Diabetologia (2025) 68:367–381
https://doi.org/10.1007/s00125-024-06320-3

ARTICLE

A polygenic risk score derived from common variants of monogenic diabetes genes is associated with young-onset type 2 diabetes and cardiovascular–kidney complications

Chun-Kwan O[1,2] · Baoqi Fan[1,2] · Sandra T. F. Tsoi[1] · Claudia H. T. Tam[1,2] · Raymond Wan[1] · Eric S. H. Lau[1,2] · Mai Shi[1,2] · Cadmon K. P. Lim[1,2] · Gechang Yu[1,2] · Jane P. Y. Ho[1] · Elaine Y. K. Chow[1,2,3] · Alice P. S. Kong[1,2,3] · Risa Ozaki[1,2,3] · Wing Yee So[1,2,3] · Ronald C. W. Ma[1,2,3] · Andrea O. Y. Luk[1,2,3] · Juliana C. N. Chan[1,2,3]

Received: 5 June 2024 / Accepted: 16 September 2024 / Published online: 23 November 2024

Abstract

Aims/hypothesis Monogenic diabetes is caused by rare mutations in genes usually implicated in beta cell biology. Common variants of monogenic diabetes genes (MDG) may jointly influence the risk of young-onset type 2 diabetes (YOD, diagnosed before the age of 40 years) and cardiovascular and kidney events.

Methods Using whole-exome sequencing data, we constructed a weighted polygenic risk score (wPRS) consisting of 135 common variants (minor allele frequency >0.01) of 34 MDG based on r^2>0.2 for linkage disequilibrium in a discovery case–control cohort of 453 adults with YOD (median [IQR] age 39.7 [34.9–46.9] years) and 405 without YOD (median [IQR] age 56.7 [50.3–61.0] years), followed by validation in an independent cross-sectional cohort with array-based genotyping for YOD and a prospective cohort of individuals with type 2 diabetes for cardiovascular and kidney events.

Results In the discovery cohort, the OR of the 135 common variants for YOD ranged from 1.00 to 2.61. In the validation cohort (920 YOD and 4910 non-YOD), top-10%-wPRS was associated with an OR of 1.42 (95% CI 1.03, 1.95, p=0.033) for YOD compared with bottom-10%-wPRS. In 2313 individuals with type 2 diabetes (median [IQR]: age 53.4 [45.4–61.7] years; disease duration 4.0 [1.0–9.0] years) observed for a median (IQR) of 17.5 (14.4–21.8) years, standardised wPRS was associated with increased HR for incident cardiovascular events (1.16 [95% CI 1.06, 1.27], p=0.001), kidney events (1.09 [95% CI 1.02, 1.16], p=0.013) and cardiovascular–kidney events (1.10 [95% CI 1.03, 1.16], p=0.003). Using the 'bottom-20%-wPRS plus baseline disease duration <5 years' group as referent, the 'top-20%-wPRS plus baseline disease duration 5 to <10 years' group had unadjusted and adjusted HR of 1.60 (95% CI 1.17, 2.19, p=0.003) and 1.62 (95% CI 1.16, 2.26, p=0.005), respectively, for cardiovascular–kidney events compared with 1.38 (95% CI 0.97, 1.98, p=0.075) and 1.06 (95% CI 0.72, 1.57, p=0.752) in the 'bottom-20%-wPRS plus baseline disease duration ≥10 years' group.

Conclusions/interpretation Common variants of MDG increased risk for YOD and cardiovascular–kidney events.

Keywords Complications · Genetics · MODY · Polygenic risk scores · Whole-exome sequencing · Young-onset diabetes

✉ Juliana C. N. Chan
jchan@cuhk.edu.hk

[1] Department of Medicine and Therapeutics, Faculty of Medicine, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong, China

[2] Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong, China

[3] Hong Kong Institute of Diabetes and Obesity, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong, China

Abbreviations

BHBHK	Better Health for Better Hong Kong
CHF	Congestive heart failure
CKD	Chronic kidney disease
CUHK	Chinese University of Hong Kong
DD	Diabetes duration
ESKD	End-stage kidney disease
GWAS	Genome-wide association studies
HKDR	Hong Kong Diabetes Register
HKFDS	Hong Kong Family Diabetes Study
LD	Linkage disequilibrium
LOD	Later-onset diabetes
MDG	Monogenic diabetes genes

Research in context

What is already known about this subject?

- Rare variants in monogenic diabetes genes (MDG) are implicated in beta cell and adipose biology
- Hundreds of risk loci contributing to the development of type 2 diabetes have been identified, including some common variants in MDG
- Young-onset diabetes (YOD) might have different underlying genetics from later-onset diabetes but data are limited, especially in Asian populations

What is the key question?

- Could common variants of MDG influence the risk of YOD and cardiovascular–kidney events?

What are the new findings?

- Common variants of MDG jointly increased the risk of YOD, where the top-10% weighted polygenic risk score (wPRS) constructed from common variants of MDG was associated with a 40% higher risk compared with the bottom-10%-wPRS
- In people with type 2 diabetes, those with the top-20%-wPRS had a 40% higher risk of incident cardiovascular–kidney events compared with those in the bottom-20%-wPRS group
- People with the top-20%-wPRS and baseline disease duration 5 to <10 years had a higher risk of incident cardiovascular–kidney events compared with those with the bottom-20%-wPRS and baseline disease duration ≥10 years

How might this impact on clinical practice in the foreseeable future?

- This study highlights the potential utility of data for common variants of MDG in increasing the precision of risk stratification for early treatment intensification to delay onset of diabetes and its complications

PVD	Peripheral vascular disease
PWH	Prince of Wales Hospital
RV	Rare variants
swPRS	Standardised wPRS
WES	Whole-exome sequencing
wPRS	Weighted polygenic risk score
YOD	Young-onset diabetes

Introduction

Monogenic diabetes, including MODY, refers to a group of diabetes of Mendelian inheritance due to rare mutations of specific genes usually implicated in developmental and beta cell biology, glucose sensing and insulin translation or processing [1]. Some mutations cause severe insulin resistance due to dysregulation of insulin signalling or fat metabolism, while some are associated with syndromic features [2]. The physiological roles of monogenic diabetes genes (MDG) are supported by experimental animal models and/or co-segregation among family members [1–3]. Given their functional significance, we argue that their common variants, albeit with smaller effect size, might contribute to young onset of type 2 diabetes. Apart from their effects on energy metabolism, the high glycaemic burden resulting from potentially poor glycaemic trajectory and long duration of exposure to abnormal milieu due to younger age at diabetes onset will increase the risk of cardiovascular–kidney events [4].

Large-scale genome-wide association studies (GWAS) across ancestries have reported associations of common variants of MDG with type 2 diabetes or related traits [5, 6]. Given potential differences in the aetiologies of young-onset diabetes (YOD, diagnosed before the age of 40 years) and later-onset diabetes (LOD), a few studies revealed specific genetic associations stratified by age of diagnosis [7, 8]. In the first GWAS targeting youth-onset diabetes in a multiethnic cohort (non-Hispanic White, African American and Hispanic) with an age of diagnosis <20 years and a mean age of 15 years, researchers identified seven genome-wide significant loci, including one novel signal in *PHF2* (encoding for PHD finger protein 2) not reported to be associated with type 2 diabetes in adults [9]. However, similar genetic studies for YOD diagnosed before the age of 40 years are lacking, especially in Asians.

We hypothesised that common variants of MDG jointly increase the risk of YOD and incident cardiovascular–kidney complications, and tested this hypothesis by constructing a weighted polygenic risk score (wPRS) with common variants of 34 MDG based on a discovery cohort (YOD vs non-YOD) with whole-exome sequencing (WES), followed by validation in independent cohorts with genotyping.

Methods

Study design and participants Participants came from three established cohorts (Fig. 1 and electronic supplementary material [ESM] Fig. 1): the Hong Kong Family Diabetes Study (HKFDS); Better Health for Better Hong Kong (BHBHK); and the Hong Kong Diabetes Register (HKDR).

The HKFDS cohort was established in 1998–2003 by the Chinese University of Hong Kong (CUHK) Diabetes Care and Research Team. There were 192 index cases with diabetes (149 with YOD), and their family members, giving a total of 1076 participants, recruited for studying genetic and environmental causes of diabetes in the Chinese population [10, 11]. The index cases were identified in the diabetes complication assessment programme at the Prince of Wales Hospital (PWH), followed by invitation of their relatives to participate in the study.

