



Clinical variable-based cluster analysis identifies novel subgroups with a distinct genetic signature, lipidomic pattern and cardio-renal risks in Asian patients with recent-onset type 2 diabetes

Jiexun Wang¹ · Jian-Jun Liu¹ · Resham L. Gurung¹ · Sylvia Liu¹ · Janus Lee¹ · Yiamunaa M¹ · Keven Ang¹ · Yi Ming Shao¹ · Justin I-Shing Tang² · Peter I. Benke³ · Federico Torta³ · Markus R. Wenk³ · Subramaniam Tavintharan⁴ · Wern Ee Tang⁵ · Chee Fang Sum⁴ · Su Chi Lim^{4,6,7}

Received: 28 February 2022 / Accepted: 25 April 2022 / Published online: 28 June 2022
© The Author(s) 2022, corrected publication 2022

Abstract

Aims/hypothesis We sought to subtype South East Asian patients with type 2 diabetes by de novo cluster analysis on clinical variables, and to determine whether the novel subgroups carry distinct genetic and lipidomic features as well as differential cardio-renal risks.

Methods Analysis by k-means algorithm was performed in 687 participants with recent-onset diabetes in Singapore. Genetic risk for beta cell dysfunction was assessed by polygenic risk score. We used a discovery–validation approach for the lipidomics study. Risks for cardio-renal complications were studied by survival analysis.

Results Cluster analysis identified three novel diabetic subgroups, i.e. mild obesity-related diabetes (MOD, 45%), mild age-related diabetes with insulin insufficiency (MARD-II, 36%) and severe insulin-resistant diabetes with relative insulin insufficiency (SIRD-RII, 19%). Compared with the MOD subgroup, MARD-II had a higher polygenic risk score for beta cell dysfunction. The SIRD-RII subgroup had higher levels of sphingolipids (ceramides and sphingomyelins) and glycerophospholipids (phosphatidylethanolamine and phosphatidylcholine), whereas the MARD-II subgroup had lower levels of sphingolipids and glycerophospholipids but higher levels of lysophosphatidylcholines. Over a median of 7.3 years follow-up, the SIRD-RII subgroup had the highest risks for incident heart failure and progressive kidney disease, while the MARD-II subgroup had moderately elevated risk for kidney disease progression.

Conclusions/interpretation Cluster analysis on clinical variables identified novel subgroups with distinct genetic, lipidomic signatures and varying cardio-renal risks in South East Asian participants with type 2 diabetes. Our study suggests that this easily actionable approach may be adapted in other ethnic populations to stratify the heterogeneous type 2 diabetes population for precision medicine.

Keywords Beta cell dysfunction · Cardiovascular disease · Chronic kidney disease · Cluster analysis · Heart failure · Lipidomics · Mortality · Polygenic risk score · Type 2 diabetes mellitus

✉ Su Chi Lim
lim.su.chi@ktph.com.sg

¹ Clinical Research Unit, Khoo Teck Puat Hospital, Singapore, Republic of Singapore

² Department of Medicine, Khoo Teck Puat Hospital, Singapore, Republic of Singapore

³ Lipidomics Incubator, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Republic of Singapore

⁴ Diabetes Centre, Admiralty Medical Centre, Singapore, Republic of Singapore

⁵ National Healthcare Group Polyclinic, Singapore, Republic of Singapore

⁶ Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Republic of Singapore

⁷ Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Republic of Singapore

Research in context

What is already known about this subject?

- Type 2 diabetes is a heterogeneous disease in terms of aetiology, pathophysiology and risk for vascular complications
- Clinical variable-based cluster analysis may allow subtyping of patients with type 2 diabetes into novel subgroups with distinct pathophysiological features

What is the key question?

- Can de novo cluster analysis on clinical variables identify novel subgroups in Asian people with type 2 diabetes that are different from those in patients of European descent?

What are the new findings?

