

Integrated clinical risk prediction of type 2 diabetes with a multifactorial polygenic risk score

Scott C. Ritchie^{1-6,*}, Henry J. Taylor^{3,4,7}, Yujian Liang⁸, Hasanga D. Manikpurage^{1,3,4}, Lisa Pennells^{3,4}, Carles Foguet^{1,3,4}, Gad Abraham^{2,9}, Joel T. Gibson^{1,3,4}, Xilin Jiang^{1,3-6,10}, Yang Liu¹⁻⁴, Yu Xu^{1,3-6}, Lois G. Kim^{3,4,11}, Anubha Mahajan^{12,13}, Mark I. McCarthy^{12,13}, Stephen Kaptoge^{3,4}, Samuel A Lambert^{1,3,4,6,14}, Angela Wood^{3-6,11,15}, Xueling Sim⁸, Francis S. Collins⁷, Joshua C. Denny^{7,16}, John Danesh^{3-6,11,17}, Adam S. Butterworth^{3-6,11}, Emanuele Di Angelantonio^{3-6,11,18}, Michael Inouye^{1-6,*}

¹Cambridge Baker Systems Genomics Initiative, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

²Cambridge Baker Systems Genomics Initiative, Baker Heart and Diabetes Institute, Melbourne, Victoria, Australia

³British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

⁴Victor Phillip Dahdaleh Heart and Lung Research Institute, University of Cambridge, Cambridge, UK

⁵British Heart Foundation Centre of Research Excellence, University of Cambridge, Cambridge, UK

⁶Health Data Research UK Cambridge, Wellcome Genome Campus and University of Cambridge, Cambridge, UK

⁷Center for Precision Health Research, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

⁸Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore, Singapore

⁹Department of Clinical Pathology, University of Melbourne, Parkville, Victoria, Australia

¹⁰Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, US

¹¹National Institute for Health and Care Research Blood and Transplant Research Unit in Donor Health and Behaviour, University of Cambridge, Cambridge, UK

¹²Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK, OX3 7BN

¹³OMNI Human Genetics, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA

¹⁴European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK

¹⁵Cambridge Centre of Artificial Intelligence in Medicine, University of Cambridge, Cambridge, UK

¹⁶All of Us Research Program, National Institutes of Health, Bethesda, MD, USA

¹⁷Department of Human Genetics, Wellcome Sanger Institute, Hinxton, UK

¹⁸Health Data Science Research Centre, Human Technopole, Milan, Italy

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

Abstract

Combining information from multiple GWASs for a disease and its risk factors has proven a powerful approach for development of polygenic risk scores (PRSs). This may be particularly useful for type 2 diabetes (T2D), a highly polygenic and heterogeneous disease where the additional predictive value of a PRS is unclear. Here, we use a meta-scoring approach to develop a metaPRS for T2D that incorporated genome-wide associations from both European and non-European genetic ancestries and T2D risk factors. We evaluated the performance of this metaPRS and benchmarked it against existing genome-wide PRS in 620,059 participants and 50,572 T2D cases amongst six diverse genetic ancestries from UK Biobank, INTERVAL, the All of Us Research Program, and the Singapore Multi-Ethnic Cohort. We show that our metaPRS was the most powerful PRS for predicting T2D in European population-based cohorts and had

43 comparable performance to the top ancestry-specific PRS, highlighting its transferability. In UK Biobank,
44 we show the metaPRS had stronger predictive power for 10-year risk than all individual risk factors apart
45 from BMI and biomarkers of dysglycemia. The metaPRS modestly improved T2D risk stratification of
46 QDiabetes risk scores for 10-year risk prediction, particularly when prioritising individuals for blood tests
47 of dysglycemia. Overall, we present a highly predictive and transferrable PRS for T2D and demonstrate
48 that the potential for PRS to incrementally improve T2D risk prediction when incorporated into UK
49 guideline-recommended screening and risk prediction with a clinical risk score.

50 Introduction

51 The global prevalence of type 2 diabetes (T2D) has quadrupled in the last 30 years, affecting approximately
52 508 million adults globally in 2021, with prevalence expected to increase a further 60% by 2050^{1,2}. The risk
53 of developing T2D is determined by a complex interplay of lifestyle, environmental, and genetic factors³.
54 Genetic studies have estimated the heritability of T2D to be 69% among adults 35–60 years of age⁴ and
55 genome-wide association studies (GWAS) have thus far identified 611 genomic loci associated with T2D
56 risk⁵.

57 Polygenic risk scores (PRS) have emerged as a powerful tool for aggregating genomic associations into a
58 single score quantifying an individual's genetic predisposition to disease^{6–8}. As they are based on the
59 germline genome, which is stable throughout the life-course, a key advantage of PRS in comparison to
60 other risk factors is early risk prediction. PRS can be used to predict disease risk at any point in a lifetime,
61 including decades before lifestyle and environmental risk factors for T2D manifest, and it has been widely
62 shown that risk prediction models can improve their ability to predict risk when PRS are integrated with
63 commonly used risk predictors^{6–8}. Numerous T2D PRS have been constructed to date, with 134 PRS from
64 40 studies published in the Polygenic Score (PGS) Catalog⁹ at the time of writing.

65 Most PRS have been developed using a single source of GWAS summary statistics. However, substantial
66 improvements in prediction have been found by studies combining multiple sources of GWAS summary
67 statistics during PRS development^{10–13}. Improvements in PRS performance have been obtained both by
68 combining information from multiple GWASs or PRSs from the disease of interest¹⁰ as well as by
69 incorporating information from GWASs for disease risk factors^{11–13}. Yet, PRS tailored specifically for T2D
70 using this strategy are currently lacking. It is unclear to what extent this will improve predictive
71 performance, transferability, and/or add value beyond existing clinical risk scores.

72 Here we utilize ancestrally diverse GWAS summary statistics from ten T2D GWAS and 34 T2D risk factor
73 GWASs to develop a PRS for T2D. This new T2D metaPRS is externally validated and compared with
74 previously published PRS in six diverse genetic ancestries from four large independent cohorts/biobanks:
75 UK Biobank^{14,15}, INTERVAL^{16,17}, the All of Us research program^{18–20}, and the Singapore Multi-Ethnic
76 Cohort²¹. We further compare the T2D metaPRS and assess its added value to conventional risk factors and
77 QDiabetes risk prediction scores²² for 10-year T2D risk prediction in UK Biobank.

78 Results

79 *Study participants*

80 A schematic of the overall study design is shown in **Figure 1**. After filtering, in total we analysed data from
81 620,059 participants, including 50,572 T2D cases, across the four study cohorts (**Methods**). Participants
82 were grouped into genetic clusters using principal components analysis and assigned ancestry labels 1KG-
83 EUR-like, 1KG-AFR-like, 1KG-AMR-like, 1KG-SAS-like, and 1KG-EAS-like based on their similarity to
84 1000 genomes (1KG) reference panel superpopulations²³ following the 2023 National Academies
85 guidelines on using population descriptors in genetics and genomics research²⁴. Importantly, these labels
86 seek to recognize (1) that genetic ancestries are distinct from and frequently do not overlap with ethnic and
87 cultural identities, (2) these groupings are defined based on genetic similarity to arbitrary sets of labelled
88 reference individuals, and (3) these groupings, while useful tools for statistical analyses, are artificial and
89 do not represent the continuum of genetic diversity that exists in the human population²⁴. Ethnic Malays in
90 the Singapore Multi-Ethnic Cohort were handled separately as their genetic ancestries are not well
91 represented by the 1KG reference panel, e.g. they do not cluster with either the 1KG EAS or SAS reference
92 populations²⁵. For consistency with the other genetic ancestry labels, here we assign the label
93 “Austronesian-like” (ASN-like) to reflect their ancestral population histories²⁵. Characteristics of each
94 genetic ancestry and cohort are described in **Table S1**.

95 *Derivation of a metaPRS for type 2 diabetes*

96 To develop the metaPRS for type 2 diabetes we split unrelated 1KG-EUR-like UK Biobank participants
97 into a PRS training dataset (N=130,816; 10,304 T2D cases) and a PRS testing dataset (N=245,117; 17,096
98 T2D cases) (**Methods, Figure 1**). To train the metaPRS, we used our previously described meta-scoring
99 approach¹¹, which leverages information from PRS trained on multiple GWAS of the target disease and its
00 risk factors (**Methods**). Summary statistics were all obtained from contemporary GWASs that did not
01 include UK Biobank participants (**Table S2**). We trained 44 PRSs to predict T2D using LDpred2²⁶ and
02 summary statistics from 10 GWAS of T2D across diverse ancestries and 34 GWAS for T2D risk factors
03 (**Figure S1, Table S3**). The 44 PRSs were subsequently combined into a single metaPRS using elasticnet
04 logistic regression²⁷ with 10-fold cross validation in the training dataset (**Figure S2, Table S4**). The T2D
05 metaPRS comprising 1.3 million SNPs is made available on the PGS Catalog⁹ with accession PGS004923.

06 *The metaPRS improves risk prediction of type 2 diabetes compared with other PRSs*

07 Using the independent 1KG-EUR-like UK Biobank testing dataset of 245,117 participants, we next
08 quantified the performance of the metaPRS for predicting prevalent T2D case status (11,080 cases) at
09 baseline and for predicting risk of incident T2D (6,016 cases from hospital episode statistics) over 10-years
10 of follow-up via survival analysis. All associations were adjusted for age, sex, and 20 genetic principal
11 components (PCs). Prevalent and incident T2D cases in UK Biobank were analysed separately due to
12 substantial differences in case identification²⁸ (**Methods**). T2D is primarily diagnosed by primary care
13 physicians, however less than half the participants had linked primary care records available. Prevalent
14 cases were identified using a combination of self-reported diabetes diagnoses, prescription medication
15 usage, and retrospective hospital records, whereas identification of incident T2D cases relied solely on
16 hospital records. The metaPRS was associated with prevalent T2D with an odds ratio of 2.30 (95% CI:
17 2.26–2.35) per standard deviation of the metaPRS, with an area under the receiver-operating characteristic
18 curve (AUC) of 0.777 (95% CI: 0.772–0.781). The metaPRS was associated with incident T2D with a
19 hazard ratio (HR) of 1.80 (95% CI: 1.75–1.85) per standard deviation of the metaPRS, with a C-index of

20 0.719 (95% CI: 0.713–0.725). When compared to other PRS (**Table S6**) that could be evaluated in 1KG-
21 EUR-like UK Biobank samples (i.e., did not include UK Biobank GWAS in PRS training), the metaPRS
22 had the strongest associations with both prevalent and incident T2D (**Figure 2 A–B, Table S5**).