The BHBHK cohort was established in 2001–2003 as part of a community-based health promotion campaign to screen for cardiovascular risk factors including obesity and diabetes in the workforce [12]. The HKFDS cohort and a random BHBHK sub-cohort (n=863) underwent structured assessment, including personal and family history, anthropometric measurements and collection of blood and urine samples for metabolic profiling [11, 13, 14]. They underwent a 2 h 75 g OGTT with measurements of plasma insulin, C-peptide and glucose, accompanied by a DNA/serum biobank. In 2012–2014, the diabetes status of both cohorts was ascertained using medical records, OGTT and HbA_{1c} [11].

The HKDR was established in 1995 by the CUHK-PWH Team as a research-driven quality improvement programme in a hospital-based setting. Patients with diabetes could be referred from all PWH medical clinics to the PWH Diabetes Centre where collection of clinical information, screening for diabetes-related complications, and data management and reporting were conducted, guided by a pre-defined protocol [15]. The participants were prospectively observed with ascertainment of clinical outcomes retrieved from the territory-wide electronic medical record system of the Hong Kong Hospital Authority. Details of the rationale, setting, team structure, procedures, database management and data-driven care were reported [16]. These studies were approved by the CUHK Clinical Research Ethics Committee.

Discovery cohort, validation cohort and HKDR prospective cohort The discovery cohort consisted of 453 individuals with YOD from the HKDR and 405 individuals without YOD from the BHBHK with WES as part of the Global Type 2 Diabetes Consortium [17]. ESM Tables 1–2 show the variants within the 34 MDG. ESM Table 3 summarises the methodology for analysing WES and array-based genotyping data used to construct a wPRS.

We validated the performance of the derived wPRS in predicting YOD in a separate cohort of 920 individuals with YOD from the HKDR and 4910 individuals without YOD from the HKDR and the BHBHK (excluding those involved in discovery cohort) with available array-based genotyping data (validation cohort). The non-YOD group included 4670 individuals with LOD from the HKDR and 240 without diabetes from the BHBHK.

We tested the association between wPRS and incident cardiovascular–kidney events in the prospective HKDR cohort of 2313 Chinese individuals with type 2 diabetes stratified by wPRS deciles and disease duration (excluding those with CVD, kidney disease and albuminuria at baseline, and those involved in discovery cohort). In a secondary analysis, we tested the associations of the wPRS with beta cell function indices and incident diabetes in 363 individuals without diabetes at baseline from the BHBHK and the HKFDS (excluding those involved in the discovery cohort, randomly picking one individual from each family for HKFDS).

Definitions and outcomes In all cohorts, sex referred to the biological sex of the individuals, and the information was defined by the sex entity recorded in official government documents such as the Hong Kong Identity Card. In the discovery cohort, YOD was defined as diabetes diagnosed before the age of 40 years and non-YOD was defined as no diabetes at the age of ≥40 years. In the validation cohort, we expanded the definition of the non-YOD group to include LOD diagnosed at age ≥40 years for a larger sample size.

In the prospective HKDR cohort, incident cardiovascular–kidney complications were defined by hospital discharge principal diagnoses and procedures coded by ICD-9 (http://www.icd9data.com/2007/Volume1/default.htm) and laboratory variables: (1) CHD; (2) stroke; (3) peripheral vascular disease (PVD); (4) congestive heart failure (CHF); (5) CVD; (6) chronic kidney disease (CKD); (7) end-stage kidney disease (ESKD); and (8) composite cardiovascular–kidney disease (ESM Table 4).

In the secondary analysis, we examined the associations of the wPRS with incident diabetes and beta cell function indices in individuals without diabetes at baseline. We calculated HOMA2-%B and HOMA2-IR using the HOMA2 calculator v2.2.3 (https://www.dtu.ox.ac.uk/homacalculator/). Insulinogenic index and disposition index as indices of beta cell function and insulin resistance were calculated as follows:

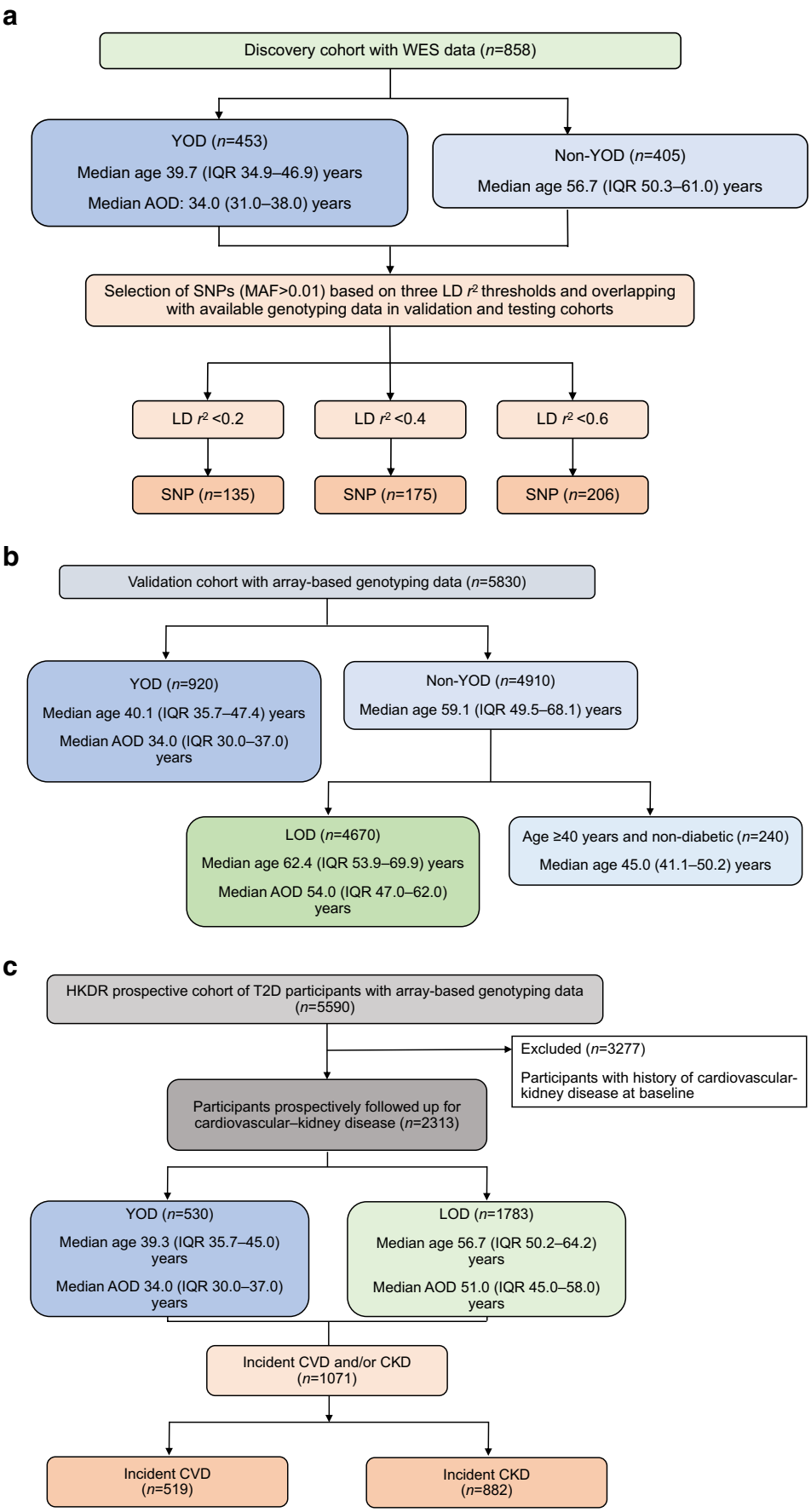

Fig. 1 Discovery cohort, validation cohort and prospective HKDR cohort. (**a**) Discovery cohort and selection of SNPs within MDG for construction of polygenic risk scores for YOD. All non-YOD participants were aged ≥40 years and without diabetes. (**b**) Validation cohort for assessing performance of wPRS for YOD. (**c**) Prospective HKDR cohort for assessing association between wPRS and incident cardiovascular–kidney complications in type 2 diabetes. AOD, age of diagnosis; MAF, minor allele frequency; T2D, type 2 diabetes

$$\text{Insulinogenic index} = \frac{\text{Ins30} - \text{Ins0}}{\text{Gluc30} - \text{Gluc0}}$$

$$\text{Disposition index} = \frac{\text{Insulinogenic index}}{6 \times \text{HOMA} - \text{IR}}$$

where Ins0, Ins30, Gluc0 and Gluc30 are plasma insulin at 0 min, insulin at 30 min, glucose at 0 min and glucose at 30 min during OGTT, respectively (units for plasma insulin and glucose are pmol/l and mmol/l, respectively).