- De novo cluster analysis on five clinical variables (diabetes onset age, BMI, HbA_{1c}, HOMA2-B and HOMA2-IR) identified three novel subgroups (mild obesity-related diabetes; mild age-related diabetes with insulin insufficiency, and severe insulin-resistant diabetes with relative insulin insufficiency) in South East Asian people with recent-onset type 2 diabetes that overlapped, but were distinct from, those in patients of European descent
- These novel subgroups demonstrated distinct genetic and lipidomic signatures in terms of genetic risk for beta cell dysfunction and plasma sphingolipid and glycerophospholipid profiles
- Importantly, the subgroups experienced considerably varying risks for cardio-renal complications over 7.3 years of follow-up

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

- The easily actionable data-driven cluster analysis may identify overlapping but distinct novel subgroups in various ethnic groups to stratify the heterogeneous type 2 diabetes population for precision medicine

Abbreviations

ANDIS	All New Diabetics in Scania
CKD	Chronic kidney disease
GWAS	Genome-wide association study
LPC	Lysophosphatidylcholine
MACE	Major adverse cardiovascular events
MARD-II	Mild age-related diabetes with insulin insufficiency
MOD	Mild obesity-related diabetes
PRS	Polygenic risk score
SIDD	Severe insulin-deficient diabetes
SIRD-RII	Severe insulin-resistant diabetes with relative insulin insufficiency

Introduction

The pathogenesis of type 2 diabetes involves a complex interplay between genetic susceptibility and environmental factors [1–3]. Comorbidities such as obesity and dyslipidaemia often co-exist with dysregulation of glucose metabolism. Hence, type 2 diabetes is highly heterogeneous in terms of aetiology, clinical presentation, and risks for vascular and non-vascular complications [4–7]. However, patients with type 2 diabetes

may be subtyped into relatively homogenous subgroups for precision medicine [1].

In a landmark study using data-driven cluster analysis, Ahlqvist et al subtyped recent-onset diabetes into five subgroups based on common clinical variables under the assumption that diabetes is clinically manifested when insulin secretion does not match decreased sensitivity [4, 8]. This novel but easily actionable subtyping approach has attracted tremendous interest, and the clustering algorithm has been replicated in several diabetic populations in recent years [9–13]. While cluster analysis on only a few clinical variables has the advantage of simplicity compared with other approaches using omics data [3], it may be argued that clusters identified from a data correlation matrix are simply the result of inter-dependency in the clinical variables [14]. One approach to address this concern is to examine whether the clusters derived from the common clinical variable have shared pathophysiological features within the subgroup but distinct from the other subgroups. Indeed, a recent study showed that inflammation biomarkers differed greatly across the novel subgroups [15]. Other studies in a European population also identified diabetic subgroups that differed in genetic risk, lipidomic and proteomic signatures [16, 17].

Compared with patients of European descent, Asians with type 2 diabetes have more severe adiposity at the same level of BMI, develop diabetes at a younger age, and demonstrate impaired insulin secretion to compensate for insulin resistance [18, 19]. South East Asia has a large population with type 2 diabetes due to the dramatic socioeconomic transition over recent decades [20]. However, to our knowledge, data on clinical variable-based cluster analysis from this region are still scarce. Most early replication studies used cluster coordinates derived from the ANDIS cohort (All New Diabetics in Scania) [8], rather than de novo cluster analysis, to subtype type 2 diabetes [12, 13, 21–23]. We hypothesise that analysis on the same set of clinical variables used in the ANDIS cohort may identify novel subgroups in Asian patients that differ from those in patients of European descent.

In the current study, we performed de novo cluster analysis in patients with recent-onset type 2 diabetes in Singapore, a city state in South East Asia with a mix of three major ethnic populations. We wished to determine whether the newly identified subgroups differ in aetiology and pathophysiology from the perspective of genetics and lipidomics. Importantly, we sought to determine whether these novel subgroups predict risks for cardio-renal complications over long-term follow-up.

