UK), n (% lost to follow-up)	164 (13.6%)	222 (13.6%)	386 (13.6%)
Years until lost to follow-up, median (IQR)	4.6 (4.0, 5.5)	4.4 (3.2, 5.4)	4.6 (3.6, 5.4)