Construction of wPRS The wPRS was constructed as follows:

$$\text{wPRS} = \beta_{1\times1} + \beta_{2\times2} + \ldots + \beta_k x_k + \ldots + \beta_n x_n$$

where β_k is the per-allele effect size for YOD associated with a single-nucleotide variant k of the 34 MODY genes, x_k is the number of effect alleles of the single-nucleotide variant k, and n is the total number of single-nucleotide variants involved in the construction of the polygenic risk score.

We employed a 'pruning-and-thresholding' approach to construct and choose a wPRS with optimal performance. Due to use of different arrays, we selected SNPs available in all cohorts and included 135, 175 and 206 SNPs located within ±1000 base-pairs of gene regions of the 34 MDG by varying r^2 thresholds of linkage disequilibrium (LD) at 0.2, 0.4 and 0.6, respectively.

Statistical analysis All data were expressed as mean ± SD or median (IQR). Between-group comparisons were made by parametric and non-parametric tests as appropriate. The wPRS was transformed into a standardised wPRS (swPRS, see equation below) or divided into categories in the association analyses.

$$\text{swPRS} = \frac{\text{wPRS of a specific individual} - \text{mean wPRS}}{\text{SD of wPRS}}$$

Binary logistic regression was used to examine the association of the wPRS with YOD expressed as OR with 95% CI in the validation cohort. Kaplan–Meier estimation accompanied by curves of one-minus-survival functions was used to describe the cumulative incidence of events with logranked test for examining differences among groups in the prospective HKDR cohort. Cox proportional hazard regression was used to examine the association of the wPRS with incident cardiovascular–kidney events expressed as HR with 95% CI accompanied by curves of one-minus-survival functions with covariates controlled at mean for continuous variables and at reference category for nominal variables. In the secondary analysis, we randomly picked one individual from each family and used multivariate linear and binary logistic regression to examine the associations of the wPRS with beta cell function indexes (HOMA-2%B, insulinogenic index, disposition index), insulin resistance (HOMA2-IR) and incident diabetes in the BHBHK-HKFDS. Missing data were handled by pairwise deletion, and the total number of individuals involved in each model of regression analysis were stated.

Results

Construction and validation of the wPRS for YOD Figure 1 and Table 1 summarise the profiles of the discovery and validation cohorts, and the prospective HKDR cohort at baseline. We analysed the WES data of 453 individuals with YOD from the HKDR (median age 39.7 [IQR 34.9–46.9] years; median age of diagnosis 34.0 [IQR 31.0–38.0] years) and 405 individuals without YOD (non-YOD) from the BHBHK (median age 56.7 [IQR 50.3–61.0] years) and estimated the effect size of each SNP of the 34 MDG for YOD (discovery cohort). Using overlapping genotyping data from the BHBHK, HKFDS and HKDR, we selected 135, 175 and 206 SNPs to construct three wPRS using LD statistics r^2 thresholds of 0.2, 0.4 and 0.6, respectively.

In the validation cohort, all swPRS were positively associated with increased odds for YOD. The wPRS constructed using SNPs with r^2<0.2 performed the best, with the swPRS having an unadjusted OR of 1.073 (95% CI 1.00, 1.15, p=0.051) and a sex- and BMI-adjusted OR of 1.07 (95% CI 0.99, 1.15, p=0.074) for YOD (Table 2 and ESM Table 5a). The OR of 1.08 (95% CI 0.94, 1.25, p=0.280) remained similar in a sensitivity analysis with strict inclusion of only those aged ≥40 years with no diabetes as the non-YOD group (ESM Table 5b). Using these SNPs (r^2<0.2), the top-10%-wPRS group had 42% higher risk of YOD than the bottom-10%-wPRS group (OR 1.42 [95% CI 1.03, 1.95], p=0.033) while the OR and significance (1.11 [95% CI 0.89, 1.39], p=0.346) were attenuated when comparing top-20% with bottom-20%-wPRS. The wPRS based on LD r^2 threshold of 0.2 was therefore used in the subsequent analysis, and the OR of the 135 SNPs for YOD ranged from 1.00 to 2.61 in the discovery cohort (ESM Table 2b).

Association of the wPRS with incident cardiovascular–kidney events in individuals with type 2 diabetes In the HKDR, 2313 individuals with no history of cardiovascular–kidney events and albuminuria at baseline (enrolled in 1994–2007) were identified. After a median follow-up of 17.5 (IQR 14.4–21.8) years, there was an accrual of 519 cardiovascular and 882 kidney events. Per-SD increase in wPRS was associated with an HR of 1.10 (95% CI 1.03, 1.16, p=0.003) for cardiovascular–kidney events (Table 3). The top-20%-wPRS group had 41% higher risk than the bottom-20%-wPRS group (HR

Table 1 Baseline characteristics of the discovery, validation and prospective HKDR cohorts

Characteristic	Discovery cohort for wPRS construction[a]				Validation cohort for YOD[b]						Prospective HKDR cohort for testing relationship of wPRS with incident cardiovascular–kidney events[b]	
	YOD (*n*=453)		Non-YOD (*n*=405)		YOD (*n*=920)		Non-YOD (*n*=4910)				YOD and LOD (*n*=2313)	
	n	Observations	*n*	Observations	*n*	Observations	LOD (*n*=4670)		No diabetes (*n*=240)		*n*	Observations
							n	Observations	*n*	Observations		
YOD	453	453 (100)	405	0 (0)	920	920 (100)	4670	0 (0)	240	0 (0)	2313	530 (22.9)
LOD	453	0 (0)	405	0 (0)	920	0 (0)	4670	4670 (100)	240	0 (0)	2313	1783 (77.1)
Without diabetes	453	0 (0)	405	405 (100)	920	0 (0)	4670	0 (0)	240	240 (100)	2313	0 (0)
Age, years	453	39.7 (34.9–46.9)	405	47.6 (42.3–54.0)[c]	920	40.1 (35.7–47.4)	4670	62.4 (53.9–69.9)	240	45.0 (41.1–50.2)[d]	2313	53.4 (45.4–61.7)
Age of diagnosis, years	451	34.0 (31.0–38.0)			920	34.0 (30.0–37.0)	4670	54.0 (47.0–62.0)			2313	47.0 (40.0–55.0)
Men	453	224 (49.4)	405	213 (52.6)	920	331 (36.0)	4670	2180 (46.7)	240	101 (42.1)	2313	1050 (45.4)
Tobacco use	452				917		4660		190		2306	
Current smoker		78 (17.3)				134 (14.6)		574 (12.3)		15 (7.9)		329 (14.3)
Ex-smoker		40 (8.8)				80 (8.7)		890 (19.1)		8 (4.2)		289 (12.5)
Non-smoker		334 (73.9)				703 (76.7)		3196 (68.6)		167 (87.9)		1688 (73.2)
Alcohol use	452				914		4648		189		2295	
Current drinker		44 (9.7)				88 (9.6)		404 (8.7)		26 (13.8)		262 (11.4)
Ex-drinker		38 (8.4)				65 (7.1)		631 (13.6)		1 (0.5)		214 (9.3)
Non-drinker		370 (81.9)				761 (83.3)		3613 (77.7)		162 (85.7)		1819 (79.3)
BMI, kg/m^2	453	25.2 (22.5–28.6)	405	24.0 (22.1–26.6)	919	25.1 (22.4–28.1)	4640	24.8 (22.5–27.2)	240	23.0 (21.2–25.6)	2308	24.5 (22.2–27.0)
Systolic BP, mmHg	453	124 (114–137)	405	119 (109–130)	920	123 (113–136)	4667	136 (124–151)	240	118 (106–128)	2313	129 (117–140)
Diastolic BP, mmHg	453	76 (69–82)	405	75 (70–82)	920	74 (68–81)	4665	76 (68–83)	240	74 (68–80)	2312	75 (68–82)
DD, years	451	6.0 (1.0–12.0)			920	7.0 (1.0–13.0)	4670	6.0 (2.0–11.0)			2313	4.0 (1.0–9.0)
<5 years		205 (45.5)				367 (39.9)		2041 (43.7)				1213 (52.4)
≥ 5 years and <10 years		89 (19.7)				188 (20.4)		1158 (24.8)				565 (24.4)
≥ 10 years		157 (34.8)				365 (39.7)		1471 (31.5)				535 (23.1)

Table 1 (continued)