Methods

Research design We focused on individuals with a diabetes duration of less than 5 years in the current study because subgroup assignment derived from cluster analysis has been shown to be relatively stable within 5 years after diabetes onset [12]. Details of the cohort used (SMART2D, Singapore Study of Macro-Angiopathy and Micro-Vascular Reactivity in Type 2 Diabetes) have been described elsewhere [24]. Briefly, 2057 participants with type 2 diabetes were recruited from outpatient clinics in a secondary hospital and an adjacent primary care medical facility between 2011 and 2014. Type 2 diabetes was diagnosed by the attending physicians after excluding type 1 diabetes and diabetes attributable to specific causes. Type 1 diabetes was diagnosed as sustained requirement for insulin treatment within 1 year after diabetes diagnosis without measurement of GAD antibody. Patients with cancer and autoimmune disease on active treatments, and those with HbA_{1c} >12% (108 mmol/mol) at screening were also excluded from the cohort. Participants were recalled for a research visit every 3 years, and also followed up by reviewing electronic health records [25]. All 687 participants with diabetes duration ≤5 years and eGFR ≥15 ml min⁻¹ 1.73 m⁻² were included in the current analysis.

This study was approved by the Singapore National Healthcare Group Domain Specific Review Board and all participants provided written informed consent.

Clinical and biochemical variables Diabetes duration was self-reported. Blood pressure was measured three times using a semi-automated blood pressure monitor, and the mean value was used. Fasting plasma glucose, HDL- and LDL-cholesterol and triacylglycerols were quantified by enzymatic methods (Roche Cobas Integra 700; Roche Diagnostics, Basel, Switzerland). HbA_{1c} was measured using a point-of-care analyser (DCA Vantage Analyzer; Siemens, Munich, Germany). Serum creatinine was measured using an enzymatic method, and GFR was estimated using the CKD-EPI equation [26]. Urinary albumin was quantified using an immunoturbidimetric assay (Roche Cobas c, Roche Diagnostics, Mannheim, Germany). Plasma C-reactive protein was quantified using an immunoassay kit (R&D Systems, Minneapolis, MN, USA). Fasting plasma C-peptide was measured using an ELISA kit (Mercodia, Uppsala, Sweden). Both intra- and inter-assay CVs were <5%. HOMA2-B (%) and HOMA2-IR were calculated based on fasting glucose and C-peptide (<https://www.dtu.ox.ac.uk/homacalculator/>, version 2.2.3).

Cluster analysis We applied the k-means algorithm as proposed by Ahlqvist et al to divide participants into subgroups [8]. Five clinical classifiers (diabetes onset age, BMI, HbA_{1c}, log-transformed HOMA2-B and HOMA2-IR) were standardised to a mean value of 0 and SD of 1. The optimal number of clusters was determined by majority voting according to 26 indices provided by the R package ‘NbClust’. Cluster stability was assessed by the Jaccard index based on bootstrapping [27].

Beta cell dysfunction, insulin resistance and type 1 diabetes polygenic risk scores Genome-wide association study (GWAS) and principal component analysis on GWAS arrays in participants of the cohort have been described before [28]. We created polygenic risk scores (PRSs) for beta cell dysfunction and insulin resistance based on 35 SNPs associated with insulin secretion and 20 SNPs associated with insulin sensitivity, respectively, in Asian populations. We weighted the SNPs by their effect on the risk of type 2 diabetes to determine whether the novel subgroups differ in genetic risk for type 2 diabetes development (see electronic supplementary material [ESM] Table 1) [29]. A high score indicates more severe beta cell dysfunction and insulin resistance. A type 1 diabetes PRS was constructed by a similar approach using nine SNPs (ESM Table 2) [30, 31]. Details on PRS derivation are given in ESM Methods. We fitted linear regression models to compare the differences in PRS across the three subgroups, in which the score was entered as a dependent variable and subgroup membership, sex and scores for the top three principal components in lieu of self-reported ethnicity were entered as covariates.