23 To replicate the metaPRS and compare to contemporary PRS trained using 1KG-EUR-like UK Biobank
24 GWAS, we analysed data from a combined 1KG-EUR-like 147,962 participants (10,795 T2D cases) from
25 the INTERVAL cohort^{16,17} and the All of Us research program^{18–20} (**Figure 2C–D, Table S5**). In
26 INTERVAL, the metaPRS was associated with incident T2D with a HR of 2.07 (95% CI: 1.92–2.23) and a
27 C-index of 0.774 (95% CI: 0.758–0.790). In All of Us, the metaPRS was associated with prevalent T2D
28 with an odds ratio of 1.92 (95% CI: 1.88–1.97) and an AUC of 0.737 (95% CI: 0.732–0.742). Importantly,
29 when compared to other genome-wide PRSs (**Table S6**), the metaPRS was the strongest predictor of T2D
30 in both cohorts. In both cohorts the second strongest PRS was that of Mars *et al.* 2022 (PGS002771)²⁹,
31 which had a HR of 2.03 (95% CI: 1.88–2.18) and C-index of 0.772 (95% CI: 0.756–0.788) in INTERVAL
32 and an odds ratio of 1.89 (95% CI: 1.85–1.94) and AUC of 0.735 (95% CI: 0.730–0.740) in All of Us.
33 Furthermore, the relative performance of PRSs was remarkably consistent across both INTERVAL and All
34 of Us (**Figure 2C–D**).

35 ***Transferability of the metaPRS across diverse genetic ancestries***

36 To assess the transferability of the metaPRS and other T2D PRS beyond 1KG-EUR-like genetic ancestries,
37 we analysed data from a combined 96,164 participants (12,377 T2D cases) clustering into five genetic
38 ancestries (1KG-AFR-like, 1KG-AMR-like, 1KG-SAS-like, 1KG-EAS-like, and ASN-like) from the UK
39 Biobank^{16,17}, the All of Us research program^{18–20}, and the Singapore Multi-Ethnic Cohort²¹. As expected,
40 we observed considerable heterogeneity in both absolute and relative strength of associations of PRS across
41 genetic ancestries and cohorts (**Table S7**). Notably, no single PRS emerged as the most predictive, even
42 within any given genetic ancestry group: the top PRS was both ancestry and cohort specific (**Figure 3**).
43 When comparing relative effect sizes across cohorts and genetic ancestries, four PRS emerged as the most
44 consistent top performers: our metaPRS, along with PRSs from Huerta-Chagoya *et al.* 2023 (weighted sum
45 of PGS003443, PGS003444, and PGS003445; **Methods**)³⁰, Shim *et al.* 2023 (PGS003867)³¹, and Mars *et*
46 *al.* 2022 (PGS002771)²⁹ (**Figure 4**). As expected^{32,33}, the predictive power of all tested PRSs weakened as
47 genetic ancestries diverged from 1KG-EUR-like: from a maximum odds ratio of 2.30 (95% CI: 2.26–2.35)
48 for any PRS in 1KG-EUR-like samples (**Table S6**), to 1.91 (95% CI: 1.78–2.05) in 1KG-SAS-like, 1.90
49 (95% CI: 1.48–2.45) in 1KG-EAS-like, 1.77 (95% CI: 1.71–1.84) in 1KG-AMR-like, 1.71 (95% CI: 1.43–
50 2.06) in ASN-like, and 1.37 (95% CI: 1.33–1.41) in 1KG-AFR-like (**Table S7**).

51 ***Comparison to conventional risk factors and QDiabetes risk scores***

52 We compared the metaPRS to established T2D risk factors and QDiabetes²², a 10-year T2D risk prediction
53 score recommended to clinicians by the UK's National Institute for Health and Care Excellence (NICE)
54 guidelines for T2D prevention³⁴ and National Health Service (NHS) health check best practice guidance³⁵.
55 For this, we utilize a subset of 190,293 1KG-EUR-like UK Biobank participants (4,064 incident T2D cases)
56 with risk factor information required for QDiabetes risk score calculation (**Figure 5A, Table S8**). The C-
57 index for the metaPRS (C-index: 0.716; 95% CI: 0.708–0.723) was larger than for all individual risk
58 factors—including family history (C-index: 0.687; 95% CI: 0.679–0.695)—except for body mass index
59 (BMI) (C-index: 0.780; 95% CI: 0.773–0.787) and glycated haemoglobin (HbA1c) (C-index: 0.826; 95%
60 CI: 0.819–0.833).

61 When added to QDiabetes and its model variants (A, B and C), the metaPRS significantly improved 10-
62 year T2D risk prediction (**Figure 5A, Table S8**). The basic QDiabetes score (model A), which incorporates
63 all risk factors that do not require taking a blood sample (**Methods**), had a C-index of 0.808 (95% CI:
64 0.802–0.814). Adding the metaPRS to QDiabetes model A increased the C-index by 0.016 (95% CI: 0.013–
65 0.019; P-value: 6×10^{-34}) yielding a total C-index of 0.824 (95% CI: 0.818–0.830). The QDiabetes score
66 incorporating fasting glucose (model B) had a C-index of 0.773 (95% CI: 0.765–0.781). The substantially
67 lower C-index for model B compared to model A of the QDiabetes score can be explained by the non-
68 fasting status of UK Biobank participants, which would lead to overestimation of risk for those who have
69 recently eaten (i.e. have higher glucose). Adding the metaPRS to QDiabetes model B led to a similar
70 increase in C-index compared to model A, with a Δ C-index of 0.019 (95% CI: 0.015–0.022; P-value:
71 6×10^{-27}), yielding a total C-index of 0.790 (95% CI: 0.783–0.798). The QDiabetes score incorporating
72 HbA1c (model C) had the largest C-index of 0.866 (95% CI: 0.861–0.872). Addition of the metaPRS to this
73 model led to a smaller, but still statistically significant, increase in C-index (Δ C-index: 0.005; 95% CI:
74 0.004–0.006; P-value: 4×10^{-15}), yielding a total C-index of 0.871 (95% CI: 0.866–0.877).

75 When incorporating the metaPRS into absolute risk predictions made by QDiabetes risk scores (**Figure 5B,**
76 **Supplementary Methods**) we observed significant improvements in risk stratification at varying risk
77 thresholds (5%, 10%, 15%) for all QDiabetes model variants (**Figure 5C, Table S9**). Consistent with the
78 above, improvements in risk stratification were strongest when adding the metaPRS to QDiabetes score
79 model A. Using a threshold of 10% absolute risk, we observed a net 8.02% improvement (95% CI: 6.83%–
80 9.22%; P-value: 1×10^{-39}) in classification of future incident T2D cases as high risk when adding the
81 metaPRS to QDiabetes score model A. Among the 4,064 incident T2D cases, the number of cases correctly
82 identified as high risk increased from 2,509 to 2,853 (an additional 11.52% of cases correctly identified as
83 high risk) with 142 cases (3.50%) incorrectly reclassified as low risk (net improvement of 8.02%). Net
84 improvements in risk stratification of cases using a 10% risk threshold were 6.92% (95% CI: 5.88%–
85 7.96%; P-value: 6×10^{-39}) for QDiabetes model B and 5.07% (95% CI: 4.13%–6.02%; P-value: 8×10^{-26}) for
86 QDiabetes model C respectively. Modest, but statistically significant, increases in the number of non-cases
87 incorrectly classified as high risk were also observed at all tested risk thresholds (**Figure 5C, Table S9**).
88 With the 10% risk threshold, the net number of non-cases incorrectly classified as high-risk increased by
89 3.01% (95% CI: 2.87%–3.14%; P-value $< 1 \times 10^{-300}$) when adding the metaPRS to QDiabetes model A, by
90 1.56% (95% CI: 1.46%–1.66%; P-value: 3×10^{-217}) when adding the metaPRS to QDiabetes model B, and
91 by 1.67% (95% CI: 1.59%–1.76%; P-value $< 1 \times 10^{-300}$) when adding the metaPRS to QDiabetes model C.

92 **Improvements in risk stratification and screening following UK guidelines**

93 NICE guidelines for T2D prevention³⁴ and NHS health check best practice guidance³⁵ recommend using
94 the basic QDiabetes score (model A) to prioritize potential high risk individuals (>5.6% risk) for fasting
95 glucose or HbA1c blood tests, which can then be used subsequently to enhance risk prediction via
96 QDiabetes models B and C²². When modifying the initial screening step by adding the metaPRS to
97 QDiabetes model A, the number of participants with >5.6% risk prioritized for blood tests increased from
98 75,153 (3,396 incident T2D cases) to 77,495 (3,517 incident T2D cases); yielding a similar number to
99 follow-up with blood tests per T2D event (number needed to screen; NNS) of 22.13 vs. 22.03 respectively
00 (Δ NNS: -0.10 , 95% CI: -0.33 – 0.14 , P-value: 0.14). Net improvements in risk stratification of T2D cases
01 after applying QDiabetes model B or C to these prioritized individuals (**Figure 5D**) of 4%–7% were
02 observed (**Figure 5D, Table S10**), similar to those observed above when systematically assessing all
03 participants with QDiabetes models B or C. Likewise, a modest but statistically significant increase in the

04 net number of non-cases incorrectly classified as high risk of 1%–2% was also observed when
05 incorporating metaPRS into both stages of the guideline-recommended screening procedure (**Figure 5D**,
06 **Table S10**).

07 When applying the QDiabetes author-recommended²² risk-threshold of 14.3% on QDiabetes model C after
08 using QDiabetes model A to prioritize individuals for HbA1c measurement (**Table S10**), a total of 10,745
09 participants (5.6%) were classified as high-risk, including 1,906 of the future T2D cases (46.9%). When
10 adding the metaPRS to both the initial screening with QDiabetes model A and subsequent risk prediction
11 with QDiabetes model C, these increased to a total of 13,564 participants (7.1%) and 2,167 cases (53.3%)
12 classified as high-risk, yielding a net absolute gain in case classification of 6.41% (95% CI: 5.43%–7.38%;
13 P-value: 7×10^{-38}) and a net increase in the number of non-cases incorrectly classified as high-risk of 1.37%
14 (95% CI: 1.30%–1.45%; P-value: 2×10^{-279}). When considering the ratio of total interventions
15 recommended per T2D case among those at high-risk (number needed to treat; NNT), a modest but
16 statistically significant increase in NNT from 5.64 to 6.26 (Δ NNT: 0.62, 95% CI: 0.50–0.74, P-value:
17 2×10^{-24}) was observed.

18 Discussion

19 In this study, we developed a PRS for T2D based on summary statistics from 44 GWASs for T2D and its
20 risk factors. We quantified the predictive power of the T2D metaPRS by performing the broadest
21 benchmarking of genome-wide T2D PRS to date (i.e. over half a million participants from six diverse
22 genetic ancestry groups from four population-based cohorts from the UK, US, and Singapore). In
23 benchmarking, we demonstrated that the T2D metaPRS is the most predictive PRS for T2D in European
24 genetic ancestries and had comparable performance to the top ancestry- and cohort- specific PRS,
25 highlighting its transferability. We further compared the T2D metaPRS to established non-genetic risk
26 factors and quantified its added value in combination with 10-year risk prediction scores in the context of
27 current UK guidelines^{34,35}.