Characteristic	Discovery cohort for wPRS construction[a]				Validation cohort for YOD[b]						Prospective HKDR cohort for testing relationship of wPRS with incident cardiovascular–kidney events[b]	
	YOD (*n*=453)		Non-YOD (*n*=405)		YOD (*n*=920)		Non-YOD (*n*=4910)				YOD and LOD (*n*=2313)	
	n	Observations	*n*	Observations	*n*	Observations	LOD (*n*=4670)		No diabetes (*n*=240)		*n*	Observations
							n	Observations	*n*	Observations		
HbA_{1c}, mmol/mol	451	57 (45–72)			920	58 (46–74)	4670	56 (46–69)			2313	54 (44–66)
HbA_{1c}, %	451	7.4 (6.3–8.7)			920	7.5 (6.4–8.9)	4670	7.3 (6.4–8.5)			2313	7.1 (6.2–8.2)
Fasting plasma glucose, mmol/l	453	7.8 (6.5–10.4)	404	4.9 (4.6–5.2)	920	8.0 (6.4–10.5)	4663	7.8 (6.4–9.9)	239	4.9 (4.6–5.2)	2312	7.7 (6.4–9.5)
LDL-cholesterol, mmol/l	453	3.1 (2.5–3.7)	285	3.1 (2.6–3.8)	843	3.0 (2.3–3.6)	4460	3.0 (2.5–3.7)	239	3.0 (2.5–3.7)	2223	3.1 (2.5–3.6)
HDL-cholesterol, mmol/l	453	1.2 (1.0–1.4)	405	1.4 (1.2–1.7)	905	1.3 (1.1–1.5)	4622	1.3 (1.1–1.5)	240	1.5 (1.3–1.8)	2286	1.3 (1.1–1.6)
Triacylglycerol, mmol/l	453	1.3 (0.9–1.9)	405	1.2 (0.8–2.0)	913	1.2 (0.8–1.9)	4648	1.4 (1.0–2.1)	240	0.9 (0.7–1.4)	2300	1.2 (0.9–1.9)
eGFR, ml/min per 1.73m^2	452	105 (92–114)	403	93 (81–103)	920	104 (88–114)	4670	78 (59–94)	240	95 (84–107)	2313	92 (80–103)

Summary statistics are expressed as median (IQR) or *n* (%)

[a]WES data

[b]Array-based genotyping data

[c]The median age of the non-YOD group in the discovery cohort here is based on the age of the participants in the BHBHK cohort when they were first recruited at baseline (2001–2003). Some individuals included in the non-YOD group had their 'no diabetes' status re-ascertained at 12 year follow-up. The median age of the non-YOD group based on the known oldest age of ascertained 'no diabetes' status of the participants was 56.7 (IQR 50.3–61.0) years

[d]The median age of the 'no diabetes' subgroup of the non-YOD group in the validation cohort here is based on the age of the participants in the BHBHK cohort when they were first recruited at baseline (2001–2003). Some individuals included in the non-YOD group had their 'no diabetes' status re-ascertained at 12 year follow-up. The median age of the 'no diabetes' subgroup based on the known oldest age of ascertained 'no diabetes' status of the participants was 54.7 (IQR 49.0–60.0) years

Table 2 Associations of wPRS, based on LD r^2 threshold of 0.2 during selection of SNPs, with YOD in validation cohort of 920 individuals with YOD and 4910 individuals without YOD

Model	Continuous swPRS	Top 10% vs bottom 10% wPRS	Top 15% vs bottom 15% wPRS	Top 20% vs bottom 20% wPRS
Unadjusted				
N	5830	1166	1750	2332
OR (95% CI)	1.07 (1.00, 1.15)	1.43 (1.04, 1.97)	1.31 (1.02, 1.70)	1.12 (0.90, 1.40)
p value	0.051	0.026*	0.037*	0.308
Adjusted for PC1 and PC2				
N	5830	1166	1750	2332
OR (95% CI)	1.07 (0.99, 1.14)	1.42 (1.04, 1.96)	1.30 (1.01, 1.69)	1.12 (0.89, 1.40)
p value	0.077	0.029*	0.044*	0.330
Add-on adjustment for sex				
N	5830	1166	1750	2332
OR (95% CI)	1.07 (0.99, 1.14)	1.41 (1.03, 1.94)	1.31 (1.01, 1.69)	1.11 (0.89, 1.39)
p value	0.082	0.034*	0.044*	0.365
Add-on adjustment for BMI				
N	5799	1157	1738	2317
OR (95% CI)	1.07 (0.99, 1.15)	1.42 (1.03, 1.95)	1.31 (1.01, 1.70)	1.11 (0.89, 1.39)
p value	0.074	0.033*	0.041*	0.346

*$p<0.05$

PC1, first principal component; PC2, second principal component

1.41 [95% CI 1.15, 1.72], $p<0.001$) after adjusting for baseline demographics, metabolic control (BMI, HbA_{1c}, systolic BP, triacylglycerol, LDL-cholesterol and HDL-cholesterol), eGFR, use of glucose-, BP- and lipid-lowering drugs, and use of tobacco and alcohol.

Analysis of individual components of cardiovascular–kidney events (Table 3) revealed that the per-SD increase in wPRS was associated with 16% higher risk of CVD (HR 1.16 [95% CI 1.06, 1.27], $p=0.001$). The top-20%-wPRS group had an HR of 1.87 (95% CI 1.38, 2.52, $p<0.001$) for a cardiovascular event compared with the bottom-20%-wPRS group. For each component of the cardiovascular events, the swPRS was associated with incident CHD (HR 1.21 [95% CI 1.07, 1.36], $p=0.003$) but not with stroke (HR 1.00 [95% CI 0.86, 1.16], $p=0.99$), PVD (HR 1.06 [95% CI 0.80, 1.39], $p=0.68$) or CHF (HR 1.08 [95% CI 0.89, 1.31], $p=0.44$) (ESM Table 6). For kidney outcomes, the per-SD increase in wPRS was associated with an HR of 1.09 (95% CI 1.02, 1.16, $p=0.013$) for CKD, with the top-20%-wPRS group having an HR of 1.34 (95% CI 1.07, 1.66, $p=0.010$) compared with the bottom-20%-wPRS group (Table 3). The swPRS was not associated with incident ESKD (HR 0.95 [95% CI 0.76, 1.19], $p=0.66$) (ESM Table 6). There was no significant interaction between disease duration and wPRS (both continuous swPRS and top-20%-wPRS vs bottom-20%-wPRS) for CVD, CKD and combined events.

Association of risk category, stratified by disease duration and wPRS, with incident cardiovascular–kidney events in individuals with type 2 diabetes We explored the importance of the wPRS relative to diabetes duration (DD) as a risk factor of diabetes-related complications [4]. Per-year increase in baseline DD was independently associated with an HR of 1.014 (95% CI 1.002, 1.025, $p=0.020$) for cardiovascular–kidney complications in the fully adjusted model (Model 4, Table 3). We stratified the HKDR cohort into six groups by wPRS (top 20% and bottom 20%) and baseline DD (<5 years, 5 to <10 years, and ≥10 years). We compared the risk association of these six groups with incident complications and explored whether those with top-20%-wPRS plus short baseline DD of less than 5–10 years would have comparable or higher risk of complications than the bottom-20%-wPRS plus long baseline DD ≥10 years group.

By Kaplan–Meier estimation (ESM Fig. 2, $p<0.01$ in all logrank tests), the 'top-20%-wPRS plus baseline DD<5 years' group with a median age of 49.0 (IQR 42.7–59.2) years had a cumulative incidence of 27% for CVD after 20 years, compared with 22% in the 'bottom-20%-wPRS group plus baseline DD≥10 years' group with a median age of 59.3 (IQR 51.2–63.8) years. Similarly, 55% and 66% of the 'top-20%-wPRS plus baseline DD 5 to <10 years' group with a median age of 54.5 (IQR 47.3–62.4) years developed CKD and cardiovascular–kidney events after 20 years, respectively, compared with 46% and 53% in the 'bottom 20%-wPRS plus baseline DD≥10 years' group. However, the 'top-20%-wPRS plus baseline DD<5 years' group had a lower cumulative incidence of CKD (38%) and cardiovascular–kidney events (47%) than the 'bottom 20%-wPRS plus baseline DD≥10 years' group after 20 years.