Lipidomics assay and data analysis

We adopted a discovery–validation approach for the lipidomics study to reduce the likelihood of false positives due to multiple comparisons. The validation study was nested in an independent cohort that has been described previously [32]. In brief, 226 participants with diabetes duration ≤ 5 years and $\text{eGFR} \geq 15 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$ were randomly selected. As measurements of HOMA2-IR and HOMA2-B were not available for the validation cohort, we used the plasma triacylglycerol/HDL-cholesterol ratio as a proxy for insulin resistance [4, 33]. Using the ‘reference’ approach [4], the coordinate of the cluster centre in the discovery cohort was calculated as the mean value of BMI, diabetes onset age, HbA_{1c} and the plasma triacylglycerol/HDL-cholesterol ratio, and participants in the validation cohort were assigned cluster membership based on minimal Euclidean distance.

Technical details for the lipidomics assay by LC-MS are described in ESM Methods. A total of 315 lipid species were included in the discovery study after excluding those with a signal-to-noise ratio < 3 and correcting for batch effect. We applied the Kruskal–Wallis test to compare the levels of lipid species across the three subgroups. Those with p values below the Bonferroni correction threshold ($p < 1.59 \times 10^{-4}$, $0.05/315$) were subjected to the Kruskal–Wallis test in the validation cohort, and a nominal p value < 0.05 was considered statistically significant. We plotted a heatmap to visualise lipid species that differed across diabetes subgroups in both discovery and validation cohorts. Furthermore, we fitted linear regression models to compare differences in lipid species between two subgroups, in which log-transformed lipid concentration was entered as a dependent variable and subgroup membership as an independent variable.

Identification of adverse clinical outcomes and statistical analysis All-cause mortality was identified from electronic medical records and cross-validated against the national death registry [34]. Cardiovascular death was identified from death certificates. Non-fatal acute myocardial infarction and stroke were identified from hospitalisation discharge summaries and surgical operation procedures. Major adverse cardiovascular events (MACE) were a composite of non-fatal acute myocardial infarction, stroke and death attributable to cardiovascular disease, whichever occurred first. Ascertainment of incident heart failure has been described previously [35]. Progressive chronic kidney disease (CKD) was defined as a decrease in eGFR of 40% or more from the baseline level, with repeated measurements at least 3 months apart as confirmation [36]. The follow-up was censored at 30 November 2019.

Incidence rates for progressive CKD, incident heart failure, MACE and all-cause mortality are presented as event number per 1000 person-years. The cumulative risk for cardio-renal events was plotted by the Kaplan–Meier approach and

compared by logrank test. We fitted Cox proportional hazard regression models to study the associations of subgroup with cardio-renal risks in the follow-up period. Index age, sex and ethnicity were included as covariates in the models. We also adjusted baseline eGFR for the study on progressive CKD. The proportional hazards assumption was tested based on Schoenfeld residuals. No violation of proportional hazard assumption was identified.

Results

Data-driven cluster analysis identified three novel subgroups in participants with recent-onset type 2 diabetes A total of 687 individuals with recent-onset diabetes were subjected to cluster analysis [8]. Majority voting according to 26 indices suggested that the study population may be optimally partitioned into three subgroups (ESM Fig. 1). The mean values of Jaccard indices were above 0.86 for all three clusters based on 5000 bootstraps, implying that the clusters were stable. Participant baseline characteristics in the three subgroups are presented in Table 1 and ESM Fig. 2.

Cluster 1 (45% of participants) was labelled as mild obesity-related diabetes (MOD). Participants in this subgroup had a high BMI ($30.1 \pm 5.0 \text{ kg/m}^2$), an elevated HOMA2-IR (median 2.2, IQR 1.8–2.9) and preserved insulin secretion as evidenced by high levels of fasting C-peptide and HOMA2-B (median 97%, IQR 78–132%).