28 Transferability is a major challenge for PRS development and a barrier to PRS utility and equitable clinical
29 application. Data availability has meant PRS have predominantly been developed using GWAS from
30 European genetic ancestries^{36,37}. This risks exacerbating health disparities as PRSs have shown reduced
31 predictive performance in individuals of non-European and complex genetic ancestries^{32,33}, whom make up
32 the majority of the global population. In our systematic benchmarking, the majority of PRSs showed
33 reduced performance relative to other PRSs when tested outside of the genetic ancestries used in their
34 development, with worsening performance as the continuum of genetic ancestries diverged. PRSs
35 developed with ancestry-specific data were also frequently out-performed by out-of-ancestry or multi-
36 ancestry PRSs, likely due to differences in available sample sizes. Surprisingly, we also found that within
37 non-1KG-EUR-like ancestries there was no single maximally predictive PRS in each ancestry group; top
38 PRSs were both ancestry- and cohort- specific. Moreover, the absolute magnitude of odds ratios weakened
39 as genetic ancestries diverged from 1KG-EUR-like, including for PRSs developed in non-1KG-EUR-like
40 samples using non-1KG-EUR-like GWAS summary statistics. Our results add to the body of evidence
41 highlighting the need for recruitment of participants from globally and genetically diverse ancestries as part
42 of large biobanks and cohorts, including and beyond high-income countries^{36,37}. Our results further

43 highlight that the relative performance of PRSs can also differ considerably between cohorts even within
44 the same genetic ancestry group, suggesting heterogeneity in environment or phenotype definition can also
45 impact transferability³⁸.

46 When compared to established risk factors, the metaPRS had stronger predictive power for 10-year risk
47 than all conventional risk factors, apart from BMI and biomarkers of dysglycemia, and captured residual
48 risk not quantified by these risk factors. The metaPRS also provided a modest, but statistically significant,
49 improvement over the QDiabetes risk scores combining established risk factors for both risk discrimination
50 and risk stratification at varying risk thresholds. Improvements from the metaPRS were less than those
51 from measurement and addition of blood biomarkers of dysglycemia (e.g. HbA1c), consistent with a
52 previous study of T2D PRSs in a population cohort of British Pakistanis and Bangladeshis³⁹.

53 However, blood tests for dysglycemia are not routinely conducted in asymptomatic individuals; UK
54 guidelines recommend using readily available lifestyle and medical history information to identify high-risk
55 individuals (i.e. using QDiabetes model A) for follow-up testing of fasting glucose or HbA1c blood tests
56 for T2D diagnosis^{34,35}. QDiabetes models B and C have been developed with a view to enhancing risk
57 prediction in those found to not be diabetic after follow-up blood tests²². PRSs may one day be included
58 among readily available factors for risk screening as they require a one-off blood sample for genotyping
59 which may be obtained at any time during a person's life, for example via initiatives like the UK Newborn
60 Genome Screening Programme⁴⁰.

61 Here, we show that, if genotypes are already available, the metaPRS can enhance this initial screening step:
62 increasing from 84% to 87% the number of future diabetics revealed to be at elevated risk by blood testing,
63 with a similar NNS of ~22. The metaPRS also improved subsequent risk prediction, increasing the number
64 of T2D cases classified as high-risk by 6.4% when used alongside HbA1c with QDiabetes model C, with a
65 modest increase in the NNT by 0.62 from 5.64 to 6.26.

66 Our study has limitations. Firstly, while we utilized multi-ancestry GWAS summary statistics whenever
67 available, a 1KG-EUR-like cohort was used for model training. As large and diverse training sets with T2D
68 outcomes become available, future studies can utilize highly diverse cohorts for PRS training alongside
69 genetically diverse GWAS, thus resulting in more powerful, and more portable, T2D PRS. The UK
70 Biobank samples used for analysis of established risk factors and QDiabetes risk scores differs from the
71 wider UK population in several key respects. UK Biobank participants are healthier than the general UK
72 population¹⁴ and thus prevalence will be higher for dichotomous risk factors (e.g. medical history) and
73 distributions will be narrower and/or shifted for continuous risk factors (e.g. BMI). Participants were also
74 non-fasting, confounding risk predictions made by QDiabetes model B which relies on fasting glucose²².
75 We also expect risk stratification to substantially differ from the general population, as T2D is primarily
76 diagnosed by primary care physicians, whereas incident T2D case identification in UK Biobank relied on
77 hospital records since less than half the cohort has linked primary care records available. Likewise, the
78 high-risk sub-population assessed subsequent to screening did not exclude those with undiagnosed T2D as
79 diagnosis was not possible as this required fasting glucose or repeated HbA1c measures^{34,35}. Our analyses
80 were also restricted to genetically homogenous 1KG-EUR-like participants self-reporting as White British
81 due to the much smaller sample sizes available for other ethnic and ancestry groups and expected
82 confounding from population stratification that would be introduced if assessing the metaPRS in the pooled
83 multi-ethnic and multi-ancestry sample^{41,42}. Despite these limitations, our analyses nevertheless indicate
84 that incorporating our T2D metaPRS could modestly improve screening and risk stratification. Further

85 studies in cohorts representative of the general UK population eligible for screening would be needed to
86 accurately quantify the precise added benefits of PRS to screening and 10-year risk prediction.

87 Overall, our study presents a new T2D PRS that is highly predictive across diverse genetic ancestries and
88 cohorts, improves risk prediction when added to established risk factors in clinical risk scores for 10-year
89 risk prediction of T2D, and has the potential to improve screening practices in the UK.

90 **Methods**

91 ***UK Biobank cohort***

92 UK Biobank is a cohort of approximately 500,000 individuals with deep phenotyping, imputed genotypes,
93 and electronic health record linkage^{14,15}. Participants were members of the general UK population between
94 40 to 69 years of age identified and recruited through primary care lists and who accepted an invitation to
95 attend one of 22 assessment centres across the UK between 2006 and 2010. Ethics were approved by the
96 North West Multi-centre Research Ethics Committee (MREC) in the UK, and this study was undertaken
97 under UK Biobank project #7439. Participants gave informed and broad consent for health-related research.

98 Recruitment included standardized questionnaires on socio-demographics, ethnicity, lifestyle factors, and
99 personal- and family- medical history. Physical measurements including height, weight, body fat
00 percentage, and systolic blood pressure were also taken at assessment, and blood samples taken for
01 genotyping and quantification of molecular phenotypes. Participants were also linked to national death and
02 cancer registries as well as hospital episode statistics. Participants were genotyped on UK BiLEVE arrays
03 and UK Biobank Axiom arrays and imputed to the 1000 genomes, UK10K, and Haplotype Reference
04 Consortium panels⁴³ using human genome build GRCh37¹⁵. Participants were filtered to a set of unrelated
05 individuals (kinship < 0.0884) identified using kinship estimates⁴⁴ supplied by UK Biobank¹⁵.

06 For the primary analyses of metaPRS derivation and validation we restricted analyses to the “White
07 British” cohort defined by UK Biobank based on self-reported ethnicity (data-field #21000) combined with
08 genetic principal components¹⁵. For consistency with other study cohorts and following the 2023 National
09 Academies guidelines on using population descriptors in genetics and genomics research²⁴ we assigned this
10 group the genetic ancestries label 1KG-EUR-like. For analyses assessing PRS transferability we similarly
11 defined genetically homogeneous populations using a combination of self-reported ethnicity and projection
12 of genetic principle components to 1KG reference ancestral superpopulations²³ using the KING software⁴⁴.
13 Participants were grouped into 1KG-SAS-like if they self-reported ethnicity as Indian, Pakistani, or
14 Bangladeshi and their KING inferred ancestry was SAS with >95% probability. Participants were grouped
15 into 1KG-EAS-like if they self-reported ethnicity as Chinese and their KING inferred ancestry was EAS
16 with >95% probability. Participants were grouped into 1KG-AFR-like if they self-reported ethnicity as
17 African, Caribbean, Black or Black British, or any other Black background and their KING inferred
18 ancestry was AFR with >95% probability.

19 As linked primary care records are only available for less than half of UK Biobank participants, prevalent
20 T2D status at baseline was adjudicated from a combination of retrospective hospital episode records, self-
21 reported history of diabetes, and baseline medication using the Eastwood *et al.* algorithms²⁸. Incident T2D
22 cases were also ascertained following the Eastwood *et al.* algorithms²⁸, on the basis of ICD-10 diagnosis

23 coding E11 in either the hospital inpatient or death registry data. Onset of incident T2D was determined as
24 the midpoint between the first hospital or death record with an ICD-10 E11 coding and the previous T2D-
25 free record (hospital record without ICD-10 E11 coding or baseline assessment)²⁸. Follow-up for incident
26 T2D events was truncated on 1st February 2020 to preclude potential confounding from SARS-CoV2
27 infection, exposure, or behavioural or environmental changes from pandemic lockdowns on metaPRS
28 training.

29 ***INTERVAL cohort***

30 INTERVAL is a cohort of approximately 50,000 participants nested within a randomized trial studying the
31 safety of varying frequency of blood donation^{16,17}. Participants were blood donors aged 18 years and older
32 (median 44 years of age; 49% women) recruited between June 2012 and June 2014 from 25 centres across
33 England. Blood samples were taken at assessment and participants consented for broad health-related
34 research^{16,17}. Electronic health record linkage was available for a maximum of 11.1 years of follow-up
35 (median 10.4 years). In total there were 38,949 participants who were diabetes free at baseline assessment
36 with linked imputed genotypes and electronic health records. Ethics were approved for this study by the
37 National Research Ethics Service (11/EE/0538).

38 Participants were genotyped using the Affymetrix UK Biobank Axiom arrays and imputed to the UK10K
39 and 1000 Genomes panel using human genome build GRCh37. Notably, a key step in the genotype QC was
40 exclusion of samples of non-European ancestry on the basis of genotype PCs⁴⁵. For consistency with the
41 other study cohorts and following the 2023 National Academies guidelines on using population descriptors
42 in genetics and genomics research²⁴ we assigned these participants the ancestry label 1KG-EUR-like.

43 Linked electronic health records from national hospital episode statistics were summarized into 301
44 endpoints from ICD-10 diagnosis codes using CALIBER rule-based phenotyping algorithms⁴⁶
45 (<https://www.caliberresearch.org/portal>) prior to being made available to analysts. The closest matching
46 CALIBER phenotype for T2D was for any diabetes; defined using ICD-10 codes E10–E14, G59.0, G63.2,
47 H28.0, H36.0, M14.2, N08.3, or O24.0–O24.3. Participants with any diabetes history were excluded from
48 the analysis. Incident diabetes events were treated as incident T2D for the purposes of analyses, consistent
49 with the rarity of adult-onset type 1 diabetes. Onset of incident T2D was determined as the midpoint
50 between the first diabetes event and the previous diabetes-free record (hospital record without a diabetes
51 coding or baseline assessment).