Table 3 Associations of wPRS, based on LD r^2 threshold of 0.2 during selection of SNPs, with incident cardiovascular–kidney complications in the HKDR cohort of 2313 individuals with type 2 diabetes

Model	Continuous swPRS	Top 10% vs bottom 10% wPRS	Top 20% vs bottom 20% wPRS	Top 30% vs bottom 30% wPRS
Association of wPRS with incident CVD				
Model 1[a]				
N	2308	459	920	1382
HR (95% CI)	1.15 (1.05, 1.25)	1.72 (1.17, 2.52)	1.72 (1.29, 2.29)	1.21 (0.96, 1.51)
	0.002**	0.006**	<0.001***	0.103
Model 2[b]				
N	2218	435	881	1327
HR (95% CI)	1.15 (1.06, 1.26)	1.72 (1.16, 2.55)	1.77 (1.32, 2.38)	1.23 (0.98, 1.55)
p value	0.002**	0.007**	<0.001***	0.077
Model 3[c]				
N	2218	435	881	1327
HR (95% CI)	1.16 (1.06, 1.27)	1.71 (1.15, 2.54)	1.85 (1.37, 2.49)	1.27 (1.01, 1.60)
p value	0.001**	0.008**	<0.001***	0.045*
Model 4[d]				
N	2200	430	873	1318
HR (95% CI)	1.16 (1.06, 1.27)	1.67 (1.12, 2.48)	1.87 (1.38, 2.52)	1.26 (1.00, 1.59)
p value	0.001**	0.012*	<0.001***	0.050*
Association of wPRS with incident CKD				
Model 1[a]				
N	2308	459	920	1382
HR (95% CI)	1.08 (1.01, 1.15)	1.41 (1.04, 1.90)	1.26 (1.02, 1.55)	1.11 (0.94, 1.32)
p value	0.023*	0.026*	0.031*	0.223
Model 2[b]				
N	2218	435	881	1327
HR (95% CI)	1.09 (1.01, 1.16)	1.41 (1.02, 1.93)	1.31 (1.05, 1.63)	1.17 (0.98, 1.40)
p value	0.017*	0.035*	0.015*	0.077
Model 3[c]				
N	2218	435	881	1327
HR (95% CI)	1.08 (1.01, 1.16)	1.40 (1.02, 1.93)	1.32 (1.06, 1.64))	1.17 (0.98, 1.39)
p value	0.018*	0.038*	0.013*	0.083
Model 4[d]				
N	2200	430	873	1318
HR (95% CI)	1.09 (1.02, 1.16)	1.42 (1.03, 1.95)	1.34 (1.07, 1.66)	1.17 (0.98, 1.40)
p value	0.013*	0.034*	0.010*	0.074
Association of wPRS with incident cardiovascular–kidney disease				
Model 1[a]				
N	2308	459	920	1382
HR (95% CI)	1.09 (1.02, 1.15)	1.48 (1.13, 1.94)	1.30 (1.07, 1.57)	1.11 (0.95, 1.30)
p value	0.007**	0.005**	0.008**	0.182
Model 2[b]				
N	2218	435	881	1327
HR (95% CI)	1.09 (1.03, 1.16)	1.48 (1.11, 1.96)	1.35 (1.11, 1.65)	1.16 (0.99, 1.37)
p value	0.005**	0.007**	0.003**	0.062
Model 3[c]				
N	2218	435	881	1327
HR (95% CI)	1.09 (1.03, 1.16)	1.47 (1.10, 1.96)	1.39 (1.14, 1.69)	1.17 (1.00, 1.38)
p value	0.004**	0.009**	0.001**	0.051

Table 3 (continued)

Model	Continuous swPRS	Top 10% vs bottom 10% wPRS	Top 20% vs bottom 20% wPRS	Top 30% vs bottom 30% wPRS
Model 4[d]				
N	2200	430	873	1318
HR (95% CI)	1.10 (1.03, 1.16)	1.45 (1.09, 1.94)	1.41 (1.15, 1.72)	1.19 (1.01, 1.39)
p value	0.003**	0.011*	<0.001***	0.038*

[a]Model 1: Adjusted for first principal component (PC1), second principal component (PC2), age, sex, BMI and DD

[b]Model 2: Model 1 + adjusted for metabolic control (HbA_{1c}, systolic BP, triacylglycerol, HDL-cholesterol, LDL-cholesterol and eGFR)

[c]Model 3: Model 2 + adjusted for medication use (oral glucose-lowering drugs, insulin, antihypertensives, lipid-regulating drugs)

[d]Model 4: Model 3 + adjusted for tobacco and alcohol use; in this model, the HR of disease duration of diabetes for incident cardiovascular–kidney disease was 1.014 (95% CI 1.002, 1.025, p=0.020)

*p<0.05; **p<0.01; ***p<0.001

Tables 4 and 5 show the unadjusted and adjusted HRs, respectively, for cardiovascular–kidney disease and its components in the six groups. ESM Fig. 3 shows the associated one-minus-survival function curves. Using Cox regression with the 'bottom-20%-wPRS plus baseline DD<5 years' group as the referent, the 'top-20%-wPRS plus baseline DD<5 years' group had an unadjusted HR of 1.87 (95% CI 1.23, 2.85, p=0.003) for CVD vs 1.43 (95% CI 0.81, 2.54, p=0.217) in the 'bottom-20%-wPRS plus baseline DD≥10 years' group. The 'top-20%-wPRS plus baseline DD 5 to <10 years' group had an unadjusted HR of 1.62 (95% CI 1.15, 2.28, p=0.006) and 1.60 (95% CI 1.17, 2.19, p=0.003) for CKD and cardiovascular–kidney disease, respectively, vs 1.29 (95% CI 0.86, 1.92, p=0.216) and 1.38 (95% CI 0.97, 1.98, p=0.075) in the 'bottom-20%-wPRS plus baseline DD≥10 years' group. The 'top-20%-wPRS plus baseline DD<5 years' group had unadjusted HR of 1.04 (95% CI

Table 4 Association of risk categories based on DD at baseline and wPRS with cardiovascular–kidney disease in prospective HKDR (unadjusted)

Risk category	No. of individuals	HR	95% CI	*p* value
CVD				
Bottom 20% wPRS and disease duration <5 years	230	Reference	Reference	Reference
Bottom 20% wPRS and disease duration 5 to <10 years	130	1.46	0.87, 2.43	0.149
Bottom 20% wPRS and disease duration ≥10 years	102	1.43	0.81, 2.54	0.217
Top 20% wPRS and disease duration <5 years	241	1.87	1.23, 2.85	0.003**
Top 20% wPRS and disease duration 5 to <10 years	116	1.96	1.20, 3.20	0.007**
Top 20% wPRS and disease duration ≥10 years	105	2.47	1.52, 4.02	<0.001***
CKD				
Bottom 20% wPRS and disease duration <5 years	230	Reference	Reference	Reference
Bottom 20% wPRS and disease duration 5 to <10 years	130	1.38	0.97, 1.94	0.071
Bottom 20% wPRS and disease duration ≥10 years	102	1.29	0.86, 1.92	0.216
Top 20% wPRS and disease duration <5 years	241	1.04	0.76, 1.42	0.812
Top 20% wPRS and disease duration 5 to <10 years	116	1.62	1.15, 2.28	0.006**
Top 20% wPRS and disease duration ≥10 years	105	1.99	1.40, 2.83	<0.001***
Cardiovascular–kidney disease				
Bottom 20% wPRS and disease duration <5 years	230	Reference	Reference	Reference
Bottom 20% wPRS and disease duration 5 to <10 years	130	1.21	0.88, 1.68	0.241
Bottom 20% wPRS and disease duration ≥10 years	102	1.38	0.97, 1.98	0.075
Top 20% wPRS and disease duration <5 years	241	1.11	0.84, 1.46	0.471
Top 20% wPRS and disease duration 5 to <10 years	116	1.60	1.17, 2.19	0.003**
Top 20% wPRS and disease duration ≥10 years	105	1.96	1.42, 2.70	<0.001***

**p<0.01; ***p<0.001

Table 5 Association of risk categories based on DD at baseline and wPRS with CVD in prospective HKDR (adjusted for covariates)

Risk category	No. of individuals	HR	95% CI	*p* value
CVD				
Bottom 20% wPRS and disease duration <5 years	213	Reference	Reference	Reference
Bottom 20% wPRS and disease duration 5 to <10 years	121	1.31	0.75, 2.23	0.349
Bottom 20% wPRS and disease duration ≥10 years	101	1.28	0.70, 2.34	0.427
Top 20% wPRS and disease duration <5 years	228	2.08	1.34, 3.24	0.001**
Top 20% wPRS and disease duration 5 to <10 years	108	2.05	1.21, 3.47	0.007**
Top 20% wPRS and disease duration ≥10 years	102	2.38	1.40, 4.05	0.001**
CKD				
Bottom 20% wPRS and disease duration <5 years	213	Reference	Reference	Reference
Bottom 20%wPRS and disease duration 5 to <10 years	121	1.08	0.73, 1.59	0.707
Bottom 20% wPRS and disease duration ≥10 years	101	1.00	0.65, 1.54	0.986
Top 20% wPRS and disease duration <5 years	228	1.13	0.81, 1.57	0.478
Top 20% wPRS and disease duration 5 to <10 years	108	1.63	1.12, 2.36	0.010*
Top 20% wPRS and disease duration ≥10 years	102	1.60	1.09, 2.36	0.016*
Cardiovascular–kidney disease				
Bottom 20% wPRS and disease duration <5 years	213	Reference	Reference	Reference
Bottom 20% wPRS and disease duration 5 to <10 years	121	0.92	0.64, 1.32	0.636
Bottom 20% wPRS and disease duration ≥10 years	101	1.06	0.72, 1.57	0.752
Top 20% wPRS and disease duration <5 years	228	1.22	0.90, 1.64	0.196
Top 20% wPRS and disease duration 5 to <10 years	108	1.62	1.16, 2.26	0.005 **
Top 20% wPRS and disease duration ≥10 years	102	1.52	1.07, 2.16	0.020 *

HRs are adjusted for first principal component (PC1), second principal component (PC2), age, sex, BMI, metabolic control (HbA_{1c}, systolic BP, triacylglycerol, HDL-cholesterol, LDL-cholesterol and eGFR), medication use (oral glucose-lowering drugs, insulin, antihypertensives, lipid-regulating drugs), and tobacco and alcohol use

*p<0.05; **p<0.01

0.76, 1.42, p=0.812) and 1.11 (95% CI 0.84, 1.46, p=0.471) for CKD and cardiovascular–kidney events, respectively.

The results were similar after adjusting for baseline covariates. The 'top-20%-wPRS plus baseline DD<5 years' group had an adjusted HR of 2.08 (95% CI 1.34, 3.24, p=0.001) for CVD compared with 1.28 (95% CI 0.70, 2.34, p=0.427) in the 'bottom-20%-wPRS plus baseline DD≥10 years' group. The 'top-20%-wPRS plus baseline DD 5 to <10 years' group had an adjusted HR of 1.63 (95% CI 1.12, 2.36, p=0.010) and 1.62 (95% CI 1.16, 2.26, p=0.005) for CKD and combined cardiovascular–kidney disease, respectively, compared with 1.00 (95% CI 0.65, 1.54, p=0.986) and 1.06 (95% CI 0.72, 1.57, p=0.752) in the 'bottom-20%-wPRS plus baseline DD≥10 years' group. (The 'top-20%-wPRS plus baseline DD<5 years' group had an adjusted HR of 1.13 (95% CI 0.81, 1.57, p=0.478) and 1.22 (95% CI 0.90, 1.64, p=0.196) for CKD and cardiovascular–kidney events, respectively.)

We repeated the analysis by restructuring the 'bottom-20%-wPRS plus baseline DD≥10 years' as the reference group where the 'top-20%-wPRS plus baseline DD 5 to <10 years' group had an adjusted HR of 1.62 (95% CI 1.04, 2.53, p=0.033) and 1.52 (95% CI 1.02, 2.27, p=0.040) for CKD and cardiovascular–kidney disease, respectively, while the 'top-20%-wPRS plus baseline DD<5 years' group had an adjusted HR of 1.63 (95% CI 0.93, 2.87, p=0.089) for CVD, with borderline significance.

Secondary analysis: association of the wPRS with beta cell function and incident diabetes in people without diabetes in the BHBHK-HKFDS We examined the associations of the wPRS with beta cell function and 12 year risk of incident diabetes in the BHBHK-HKFDS. swPRS were negatively associated with beta cell function indices and positively associated with risk of incident diabetes with an adjusted OR of 1.37 (95% CI 0.89, 2.12, p=0.150), albeit short of significance (ESM Table 7a).

Discussion

Based on prior knowledge regarding the functions of MDG, we successfully constructed and validated a wPRS, derived from common variants of MDG, which was associated with YOD and incident cardiovascular–kidney complication in

type 2 diabetes. After nearly two decades of follow-up, the top-20%-wPRS group with less than 5–10 years of DD at baseline had higher risk of cardiovascular–kidney complication and its components than the bottom-20%-wPRS group with baseline DD ≥10 years. Our findings highlight the potential utility of information of MDG common variants in increasing the precision of risk stratification for early treatment intensification to delay onset of diabetes and its complications.

Known importance of MDG in diabetes Genetic and experimental studies demonstrated that rare variants (RV) in MDG could cause abnormal beta cell or adipose biology resulting in familial YOD or syndromic diabetes with high penetrance [1–3]. Aside from monogenic diabetes, RV of MDG increased the risk of the common form of type 2 diabetes. In the largest WES study, including 20,791 individuals with type 2 diabetes and 24,440 individuals without diabetes from five ancestries, RV located in *PDX1* (encoding for pancreatic and duodenal homeobox 1), *GCK* (encoding for glucokinase) and *HNF1A* (encoding for hepatic nuclear factor 1α) were associated with 1.5- to 3.5-increased odds of type 2 diabetes by weighted burden testing [17]. Some of the risk loci for type 2 diabetes were common variants of MDG, despite their small effect size [5, 18, 19]. Type 2 diabetes is a polygenic disease due to common variants and RVs implicated in pancreatic, adipose and muscle biology [19]. In family-based cohort studies, common variants of MDG were shown to modulate the age of diagnosis of MODY. For example, the common *HNF1A* variant I27L advanced age of diagnosis of the protein-truncating subtype of *HNF1A*-MODY [20]. On the other hand, the polygenic risk score for the common form of type 2 diabetes also jointly advanced the age of diagnosis of *HNF1A*-MODY [21]. We and others reported that common variants of type 2 diabetes, including some in MDG, predicted younger age of diagnosis of type 2 diabetes and earlier insulin requirement in both white European and Asian individuals after adjusting for clinical covariables [22, 23].

Distinct genetic profile of YOD from LOD Many research groups have reported the heterogeneous phenotypes and aggressive clinical course of YOD including in Chinese individuals [24–26], although large-scale genetic association studies specific to YOD are lacking. Most studies examined the association of known risk variants of type 2 diabetes with age of diagnosis of diabetes or YOD instead of creating a discovery cohort of YOD cases vs either healthy or LOD control individuals. In our study, we used WES data to construct a wPRS for YOD including estimation of effect size of all common variants within MDG from a case–control cohort of YOD vs non-YOD. This was followed by validation and testing of the association of the wPRS with incident cardiovascular–kidney complications in independent Chinese cohorts. We and others had reported that different ages of diagnosis might be attributed to differences in genetics. In the Botnia Family Study, heritability (h^2) of type 2 diabetes was 0.69 in people diagnosed at age 35–60 years and dropped to 0.31 when including those diagnosed at up to age 75 years [27]. In Pima Indians with high prevalence of YOD, there was bimodal distribution of 2 h plasma glucose with strong heritability of acute insulin secretion and body fat [28]. In Hong Kong Chinese individuals, family history of YOD was associated with six- to eightfold increased risk of diabetes vs <1.6-fold for family history of diabetes diagnosed after the age of 50 years compared with no family history of diabetes [11].

In a small-scale GWAS study, we first reported the genetic association of *DACH1*, a transcription factor, with familial YOD in Chinese with replication in a multi-ethnic Asian population [29]. This SNP, implicated in insulin secretion and islet development, was also associated with systolic BP, insulin resistance and CVD in the Chinese population. Using familial YOD as a discovery cohort, we first discovered its genetic association with *PAX4*, another transcription factor, with replication in other Asian cohorts [30]. Likewise, using a pathway approach, SNPs located in *CPE* (encoding for carboxypeptidase E) and *IDE* (encoding for insulin-degrading enzyme), implicated in human islet amyloid biology, were associated with YOD with replication in Asians [31]. These findings supported the utility of our prospective cohorts with extensive phenotypes and the importance of genetic factors in YOD in Chinese individuals. The latter had lower beta cell function and more rapid decline with disease duration than LOD [26].