52 ***All of Us research program cohort***

53 All of Us is a longitudinal cohort aiming to recruit one million participants from across the USA¹⁸. In the
54 v7 data freeze, there were approximately 206,000 participants with deep phenotyping, whole genome
55 sequencing, and electronic health record linkage¹⁹. Participants were members of the general USA
56 population ≥ 18 years of age with recruitment focused on groups underrepresented in biomedical research²⁰.
57 Research was conducted on the All of Us Researcher Workbench under the guidelines defined by the All of
58 Us Ethical Conduct of Research Policy.

59 Details of whole genome sequencing and quality control are described extensively in the All of Us
60 Genomic Research Data Quality Report C2022Q4R9 at [https://support.researchallofus.org/hc/en-
61 us/articles/4617899955092-All-of-Us-Genomic-Quality-Report](https://support.researchallofus.org/hc/en-us/articles/4617899955092-All-of-Us-Genomic-Quality-Report). Computation of kinship relatedness and
62 clustering of participants by genetic similarity to 1KG AFR, EUR, and AMR reference ancestral
63 superpopulations²³ are also described in the report. Additional downstream quality control and filtering of

64 sequence data is as described in Suzuki *et al.* 2024⁵. Briefly, related individuals were pruned to obtain a
65 maximal independent set (kinship score > 0.1), and variants were filtered to high-quality SNPs with MAF >
66 1% or MAC > 100 in at least one of the genetic ancestry clusters. SNPs with MAF < 1% that deviated from
67 Hardy-Weinberg equilibrium ($P < 1 \times 10^{-6}$) were removed. Principal components used for correction of
68 population structure were calculated in each ancestry group separately using SNPs present in the 1000
69 Genomes project phase 3 release. Samples whose sex could not be imputed from genotypes were excluded.

70 Phenotyping of T2D case and control status was performed using the PheKB algorithm
71 (<https://phekb.org/phenotype/type-2-diabetes-mellitus>) as described in Suzuki *et al.* 2024⁵. T2D cases were
72 ascertained based on a combination of hospital diagnosis codes, prescription medication, and lab results
73 from blood tests occurring prior to baseline sample assessment. Participants were considered controls if
74 they had no history of any diabetes diagnoses, T2D medication, or abnormal glucose or HbA1c lab results.
75 Participants with T1D or uncertain diabetes status were excluded from analysis.

76 ***Singapore Multi-Ethnic Cohort***

77 The Singapore Multi-Ethnic Cohort is a population-based cohort studying how genes and lifestyle influence
78 disease risk differently in participants from three major ethnic groups in Singapore: Chinese, Indian, and
79 Malay (<https://blog.nus.edu.sg/sphs/population-studies/multi-ethnic-cohort-phase-1-mec1/>)²¹. Participants
80 were recruited between 2004 and 2010 and invited for follow-up assessment between 2011–2016 (mean
81 follow-up 6.3 years). In total there were 2,871 participants with whole-genome sequencing who were
82 disease-free at baseline. Written consent was obtained from all participants, and this study was approved by
83 the National University of Singapore Institutional Review Board (reference codes: B-16-158 and N-18-
84 059).

85 Details of whole genome sequencing and quality control are as previously described by the Singapore
86 National Precision Medicine program strategy report⁴⁷. Briefly, sequencing was performed to an average
87 depth of 15x coverage. Reads were aligned with BWA-MEM v.0.7.17 and genotyped using GATK
88 v.4.0.6.0. Variants were filtered to retain VQSR-PASS and non-STAR allele variants. Samples with call
89 rate <95%, BAM cross-contamination rate >2%, BAM error rate > 1.5% were excluded. Genotypes with
90 depth coverage (DP) < 5, genotype quality (GQ) < 20, or allele balance (AB) > 0.8 were set to null, and
91 samples with abnormal ploidy excluded. Genetic variants were filtered to exclude those with robust, unified
92 test for Hardy-Weinberg equilibrium (RUTH) P-value <0.01, a variant call rate <90%, being monomorphic,
93 or having a minor allele count (MAC) <2 prior to phasing with Eagle version 2.4^{43,48}. After quality control,
94 the dataset included 39,967,216 genetic variants in 2,871 samples. Samples were clustered into three groups
95 by genetic similarity using the k-means algorithm on the first 15 genetic principal components calculated
96 on the verifyBamID2 variant panel (1KG phase 3)⁴⁹.

97 Genetic ancestry labels for each cluster were based on the majority reported ethnicity in each group, here
98 labelled as 1KG-EAS-like, 1KG-SAS-like, and ASN-like for consistency with other the study cohorts and
99 following the 2023 National Academies guidelines on using population descriptors in genetics and
00 genomics research²⁴. The ASN-like label was used here to label the genetic cluster with Malay as the
01 majority reported ethnicity, as their genetic ancestries were not well represented by either the EAS or SAS
02 super populations in the 1KG reference panel²⁵. The label Austronesian-like (ASN-like) was chosen to
03 reflect the ancestral population histories of this group²⁵.

04 Incident T2D (N=577) was ascertained as previously described⁵⁰ through a combination of linkage to
05 national healthcare records, self-reported medical history at follow-up assessment (either diagnosis from a
06 primary care physician or current diabetes medication usage), or with blood biomarker concentrations
07 indicative of diabetes following the American Diabetes Association criteria (fasting glucose ≥ 7 mmol/L or
08 HbA1c $\geq 6.5\%$ or random blood glucose ≥ 11 mmol/L)⁵¹.

09 *MetaPRS training*

10 The T2D metaPRS was trained in a subset of 130,816 UK Biobank participants in the “White British”
11 genetic ancestry cluster¹⁵. MetaPRS training comprised two key steps (**Figure 1**): (1) training of 44
12 individual component risk factor and diverse-ancestry T2D PRSs using LDpred2²⁶, and (2) training the
13 joint model combining the 44 T2D and related risk factor PRSs into a single meta-PRS using elasticnet
14 penalized logistic regression²⁷.

15 For step 1, we trained 44 PRSs for T2D using summary statistics from ten GWAS (or exome-wide
16 association studies) for T2D across diverse ancestries and 34 GWAS for T2D risk factors (**Table S2**). To
17 prevent overfitting, we selected contemporary GWAS that did not include UK Biobank participants⁵². The
18 selected GWAS also did not include samples from any of the cohorts used for metaPRS evaluation in this
19 study. Due to computational limitations of LDpred2, summary statistics were restricted to 1.6 million
20 autosomal bi-allelic SNPs that were present in either the HapMap3 reference panel⁵³ or in the two exome-
21 wide association studies among the 44 GWASs (**Table S2**). When mapping GWAS summary statistics and
22 HapMap3 variants to UK Biobank the UCSC Genome Browser⁵⁴ liftOver tool was used to map positions
23 from GRCh36 or GRCh38 to GRCh37 as needed. SNPs were further filtered on a per-GWAS basis
24 following LDpred2 recommendations to remove variants with low power or divergent MAF between the
25 GWAS and UK Biobank. LDpred2 was used to reweight GWAS summary statistics based on the linkage-
26 disequilibrium of a subset of 11,074 UK Biobank participants enriched for T2D (1,202 cases) under
27 multiple possible parameterisations of trait polygenicity and heritability (i.e. LDpred2 infinitesimal, grid-
28 search, automatic, and lassosum models²⁶). The remaining 120,464 UK Biobank participants (9,102 T2D
29 cases) were then used to determine the optimal LDpred2 parameter choice for T2D prediction by assessing
30 the AUC of logistic regression for combined prevalent and incident T2D case status (**Figure S1, Table S3**).
31 Logistic regressions were fit adjusting for age and sex, and candidate PRSs were adjusted for 20 genetic
32 PCs and standardized prior to model fitting.

33 Elasticnet penalized logistic regression²⁷ was subsequently used in the 120,464 UK Biobank participants
34 not used for LDpred2 parameter tuning to estimate the relative contributions of the 44 PRSs to T2D
35 prediction and for deriving a single metaPRS (**Figure S2, Table S4**). The PC-adjusted 44 PRSs trained
36 above were standardised and used as predictor variables along with age and sex in the regression. A range
37 of elasticnet mixing parameters were tested (0, 0.1, 0.25, 0.5, 0.75, 0.9, and 1) with 10-fold cross-validation
38 performed for each mixing parameter to tune the respective lambda penalty. The optimal regression fit was
39 chosen as the combination of elasticnet mixing parameter and lambda penalty that had, across the 10-cross
40 validation folds, the greatest mean AUC combined prevalent and incident T2D case status.

41 Per-SNP weights for the T2D metaPRS were subsequently derived via a weighted sum; where for each
42 SNP i , the effect size was calculated as the sum of the per-SNP effect sizes γ derived from LDpred2 for
43 each PRS j multiplied by the β coefficient estimated for the PRS in the optimal elasticnet regression:

$$SNP_i^{T2D\ metaPRS} = \sum_{j=1}^{44} \beta_j \gamma_{i,j}$$

44 The T2D metaPRS comprised 1,349,896 SNPs, which we make available along with the respective weights
45 through the PGS Catalog⁹ with accession PGS004923.

46 *Assessment of PRSs for prediction of T2D risk*

47 For comparison with the metaPRS, genome-wide T2D PRS were obtained from PGS Catalog⁹ and
48 separated into two groups: PRS whose training samples included UK Biobank participants of White
49 European ancestries, and PRS whose training samples had no overlap with any of the cohorts analysed in
50 this study (**Table S6**). PRS were calculated in each cohort as the weighted sum of their SNP weights
51 multiplied by the dosages of the respective effect alleles. The Huerta-Chagoya et al. 2023 PRS was
52 calculated as the weighted sum of their three component scores deposited as PGS003443, PGS003444, and
53 PGS003445 following their formula of $0.531117 \times \text{PGS003443} + 0.5690198 \times \text{PGS003444} +$
54 $0.1465538 \times \text{PGS003445}$ after standardising each component score to have mean 0 and standard deviation 1
55 in the target cohort³⁰. Each PRS was subsequently adjusted for population structure by taking the residuals
56 of a linear regression of the PRS on a cohort-specific number of genotype PCs (20 PCs for UK Biobank,
57 INTERVAL, and All of Us).

58 Prediction of T2D case status at baseline assessment was assessed for each PRS separately using logistic
59 regression adjusting for age at baseline and sex as covariates. In UK Biobank, assessment centre was also
60 used as a covariate. PRS were standardized when fitting the regression so that reported odds ratios were per
61 standard deviation increase and comparable across PRS. 95% confidence intervals for area under the
62 receiver operating characteristic curves (AUC) were calculated using 2000 stratified bootstrap replicates⁵⁵.
63 Logistic regression was also used to assess incident T2D prediction in the Singapore Multi-Ethnic Cohort,
64 as time to T2D onset (or T2D-free survival) was not available due to the heterogeneity of incident T2D
65 ascertainment.