In a GWAS of 24,986 cases of type 2 diabetes and 187,130 controls in the UK Biobank stratified by age of diagnosis, there were subgroup-specific type 2 diabetes risk loci and subgroup-specific effect size of the common variants [7]. Seventeen independent SNPs had different effect size in different age of diagnosis subgroups where SNPs mapped to *SLCO4C1*, *SLC6A1*, *RP11–58B2.1*, *PAM* and *CCND2-AS1* were more strongly associated with cases diagnosed before the age of 50 years than with cases diagnosed after the age of 70 years, although none of them were traditional MDG. The gene encoding for peptidylglycine α-amidating monooxygenase (*PAM*) was recently identified as a novel MDG [32]. In another UK Biobank analysis of type 2 diabetes stratified by BMI and age of diagnosis (BMI>30 kg/m^2, BMI<30 kg/m^2 and age of diagnosis <60 years, BMI<30 kg/m^2 and age of diagnosis >60 years), 277 lead SNPs were identified with 18 of them, including one in *NEUROG3* (encoding for neurogenin-3), showing subgroup difference [8]. *NEUROG3* mutations can cause permanent neonatal diabetes and childhood-onset diabetes with severe insulin deficiency [33]. Similarly, Swedish researchers used cluster analysis to categorise individuals with diabetes by various combinations of age, age of

diagnosis, BMI, HOMA-B, HOMA-IR and autoantibodies. These clusters had different patterns of genetic factors with prognostic significance for insulin requirement and CKD [34–36]. Taken together, the phenotypic heterogeneity and high risk of complications in individuals with YOD, with likely distinct genetic background from LOD, call for more precise classification and stratification for early and personalised prevention [4, 37].

Risk stratification by MDG wPRS Our study using MDG wPRS has filled some of this knowledge gap. Our discovery cohort included only individuals with YOD as cases, in contrast to most discovery cohorts that include type 2 diabetes cases irrespective of age of diagnosis. Based on a candidate gene approach, we used all variants and their corresponding effect estimates to construct a wPRS for YOD with independent validation. In the main analysis, the wPRS was associated with YOD and incident cardiovascular–kidney events. The top-20%-wPRS group had 87%, 34% and 41% higher risk of incident CVD, CKD and composite outcomes, respectively, compared with the lowest-20%-wPRS group.

Although DD is a key driving factor of diabetes complications due to accumulating glycaemic burden and other factors, the MDG wPRS was no less, if not more, important than DD for risk stratification. Our data showed that those with the top 20% risk by genetics, despite short baseline DD (<5 to 10 years), had a higher hazard of cardiovascular–kidney complications than those in the bottom 20% with long baseline DD≥10 years. Since MDG are implicated in beta cell biology, insulin resistance and other syndromic features, we postulated that poor glycaemic trajectory and hence rapid accumulation of glycaemic burden within a short period might advance these complications.

More detailed analysis showed that the high genetic risk group with short DD were of younger age and had more 'optimal' glycaemic control with HbA_{1c} <53.0 mmol/mol (7%) than the low genetic risk group with long DD at baseline (ESM Table 8). This might give an illusion of a low-risk profile to their attending physicians in routine care settings where biogenetic markers are not yet in clinical use and might lead to less-frequent monitoring of glycaemic control and incident complications. This low level of vigilance could mean missed opportunity for early intensification of glucose-lowering therapy, control of other risk factors and introduction of organ-protective therapy. Of note, individuals in the high genetic risk group with short DD had higher BMI, BP, dyslipidaemia and were more likely to be active smokers than the low genetic risk group with long DD. More studies are needed to confirm the use of genetic markers to identify high-risk individuals and use this personalised information to inform practice and motivate behavioural change. To this end, in a 1 year RCT of 420 Chinese individuals with type 2 diabetes, provision of information on their genetic risks for complications improved empowerment and reduced distress, albeit without effect on metabolic control [38].

Limitations and prospects In this study, we utilised prospective cohorts enriched with YOD and clinical events to ascertain the utility of the MDG wPRS. To our knowledge, there are no comparable prospective cohorts in the Chinese population for external replication, which is one of the weaknesses, and the results might not be generalisable to other ethnicities. We do not have sufficient sample size to obtain precise estimates of effect size and confirm the significance of individual SNPs. The limited sample size also disabled hypothesis-free discovery of variants across the genome to derive an inclusive wPRS for YOD, and large-scale studies are required. Similarly, we did not perform sex-specific analysis in the discovery cohort for derivation of sex-specific wPRS in view of the limited sample size. As the wPRS was not sex-specific, it was applied to analysis of the whole validation cohort for YOD and HKDR prospective cohort for cardiovascular–kidney complications. Nevertheless, sex was included as a covariate for adjustment across all main analyses. Whether sex-specific wPRS would have different performance from non-sex-specific wPRS in this regard requires further exploration. Existing genetic discovery studies for type 2 diabetes involve mostly individuals with LOD. Based on the potential difference in genetics of YOD and LOD, we speculated that the general type 2 diabetes polygenic risk scores derived from these cohorts might be less efficient in predicting YOD compared with a YOD polygenic risk score. Comparative performance in other traits, including glycaemic deterioration and complications, would require further examination.

Conclusion Common variants of MDG were associated with YOD and cardiovascular–kidney complications in type 2 diabetes in Chinese individuals. Availability of the genetic information might help improve risk stratification for primary or secondary prevention purposes.

Supplementary Information The online version contains peer-reviewed but unedited supplementary material available at https://doi.org/10.1007/s00125-024-06320-3.

Acknowledgements We thank all the involved research and clinical staff in the Hospital Authority and Department of Medicine and Therapeutics, the Chinese University of Hong Kong, for their professionalism and dedication. Special thanks are extended to all patients, their relatives and volunteers for participating in these studies. We are grateful to the support of the Hong Kong Hospital Authority. Some data in this manuscript were presented in the American Diabetes Association 84th Scientific Sessions in the form of a published abstract and a poster in 2024.

Data availability Due to local law and regulation, no data can be shared with external parties. Summary statistics may be shared upon reasonable request to the corresponding author.

Funding The study is funded by the Hong Kong Genome Institute, Hong Kong Government Health and Medical Research Fund Commissioned Research (CFS-CUHK2) and Research Grants Council Impact Research Fund.

Author's relationships and activities JCNC has received research grants (through institutions) and/or honoraria for consultancy and/or giving lectures from Applied Therapeutics, AstraZeneca, Bayer, Boehringer Ingelheim, Eli Lilly, Hua Medicine, Powder Pharmaceuticals, Merck Serono, MSD, Pfizer, Sanofi, Viatris and Zuellig Pharma. AOYL has served as an advisory committee member for AstraZeneca, Boehringer Ingelheim, Sanofi and Amgen, and has received research grants and travel grants from AstraZeneca, Boehringer Ingelheim, MSD, Novartis, Novo Nordisk, Sanofi and Amgen. RCWM has received research grants for clinical trials from AstraZeneca, Bayer, MSD, Novo Nordisk, Sanofi, Roche and Tricida and honoraria for consultancy or lectures from AstraZeneca and Boehringer Ingelheim. APSK has received research grants and/or speaker honoraria from Abbott, AstraZeneca, Bayer, Boehringer Ingelheim, Eli Lilly, Merck Serono, Nestle, Novo Nordisk and Sanofi. JCNC, WYS and RCWM hold patents for using biogenetic markers to predict risk of diabetes and its complications. JCNC, CKPL, RCWM and WYS are co-founders of GemVCare, a biotech start-up supported by the Technology Start-up Support Scheme for Universities of the Hong Kong Government Innovation and Technology Commission. RCWM is an advisory board member of the editorial board of Diabetologia. The authors declare that there are no other relationships or activities that might bias, or be perceived to bias, their work.