66 Prediction of incident T2D risk was assessed using Cox proportional hazards regression using time-in-study
67 as the time scale and adjusting for age at baseline assessment and sex as covariates. Participants with T2D
68 at baseline assessment were excluded. In UK Biobank, assessment centre was also used as a covariate. PRS
69 were standardized when fitting the regression so that reported hazard ratios were per standard deviation
70 increase and comparable across PRS. 95% confidence intervals for the Harrell's C-index were calculated
71 from the standard errors obtained using the infinitesimal jackknife method⁵⁶.

72 *Comparison to established risk factors and risk scores*

73 The metaPRS was compared to established T2D risk factors and 10-year T2D risk prediction scores
74 (QDiabetes) in a subset of 190,293 UK Biobank participants from the metaPRS testing set that were free of
75 T2D at baseline assessment and had quantified measurements for glucose and glycated haemoglobin
76 (HbA1c).

77 We compared the metaPRS to the QDiabetes scores that are recommended for 10-year T2D risk prediction
78 in the UK by the National Institute for Health and Care Excellence (NICE) guidelines in the UK²². Three
79 QDiabetes scores are recommended depending on the availability of blood samples and fasting status:
80 QDiabetes model A, which incorporates all risk factors that do not require taking a blood sample;
81 QDiabetes model B, which additionally incorporates fasting glucose; and QDiabetes model C, which
82 additionally incorporates HbA1c (but not fasting glucose). Risk factors used by all three QDiabetes models

83 were participant age, sex, BMI, smoking status, Townsend deprivation index, family history of diabetes,
84 antihypertensive medication usage, history of CVD, systematic corticosteroid medication usage, lipid
85 lowering medication usage, history of gestational diabetes, history of polycystic ovary syndrome, history of
86 learning difficulties, history of bipolar or schizophrenia disorders, and usage of 2nd generation atypical
87 antipsychotic medications. UK Biobank participants were non-fasting (median fasting time of 3 hours) so
88 non-fasting glucose was used for QDiabetes model B. Details on each risk factor definition in UK Biobank
89 are given in the **Supplementary Methods**.

90 Prediction of 10-year risk of incident T2D for each risk factor, risk score, and the metaPRS, were assessed
91 using Cox proportional hazards regression using time-in-study as the time scale and adjusting for age, sex,
92 and assessment centre. The metaPRS was adjusted for 20 genotype PCs prior to model fitting. Prior to
93 model fitting BMI, glucose, and HbA1c were log transformed, and Townsend deprivation index was
94 inverse rank normalized. All predictor variables were standardized when fitting Cox proportional hazard
95 regressions. 95% confidence intervals for the Harrell's C-index were calculated from the standard errors
96 obtained using the infinitesimal jackknife method⁵⁶.

97 Incremental improvement of the metaPRS over QDiabetes 2018 model C was assessed using multivariable
98 Cox proportional hazards regression adjusting for age and sex. The change in C-index (Δ C-index) was
99 calculated as the difference in C-index over a Cox proportional hazards regression fit for QDiabetes 2018
00 model C adjusting for age and sex. A bootstrap procedure with 1,000 bootstraps was used to estimate the
01 standard error for the Δ C-index. Bootstrap resampling was performed using methods appropriate for right-
02 censored data⁵⁷. The 95% confidence interval and two-sided P-value were computed from the bootstrap
03 standard error using the first order normal approximation method.

04 Incremental improvements in risk stratification when adding the metaPRS to QDiabetes risk scores
05 (**Supplementary Methods**) were assessed at varying risk thresholds using categorical net reclassification
06 improvement (NRI) analysis^{58,59}. Categorical NRI analysis was used to assess relative to QDiabetes risk
07 scores alone (1) the % of incident T2D cases correctly reclassified from low risk to high risk, and (2) the %
08 of non-cases correctly reclassified from high risk to low risk. Bootstrap resampling of the categorical NRI
09 analysis was performed using the `nricens` R package version 1.6, and 95% confidence intervals and P-
10 values were subsequently calculated from the bootstrap standard error using the first order normal
11 approximation method.

12 **Data Availability**

13 Data from UK Biobank are available for health-related research subject to approval from the UK Biobank
14 access committee. See <https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access> for further
15 details.

16 Data from the INTERVAL cohort can be requested by researchers for health-related research subject to
17 approval from the INTERVAL Data Access Committee. Data will be shared through an institutional data
18 sharing agreement. The INTERVAL Data Access Committee can be contacted via email at
19 helpdesk@intervalstudy.org.uk. Further information on the data access policy can be found
20 at <http://www.donorhealth-btru.nihr.ac.uk/project/bioresource>.

21 Data from All of Us are available to researchers via the All of Us research hub subject to institutional data
22 sharing agreement. For more information, see [https://allofus.nih.gov/get-involved/opportunities-
23 researchers](https://allofus.nih.gov/get-involved/opportunities-researchers).

24 Data from the Singapore Multi-Ethnic Cohort study can be requested by researchers for scientific purposes
25 through an application process at the listed website ([https://blog.nus.edu.sg/sphs/data-and-samples-
request/](https://blog.nus.edu.sg/sphs/data-and-samples-
26 request/)). Data will be shared through an institutional data sharing agreement.

27 ***Code Availability***

28 Code underlying this paper are available at https://github.com/sritchie73/T2D_metaPRS_paper. This
29 repository and specific release for this paper are permanently archived by Zenodo at
30 <https://doi.org/10.5281/zenodo.13362823>.

31 Acknowledgements

32 The authors are grateful to UK Biobank, INTERVAL, All of Us, and the Singapore Multi-ethnic cohort for
33 access to data to undertake this study.

34 Participants in the INTERVAL randomized controlled trial were recruited with the active collaboration of
35 NHS Blood and Transplant (www.nhsbt.nhs.uk), which has supported fieldwork and other elements of the
36 trial. DNA extraction and genotyping were co-funded by the National Institute for Health and Care
37 Research (NIHR), the NIHR BioResource (<http://bioresource.nihr.ac.uk>) and the NIHR Cambridge
38 Biomedical Research Centre (BRC) (no. BRC-1215-20014). The academic coordinating centre for
39 INTERVAL was supported by core funding from the NIHR Blood and Transplant Research Unit in Donor
40 Health and Genomics (no. NIHR BTRU-2014-10024), UK Medical Research Council (MRC) (no.
41 MR/L003120/1), British Heart Foundation (nos SP/09/002, RG/13/13/30194 and RG/18/13/33946) and the
42 NIHR Cambridge BRC (no. BRC-1215-20014). A complete list of the investigators and contributors to the
43 INTERVAL trial is provided in ref. 17. The academic coordinating centre thanks blood donor centre staff
44 and blood donors for participating in the INTERVAL trial.

45 The All Of Us Research Program is supported by the National Institutes of Health, Office of the Director:
46 Regional Medical Centers: 1 OT2 OD026549; 1 OT2 OD026554; 1 OT2 OD026557; 1 OT2 OD026556; 1
47 OT2 OD026550; 1 OT2 OD 026552; 1 OT2 OD026553; 1 OT2 OD026548; 1 OT2 OD026551; 1 OT2
48 OD026555; IAA #: AOD 16037; Federally Qualified Health Centers: HHSN 263201600085U; Data and
49 Research Center: 5 U2C OD023196; Biobank: 1 U24 OD023121; The Participant Center: U24 OD023176;
50 Participant Technology Systems Center: 1 U24 OD023163; Communications and Engagement: 3 OT2
51 OD023205; 3 OT2 OD023206; and Community Partners: 1 OT2 OD025277; 3 OT2 OD025315; 1 OT2
52 OD025337; 1 OT2 OD025276. In addition, the All of Us Research Program would not be possible without
53 the partnership of its participants.

54 The MEC study is supported by individual research and clinical scientist award schemes from the National
55 Medical Research Council (NMRC) and the Biomedical Research Council (BMRC) of Singapore, and
56 infrastructure funding from the Singapore Ministry of Health (Population Health Metrics and Analytics
57 PHMA), National University of Singapore and National University Health System, Singapore. The MEC
58 whole-genome sequencing data made use of data generated as part of the Singapore National Precision
59 Medicine (NPM) program funded by the Industry Alignment Fund (Pre-Positioning) (IAF-PP:
60 H17/01/a0/007).

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

66 This work was supported by core funding from the British Heart Foundation (RG/18/13/33946:
67 RG/F/23/110103), NIHR Cambridge Biomedical Research Centre (NIHR203312) [*], BHF Chair Award
68 (CH/12/2/29428) and by Health Data Research UK, which is funded by the UK Medical Research Council,
69 Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department
70 of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social

71 Care Directorates, Health and Social Care Research and Development Division (Welsh Government),
72 Public Health Agency (Northern Ireland), British Heart Foundation and the Wellcome Trust.

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

75 L.P. was supported by a Rutherford Fund Fellowship from the Medical Research Council grant
76 MR/S003746/1. X.J. was funded by British Heart Foundation (CH/12/2/29428) and Wellcome Trust
77 (227566/Z/23/Z). Y.X. and M.I. are supported by the UK Economic and Social Research Council
78 (ES/T013192/1). L. K. is funded by the NIHR BTRU in Donor Health and Behaviour (NIHR203337) and a
79 BHF Chair award (CH/12/2/29428). S.A.L. was supported by a Canadian Institutes of Health Research
80 postdoctoral fellowship (MFE-171279). F.S.C. acknowledges support from United States' National
81 Institutes of Health (NIH) grant ZIA-HG000024. J.C.D. acknowledges support from United States'
82 National Institutes of Health (NIH) grant ZIA-HG200417. J.D. holds a BHF Professorship and a NIHR
83 Senior Investigator Award. E.D.A. holds a NIHR Senior Investigator Award. M.I. is supported by the Munz
84 Chair of Cardiovascular Prediction and Prevention and the NIHR Cambridge Biomedical Research Centre
85 (NIHR203312).

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

88 **Competing Interests**

89 During the course of this project G.A. became a full-time employee of CSL Ltd. All significant
90 contributions to this study were made prior to this role and CSL Ltd had no input to the study. A.M. and
91 M.I.M are both employees of Genentech Ltd, and holders of Roche stock. Genentech Ltd was not involved
92 with the described work. J.D. serves on scientific advisory boards for AstraZeneca, Novartis, and UK
93 Biobank, and has received multiple grants from academic, charitable and industry sources outside of the
94 submitted work. A.S.B. reports institutional grants from AstraZeneca, Bayer, Biogen, BioMarin,
95 Bioverativ, Novartis, Regeneron and Sanofi. M.I. is a trustee of the Public Health Genomics (PHG)
96 Foundation, a member of the Scientific Advisory Board of Open Targets, and has research collaborations
97 with AstraZeneca, Nightingale Health and Pfizer which are unrelated to this study. The remaining authors
98 declare no competing interests.

99 **References**

- 00 1. Zhou, B. *et al.* Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based
01 studies with 4.4 million participants. *Lancet* **387**, 1513–1530 (2016).
- 02 2. Ong, K. L. *et al.* Global, regional, and national burden of diabetes from 1990 to 2021, with projections
03 of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* **402**,
04 203–234 (2023).
- 05 3. Wu, Y., Ding, Y., Tanaka, Y. & Zhang, W. Risk factors contributing to type 2 diabetes and recent
06 advances in the treatment and prevention. *Int. J. Med. Sci.* **11**, 1185–1200 (2014).
- 07 4. Almgren, P. *et al.* Heritability and familiarity of type 2 diabetes and related quantitative traits in the
08 Botnia Study. *Diabetologia* **54**, 2811–2819 (2011).
- 09 5. Suzuki, K. *et al.* Genetic drivers of heterogeneity in type 2 diabetes pathophysiology. *Nature* **627**,
10 347–357 (2024).
- 11 6. Wray, N. R. *et al.* From Basic Science to Clinical Application of Polygenic Risk Scores: A Primer.
12 *JAMA Psychiatry* **78**, 101–109 (2021).
- 13 7. McCarthy, M. I. & Mahajan, A. The value of genetic risk scores in precision medicine for diabetes.
14 *Expert Review of Precision Medicine and Drug Development* **3**, 279–281 (2018).
- 15 8. Lambert, S. A., Abraham, G. & Inouye, M. Towards clinical utility of polygenic risk scores. *Hum.*
16 *Mol. Genet.* **28**, R133–R142 (2019).
- 17 9. Lambert, S. A. *et al.* The Polygenic Score Catalog as an open database for reproducibility and
18 systematic evaluation. *Nat. Genet.* **53**, 420–425 (2021).
- 19 10. Inouye, M. *et al.* Genomic Risk Prediction of Coronary Artery Disease in 480,000 Adults:
20 Implications for Primary Prevention. *J. Am. Coll. Cardiol.* **72**, 1883–1893 (2018).
- 21 11. Abraham, G. *et al.* Genomic risk score offers predictive performance comparable to clinical risk
22 factors for ischaemic stroke. *Nat. Commun.* **10**, 5819 (2019).

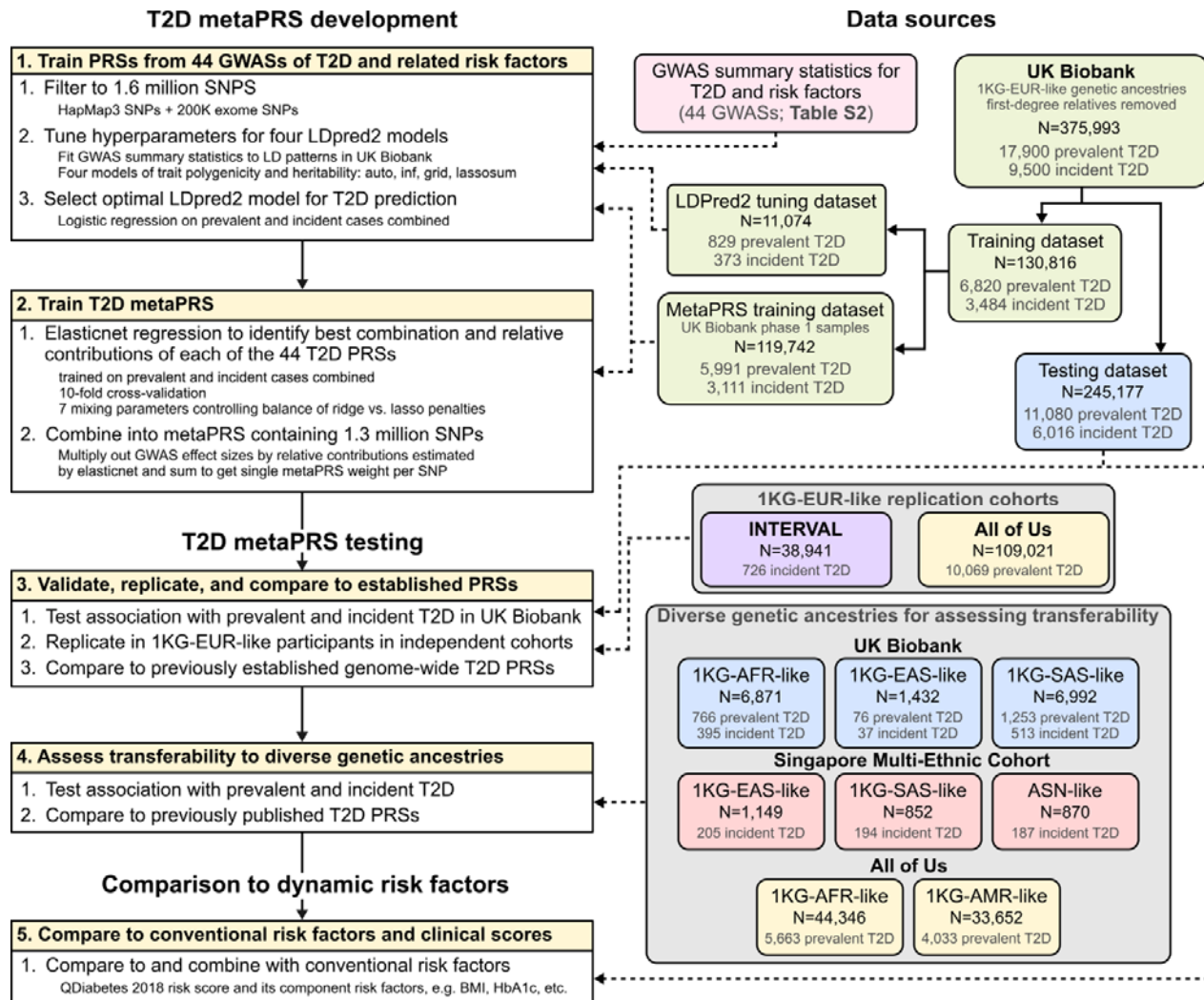
- 23 12. Truong, B. *et al.* Integrative polygenic risk score improves the prediction accuracy of complex traits
24 and diseases. *Cell Genom* **4**, 100523 (2024).
- 25 13. Kelemen, M., Vigorito, E., Fachal, L., Anderson, C. A. & Wallace, C. shaPRS: Leveraging shared
26 genetic effects across traits or ancestries improves accuracy of polygenic scores. *Am. J. Hum. Genet.*
27 **111**, 1006–1017 (2024).
- 28 14. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of
29 complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
- 30 15. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**,
31 203–209 (2018).
- 32 16. Moore, C. *et al.* The INTERVAL trial to determine whether intervals between blood donations can be
33 safely and acceptably decreased to optimise blood supply: study protocol for a randomised controlled
34 trial. *Trials* **15**, 363 (2014).
- 35 17. Di Angelantonio, E. *et al.* Efficiency and safety of varying the frequency of whole blood donation
36 (INTERVAL): a randomised trial of 45 000 donors. *Lancet* **390**, 2360–2371 (2017).
- 37 18. All of Us Research Program Investigators *et al.* The “All of Us” Research Program. *N. Engl. J. Med.*
38 **381**, 668–676 (2019).
- 39 19. All of Us Research Program Genomics Investigators. Genomic data in the All of Us Research
40 Program. *Nature* **627**, 340–346 (2024).
- 41 20. Ramirez, A. H. *et al.* The All of Us Research Program: Data quality, utility, and diversity. *Patterns (N*
42 *Y)* **3**, 100570 (2022).
- 43 21. Tan, K. H. X. *et al.* Cohort Profile: The Singapore Multi-Ethnic Cohort (MEC) study. *Int. J.*
44 *Epidemiol.* **47**, 699–699j (2018).
- 45 22. Hippisley-Cox, J. & Coupland, C. Development and validation of QDiabetes-2018 risk prediction
46 algorithm to estimate future risk of type 2 diabetes: cohort study. *BMJ* **359**, j5019 (2017).

- 47 23. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**,
48 68–74 (2015).
- 49 24. Committee on the Use of Race, Ethnicity, and Ancestry as Population Descriptors in Genomics
50 Research *et al. Using Population Descriptors in Genetics and Genomics Research*. (National
51 Academies Press, Washington, D.C., 2023).
- 52 25. Wong, L.-P. *et al.* Deep whole-genome sequencing of 100 southeast Asian Malays. *Am. J. Hum.*
53 *Genet.* **92**, 52–66 (2013).
- 54 26. Privé, F., Arbel, J. & Vilhjálmsson, B. J. LDpred2: better, faster, stronger. *Bioinformatics* **36**, 5424–
55 5431 (2021).
- 56 27. Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via
57 Coordinate Descent. *J. Stat. Softw.* **33**, 1–22 (2010).
- 58 28. Eastwood, S. V. *et al.* Algorithms for the Capture and Adjudication of Prevalent and Incident Diabetes
59 in UK Biobank. *PLoS One* **11**, e0162388 (2016).
- 60 29. Mars, N. *et al.* Systematic comparison of family history and polygenic risk across 24 common
61 diseases. *Am. J. Hum. Genet.* **109**, 2152–2162 (2022).
- 62 30. Huerta-Chagoya, A. *et al.* The power of TOPMed imputation for the discovery of Latino-enriched rare
63 variants associated with type 2 diabetes. *Diabetologia* **66**, 1273–1288 (2023).
- 64 31. Shim, I. *et al.* Clinical utility of polygenic scores for cardiometabolic disease in Arabs. *Nat. Commun.*
65 **14**, 6535 (2023).
- 66 32. Ding, Y. *et al.* Polygenic scoring accuracy varies across the genetic ancestry continuum. *Nature* **618**,
67 774–781 (2023).
- 68 33. Martin, A. R. *et al.* Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat.*
69 *Genet.* **51**, 584–591 (2019).
- 70 34. *Type 2 Diabetes: Prevention in People at High Risk*. (National Institute for Health and Care
71 Excellence, 2017).

- 72 35. Hylton, K., Thompson, K., Kearney, M. & Lagord, C. *NHS Health Check: Best Practice Guidance*.
73 (Public Health England, 2019).
- 74 36. Fatumo, S. *et al.* A roadmap to increase diversity in genomic studies. *Nat. Med.* **28**, 243–250 (2022).
- 75 37. Gurdasani, D., Barroso, I., Zeggini, E. & Sandhu, M. S. Genomics of disease risk in globally diverse
76 populations. *Nat. Rev. Genet.* **20**, 520–535 (2019).
- 77 38. Monti, R. *et al.* Evaluation of polygenic scoring methods in five biobanks shows larger variation
78 between biobanks than methods and finds benefits of ensemble learning. *Am. J. Hum. Genet.* **0**,
79 (2024).
- 80 39. Hodgson, S. *et al.* Integrating polygenic risk scores in the prediction of type 2 diabetes risk and
81 subtypes in British Pakistanis and Bangladeshis: A population-based cohort study. *PLoS Med.* **19**,
82 e1003981 (2022).
- 83 40. Genomics England. Newborn Genomes Programme.
84 <https://www.genomicsengland.co.uk/initiatives/newborns> (2023).
- 85 41. Kachuri, L. *et al.* Principles and methods for transferring polygenic risk scores across global
86 populations. *Nat. Rev. Genet.* **25**, 8–25 (2024).
- 87 42. Lennon, N. J. *et al.* Selection, optimization and validation of ten chronic disease polygenic risk scores
88 for clinical implementation in diverse US populations. *Nat. Med.* **30**, 480–487 (2024).
- 89 43. Loh, P.-R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat.*
90 *Genet.* **48**, 1443–1448 (2016).
- 91 44. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies.
92 *Bioinformatics* **26**, 2867–2873 (2010).
- 93 45. Sun, B. B. *et al.* Genomic atlas of the human plasma proteome. *Nature* **558**, 73–79 (2018).
- 94 46. Kuan, V. *et al.* A chronological map of 308 physical and mental health conditions from 4 million
95 individuals in the English National Health Service. *Lancet Digit Health* **1**, e63–e77 (2019).
- 96 47. Wong, E. *et al.* The Singapore National Precision Medicine strategy. *Nat. Genet.* **55**, 178–186 (2023).