Contribution statement JCNC and AOYL conceptualised the research question. CKO and BF performed the data analysis. CKO, BF, JCNC, AOYL, STFT, CHTT, RW, ESHL, MS, CKPL, GY, JPYH, EYKC, APSK, RO, WYS and RCWM contributed to the interpretation of the data. CKO wrote the first draft. All authors critically reviewed the manuscript and approved the final version. JCNC is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Bonnefond A, Unnikrishnan R, Doria A et al (2023) Monogenic diabetes. Nat Rev Dis Primers 9(1):12. https://doi.org/10.1038/s41572-023-00421-w
2. Zhang H, Colclough K, Gloyn AL, Pollin TI (2021) Monogenic diabetes: a gateway to precision medicine in diabetes. J Clin Invest 131(3):e142244. https://doi.org/10.1172/jci142244
3. De Franco E (2020) From biology to genes and back again: gene discovery for monogenic forms of beta-cell dysfunction in diabetes. J Mol Biol 432(5):1535–1550. https://doi.org/10.1016/j.jmb.2019.08.016
4. Chan JC, Lau ES, Luk AO et al (2014) Premature mortality and co-morbidities in young-onset diabetes - a 7 year prospective analysis. Am J Med 127(7):616–624. https://doi.org/10.1016/j.amjmed.2014.03.018
5. Vujkovic M, Keaton JM, Lynch JA et al (2020) Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. Nat Genet 52(7):680–691. https://doi.org/10.1038/s41588-020-0637-y
6. Mahajan A, Taliun D, Thurner M et al (2018) Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet 50(11):1505–1513. https://doi.org/10.1038/s41588-018-0241-6
7. Noordam R, Läll K, Smit RAJ et al (2021) Stratification of type 2 diabetes by age of diagnosis in the UK biobank reveals subgroup-specific genetic associations and causal risk profiles. Diabetes 70(8):1816–1825. https://doi.org/10.2337/db20-0602
8. Christiansen CE, Arathimos R, Pain O, Molokhia M, Bell JT, Lewis CM (2023) Stratified genome-wide association analysis of type 2 diabetes reveals subgroups with genetic and environmental heterogeneity. Hum Mol Genet 32(16):2638–2645. https://doi.org/10.1093/hmg/ddad093
9. Srinivasan S, Chen L, Todd J et al (2021) The first genome-wide association study for type 2 diabetes in youth: the progress in diabetes genetics in youth (ProDiGY) consortium. Diabetes 70(4):996–1005. https://doi.org/10.2337/db20-0443
10. Li JKY, Ng MCY, So WY et al (2006) Phenotypic and genetic clustering of diabetes and metabolic syndrome in Chinese families with type 2 diabetes mellitus. Diabetes Metab Res Rev 22:46–52. https://doi.org/10.1002/dmrr.577
11. Zhang Y, Luk AOY, Chow E et al (2017) High risk of conversion to diabetes in first-degree relatives of individuals with young-onset type 2 diabetes: a 12-year follow-up analysis. Diabet Med 34(12):1701–1709. https://doi.org/10.1111/dme.13516
12. Ko GT, Chan JC, Chan AW et al (2007) Low levels of awareness of suboptimal health conditions in a high-risk working population: the "better health for better Hong Kong" health promotion campaign. Int J Behav Med 14(2):63–69. https://doi.org/10.1007/BF03004170
13. Ko GT, Chan JC, Chan AW et al (2007) Association between sleeping hours, working hours and obesity in Hong Kong Chinese: the "better health for better Hong Kong" health promotion campaign. Int J Obes (Lond) 31(2):254–260. https://doi.org/10.1038/sj.ijo.0803389
14. Ko GT, So WY, Chow CC et al (2010) Risk associations of obesity with sugar-sweetened beverages and lifestyle factors in Chinese: the "Better Health for Better Hong Kong" health promotion campaign. Eur J Clin Nutr 64(12):1386–1392. https://doi.org/10.1038/ejcn.2010.181
15. Piwernetz K, Home PD, Snorgaard O, Antsiferov M, Staehr-Johansen K, Krans M (1993) Monitoring the targets of the St Vincent Declaration and the implementation of quality management in diabetes care: the DIABCARE initiative. The DIABCARE Monitoring Group of the St Vincent Declaration Steering Committee. Diabet Med 10(4):371–377. https://doi.org/10.1111/j.1464-5491.1993.tb00083.x
16. Chan JCN, Lim LL, Luk AOY et al (2019) From Hong Kong diabetes register to JADE program to RAMP-DM for data-driven actions. Diabetes Care 42(11):2022–2031. https://doi.org/10.2337/dci19-0003
17. Flannick J, Mercader JM, Fuchsberger C et al (2019) Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls. Nature 570(7759):71–76. https://doi.org/10.1038/s41586-019-1231-2

18. DeForest N, Majithia AR (2022) Genetics of type 2 diabetes: implications from large-scale studies. Curr Diab Rep 22(5):227–235. https://doi.org/10.1007/s11892-022-01462-3
19. Spracklen CN, Horikoshi M, Kim YJ et al (2020) Identification of type 2 diabetes loci in 433,540 East Asian individuals. Nature 582(7811):240–245. https://doi.org/10.1038/s41586-020-2263-3
20. Locke JM, Saint-Martin C, Laver TW et al (2018) The common HNF1A variant I27L is a modifier of age at diabetes diagnosis in individuals with HNF1A-MODY. Diabetes 67(9):1903–1907. https://doi.org/10.2337/db18-0133
21. Kettunen JLT, Rantala E, Dwivedi OP et al (2022) A multigenerational study on phenotypic consequences of the most common causal variant of HNF1A-MODY. Diabetologia 65(4):632–643. https://doi.org/10.1007/s00125-021-05631-z
22. Jiang G, Luk AO, Tam CHT et al (2020) Obesity, clinical, and genetic predictors for glycemic progression in Chinese patients with type 2 diabetes: a cohort study using the Hong Kong Diabetes Register and Hong Kong Diabetes Biobank. PLoS Med 17(7):e1003209. https://doi.org/10.1371/journal.pmed.1003209
23. Zhou K, Donnelly LA, Morris AD et al (2014) Clinical and genetic determinants of progression of type 2 diabetes: a DIRECT study. Diabetes Care 37(3):718–724. https://doi.org/10.2337/dc13-1995
24. Misra S, Ke C, Srinivasan S et al (2023) Current insights and emerging trends in early-onset type 2 diabetes. Lancet Diabetes Endocrinol 11(10):768–782. https://doi.org/10.1016/s2213-8587(23)00225-5
25. Ke C, Stukel TA, Shah BR et al (2020) Age at diagnosis, glycemic trajectories, and responses to oral glucose-lowering drugs in type 2 diabetes in Hong Kong: a population-based observational study. PLoS Med 17(9):e1003316. https://doi.org/10.1371/journal.pmed.1003316
26. Fan Y, Fan B, Lau ESH et al (2023) Comparison of beta-cell function between Hong Kong Chinese with young-onset type 2 diabetes and late-onset type 2 diabetes. Diabetes Res Clin Pract 205:110954. https://doi.org/10.1016/j.diabres.2023.110954
27. Almgren P, Lehtovirta M, Isomaa B et al (2011) Heritability and familiality of type 2 diabetes and related quantitative traits in the Botnia Study. Diabetologia 54(11):2811–2819. https://doi.org/10.1007/s00125-011-2267-5
28. Looker HC, Chang DC, Baier LJ, Hanson RL, Nelson RG (2023) Diagnostic criteria and etiopathogenesis of type 2 diabetes and its complications: lessons from the Pima Indians. Presse Med 52(1):104176. https://doi.org/10.1016/j.lpm.2023.104176
29. Ma RC, Lee HM, Lam VK et al (2014) Familial young-onset diabetes, pre-diabetes and cardiovascular disease are associated with genetic variants of DACH1 in Chinese. PLoS One 9(1):e84770. https://doi.org/10.1371/journal.pone.0084770
30. Ma RC, Hu C, Tam CH et al (2013) Genome-wide association study in a Chinese population identifies a susceptibility locus for type 2 diabetes at 7q32 near PAX4. Diabetologia 56(6):1291–1305. https://doi.org/10.1007/s00125-013-2874-4
31. Lam VK, Ma RC, Lee HM et al (2013) Genetic associations of type 2 diabetes with islet amyloid polypeptide processing and degrading pathways in asian populations. PLoS One 8(6):e62378. https://doi.org/10.1371/journal.pone.0062378
32. Feiner J, Perrot N, Chong M et al (2023) 221-LB: identification of PAM as novel monogenic diabetes gene. Diabetes 72(Supplement_1):221-LB. https://doi.org/10.2337/db23-221-LB
33. Sanchez Caballero L, Gorgogietas V, Arroyo MN, Igoillo-Esteve M (2021) Molecular mechanisms of β-cell dysfunction and death in monogenic forms of diabetes. Int Rev Cell Mol Biol 359:139–256. https://doi.org/10.1016/bs.ircmb.2021.02.005
34. Ahlqvist E, Storm P, Karajamaki A et al (2018) Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. Lancet Diabetes Endocrinol 6(5):361–369. https://doi.org/10.1016/S2213-8587(18)30051-2
35. Mansour Aly D, Dwivedi OP, Prasad RB et al (2021) Genome-wide association analyses highlight etiological differences underlying newly defined subtypes of diabetes. Nat Genet 53(11):1534–1542. https://doi.org/10.1038/s41588-021-00948-2
36. Zaghlool SB, Halama A, Stephan N et al (2022) Metabolic and proteomic signatures of type 2 diabetes subtypes in an Arab population. Nat Commun 13(1):7121. https://doi.org/10.1038/s41467-022-34754-z
37. Ke C, Lau E, Shah BR et al (2019) Excess burden of mental illness and hospitalization in young-onset type 2 diabetes: a population-based cohort study. Ann Intern Med 170(3):145–154. https://doi.org/10.7326/M18-1900
38. Ma RCW, Xie F, Lim CKP et al (2022) A randomized clinical trial of genetic testing and personalized risk counselling in patients with type 2 diabetes receiving integrated care -the genetic testing and patient empowerment (GEM) trial. Diabetes Res Clin Pract 189:109969. https://doi.org/10.1016/j.diabres.2022.109969