- 97 48. Kwong, A. M. *et al.* Robust, flexible, and scalable tests for Hardy-Weinberg equilibrium across
98 diverse ancestries. *Genetics* **218**, (2021).
- 99 49. Zhang, F. *et al.* Ancestry-agnostic estimation of DNA sample contamination from sequence reads.
00 *Genome Res.* **30**, 185–194 (2020).
- 01 50. Liang, Y. *et al.* Circulating proteomic profiles are associated with the onset of type 2 diabetes in a
02 multi-ethnic Asian population — a longitudinal study. *medRxiv* 2024.05.28.24308009 (2024)
03 doi:10.1101/2024.05.28.24308009.
- 04 51. American Diabetes Association Professional Practice Committee. 2. Classification and diagnosis of
05 diabetes: Standards of Medical Care in diabetes-2022. *Diabetes Care* **45**, S17–S38 (2022).
- 06 52. Wand, H. *et al.* Improving reporting standards for polygenic scores in risk prediction studies. *Nature*
07 **591**, 211–219 (2021).
- 08 53. International HapMap 3 Consortium *et al.* Integrating common and rare genetic variation in diverse
09 human populations. *Nature* **467**, 52–58 (2010).
- 10 54. Nassar, L. R. *et al.* The UCSC Genome Browser database: 2023 update. *Nucleic Acids Res.* **51**,
11 D1188–D1195 (2023).
- 12 55. Robin, X. *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves.
13 *BMC Bioinformatics* **12**, 77 (2011).
- 14 56. Therneau, T. M. & Grambsch, P. M. *Modeling Survival Data: Extending the Cox Model*. (Springer
15 Science & Business Media, 2013).
- 16 57. Efron, B. Censored Data and the Bootstrap. *J. Am. Stat. Assoc.* **76**, 312–319 (1981).
- 17 58. Pencina, M. J., D’Agostino, R. B., Sr & Steyerberg, E. W. Extensions of net reclassification
18 improvement calculations to measure usefulness of new biomarkers. *Stat. Med.* **30**, 11–21 (2011).
- 19 59. Uno, H., Tian, L., Cai, T., Kohane, I. S. & Wei, L. J. A unified inference procedure for a class of
20 measures to assess improvement in risk prediction systems with survival data. *Stat. Med.* **32**, 2430–
21 2442 (2013).

22 Figures

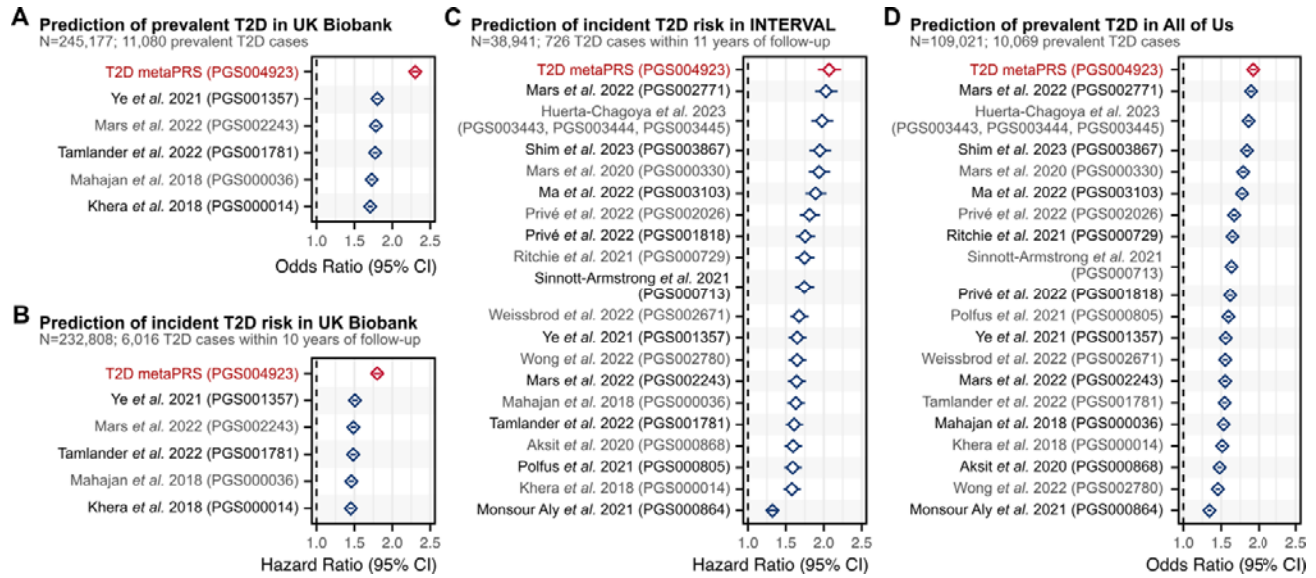


23

24 Figure 1: Study Design

25 Acronyms are as follows. T2D: type 2 diabetes. PRS: polygenic risk scores. SNPs: single nucleotide
 26 polymorphisms. LD: linkage disequilibrium. GWAS: genome-wide association study. 1KG-AFR-like,
 27 1KG-AMR-like, 1KG-EAS-like, 1KG-EUR-like, 1KG-SAS-like: genetic ancestry labels, defined based on
 28 clustering of participants by genetic principal components and the similarity of those clusters to 1000
 29 Genomes reference panel superpopulations following the 2023 National Academies guidelines on using
 30 population descriptors in genetics and genomics research. ASN-like: genetic ancestry label chosen for
 31 ethnic Malays in the Singapore Multi-Ethnic cohort to represent their ancestral population history due to
 32 lack of representation of Austronesian populations in the 1000 Genomes reference panel.

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

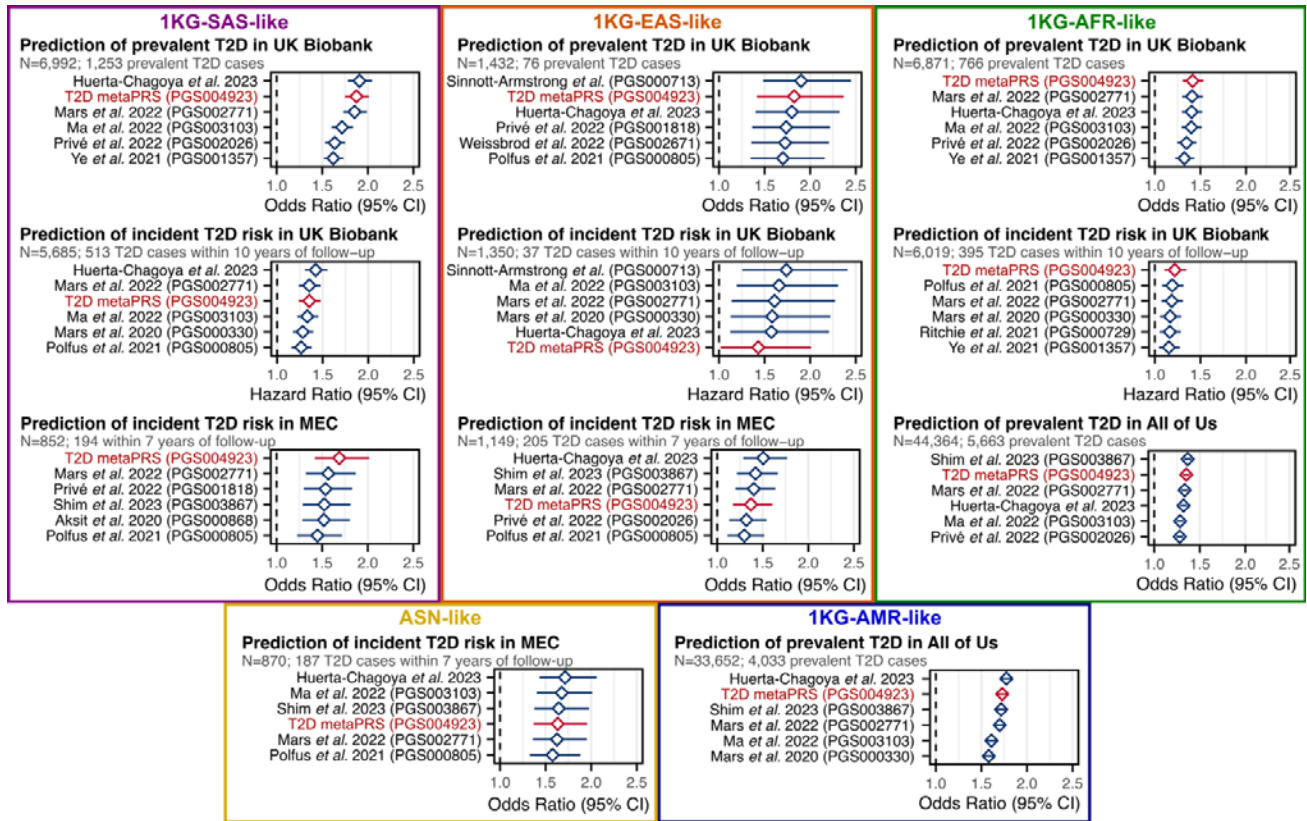


33

34 **Figure 2: Comparison of T2D PRSs in people of 1KG-EUR-like genetic ancestries across three**
35 **cohorts**

36 Comparison of PRSs for association with prevalent T2D status or time-to-onset of incident T2D in
37 participants of 1KG-EUR-like genetic ancestries from the UK Biobank, INTERVAL, and All of Us
38 research program cohorts. In UK Biobank incident and prevalent T2D were analysed separately due to
39 significant difference in phenotype severity (**Methods**). Analyses of UK Biobank excluded participants
40 used for metaPRS training, and PRSs derived from GWAS performed in UK Biobank samples. The limited
41 number of PRSs tested in UK Biobank compared to INTERVAL and All of Us reflects that the majority of
42 contemporary PRSs utilize GWAS performed in UK Biobank samples for PRS development. PRSs were
43 adjusted for 20 genetic principal components in each cohort prior to model fitting. Diamonds show the odds
44 ratios or hazard ratios, and horizontal bars show the 95% confidence intervals. Odds ratios and hazard
45 ratios are per standard deviation increase in the respective PC-adjusted PRS. Logistic and Cox proportional
46 hazards regressions were adjusted for age, sex, and cohort specific covariates (e.g., assessment centre).
47 Odds ratios and hazard ratios are detailed in **Table S6**. Details on comparison PRSs are provided in **Table**
48 **S5**.

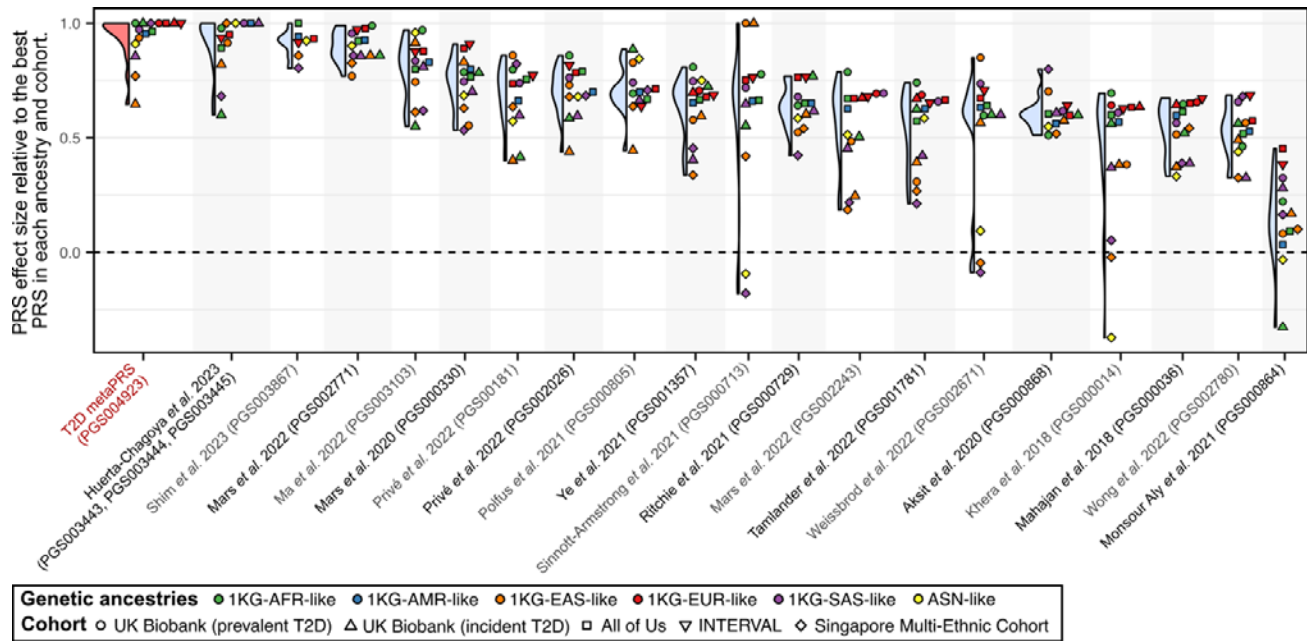
It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).



49

50 **Figure 3: Top six T2D PRSs in people of five diverse ancestry groups across three cohorts**

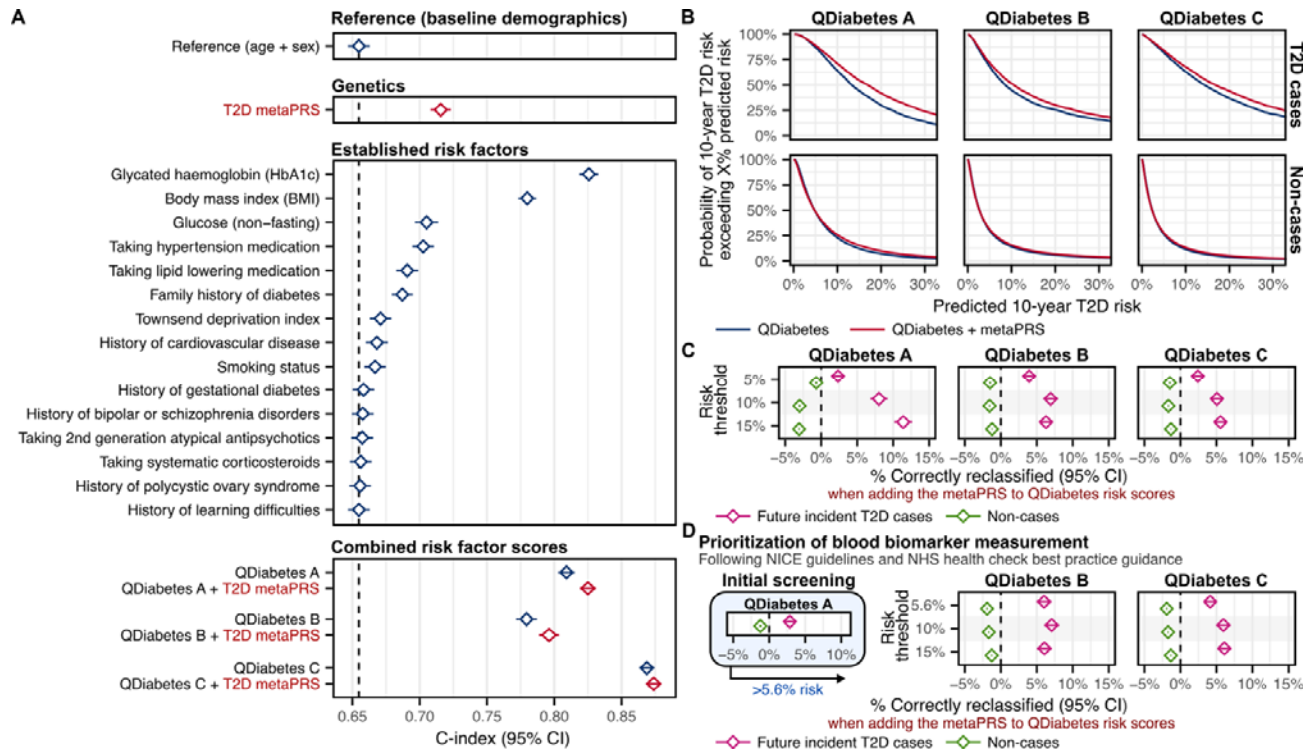
51 Comparison of PRSs for association with T2D status or time-to-onset in participants clustering into five
 52 diverse ancestry groups from UK Biobank, the All of Us research program cohort, and the Singapore Multi-
 53 Ethnic Cohort (MEC). Participants in each cohort were clustered by genetic similarity with the 1KG
 54 reference population participants (**Methods**), except for ethnic Malays in the MEC study, as their genetic
 55 ancestries are distinct from and not represented by 1KG reference populations, and here labelled as
 56 Austronesian (ASN)-like to reflect their ancestral population histories. The top six PRS for each cohort and
 57 genetic ancestry group are shown; in UK Biobank incident and prevalent T2D were analysed separately due
 58 to significant difference in phenotype severity (**Methods**). Odds ratios and hazard ratios for all tested PRS
 59 are detailed in **Table S7**. PRSs were adjusted for 20 genetic principal components in each cohort prior to
 60 model fitting. Diamonds show the odds ratios or hazard ratios, and horizontal bars show the 95%
 61 confidence intervals. Odds ratios and hazard ratios are per standard deviation increase in the respective PC-
 62 adjusted PRS. Logistic and Cox proportional hazards regressions were adjusted for age, sex, and cohort
 63 specific covariates (e.g., assessment centre). Logistic regression was used to assess associations with
 64 incident T2D in MEC as time to T2D onset (or T2D-free survival) was not available due to the
 65 heterogeneity of incident T2D ascertainment (**Methods**). Note the Shim *et al.* 2023 PRS (PGS003867)
 66 could not be tested in UK Biobank as it was derived from multi-ancestry GWAS performed in UK Biobank
 67 samples.



68

69 **Figure 4: Relative rank of PRSs by effect size within each genetic ancestry and cohort**

70 Violin and dotplot comparing PRS effect size relative to the PRS with the maximum effect size within each
71 cohort and genetic ancestry group. A score of 1.0 is given to the PRS with the maximum odds ratio or
72 hazard ratio in each cohort and genetic ancestry combination. Other PRSs were then assigned a relative
73 prediction value based on the ratio of their log odds ratio (or log hazard ratio) to that of the strongest PRS.
74 PRS are ordered left to right based on their median score.



75

76 **Figure 5: Comparison to established risk factors and risk prediction scores**

77 **A)** Comparison of C-index over age and sex alone for the metaPRS to individual type 2 diabetes risk
 78 factors and 10-year type 2 diabetes risk prediction scores (QDiabetes) in 190,293 1KG-EUR-like UK
 79 Biobank participants (4,064 incident T2D cases). The QDiabetes 2018 model A score is calculated from all
 80 listed individual risk factors, excepting glucose and HbA1c. The QDiabetes 2018 model B and model C
 81 scores additionally incorporates fasting glucose and HbA1c respectively. Note UK Biobank participants are
 82 non-fasting leading to likely underestimation of QDiabetes B. For comparison purposes the set of
 83 participants analysed here was selected as the subset in which the QDiabetes 2018 model C risk score could
 84 be computed (complete risk factor information, with height between 1.4 and 2.1 meters, weight \leq 180 kg,
 85 and HbA1c between 15 and 48 mmol/mol). Diamonds show the C-index and horizontal bars show the 95%
 86 confidence intervals. C-indices and individual risk factor hazard ratios are detailed in **Table S8**. **B)**
 87 Probability of predicted 10-year risk exceeding X% when using QDiabetes risk with or without the T2D
 88 metaPRS. Probabilities were calculated as one minus the empirical cumulative distributive function across
 89 cases and non-cases combined. Probability curves extend to the right of each plot up, to 100% predicted
 90 risk, but are truncated here for clarity. **C)** Categorical net reclassification improvement (NRI) when adding
 91 the metaPRS to QDiabetes risk scores for stratifying participants into high and low risk groups at varying
 92 risk thresholds. % correctly reclassified: net % of cases that were correctly reclassified from the low-risk
 93 group into the high-risk group when adding the metaPRS (pink) or the net % of non-cases that were
 94 correctly reclassified from the high-risk group into the low-risk group when adding the metaPRS (green).
 95 95% confidence intervals were estimated via a bootstrap sampling procedure with 1000 bootstraps.
 96 Diamonds show the net % correctly reclassified and horizontal bars show the 95% confidence intervals.
 97 Categorical NRI details and numbers allocated to each risk category are provided in **Table S9**. **D)**
 98 Categorical NRI when incorporating the metaPRS into a two-stage procedure in which QDiabetes model A
 99 is used to prioritize potential high-risk individuals for fasting glucose or HbA1c blood tests for subsequent

00 risk prediction and stratification. Categorical NRI details and numbers allocated to each risk category are
01 provided in **Table S10**.