Cluster 2 (19.0% of participants) was labelled as severe insulin-resistant diabetes with relative insulin insufficiency (SIRD-RII). These participants had the highest level of HOMA2-IR (median 2.7, IQR 1.9–3.8), the highest BMI ($31.6 \pm 5.9 \text{ kg/m}^2$) and the worst glycaemic control among the three subgroups. They also had the highest level of triacylglycerols, the lowest level of HDL-cholesterol, the highest level of C-reactive protein and the youngest age at diabetes diagnosis. Their HOMA2-B index was low (median 44%, IQR 32–62%) but their fasting C-peptide remained at a high level compared with the other two subgroups.

Cluster 3 (36% of participants) was labelled as mild age-related diabetes with insulin insufficiency (MARD-II). These participants were slightly older (56 ± 9.9 years) and had a low HOMA2-B index (median 53%, IQR 40–66%) at diabetes diagnosis. Their fasting C-peptide was 45% lower than the other two subgroups. They had no overt obesity ($\text{BMI} 24.9 \pm 3.4 \text{ kg/m}^2$) and only moderately elevated HOMA2-IR (median 1.3, IQR 0.9–1.6).

The MOD subgroup was taken as the reference in the subsequent analyses because it was the largest subgroup in the study population.

High PRS for beta cell dysfunction in the MARD-II subgroup Compared with the MOD subgroup, the participants in the

Table 1 Baseline characteristics of three subgroups derived from k-means cluster analysis in individuals with recent-onset type 2 diabetes ($n=687$)

Variable	MOD (45%)	SIRD-RII (19%)	MARD-II (36%)	<i>p</i> value
Number of participants	307	130	250	–
Index age (years)	55 ± 9.6	43 ± 11.5	59 ± 9.9	<0.001
Age at diabetes onset (years)	52 ± 9.4	40 ± 10.9	56 ± 9.9	<0.001
Male	144 (46.9)	74 (56.9)	131 (52.4)	0.13
Ethnicity				<0.001
Chinese	123 (40.1)	62 (47.7)	157 (62.8)	
Malay	104 (33.9)	37 (28.5)	49 (19.6)	
Asian Indian	80 (26.1)	31 (23.8)	44 (17.6)	
Diabetes duration (years)	3.0 (1.0–4.0)	3.0 (2.0–5.0)	3.0 (1.0–5.0)	0.19
BMI (kg/m ²)	30.1 ± 5.0	31.6 ± 5.9	24.9 ± 3.4	<0.001
HbA _{1c} (%)	6.9 ± 0.7	9.2 ± 1.2	7.1 ± 0.9	<0.001
HbA _{1c} (mmol/mol)	52 ± 5.3	77 ± 10	54 ± 6.8	–
Fasting plasma glucose (mmol/l)	6.5 ± 1.4	10.7 ± 2.6	7.4 ± 1.7	<0.001
Fasting C-peptide (pmol/l)	931 (771–1185)	920 (709–1305)	523 (386–644)	<0.001
HOMA2-B (%)	96.9 (77.5–131.7)	43.7 (32.4–61.9)	53.3 (40.0–66.0)	<0.001
HOMA2-IR	2.2 (1.8–2.9)	2.7 (1.9–3.8)	1.3 (0.9–1.6)	<0.001
Triacylglycerol/HDL ratio	1.1 (0.8–1.8)	1.6 (1.1–2.3)	0.9 (0.6–1.4)	<0.001
Blood pressure (mmHg)				
Systolic	137 ± 16.8	136 ± 14.9	139.3 ± 18.0	0.07
Diastolic	80.2 ± 9.3	82.0 ± 9.2	78.3 ± 9.5	0.001
Kidney function				
eGFR (ml min ⁻¹ 1.73 m ⁻²)	92 ± 19.2	109 ± 19.7	92 ± 21.0	<0.001
Urine ACR (mg/mmol)	1.5 (0.5–4.1)	2.7 (0.7–8.4)	1.1 (0.1–4.6)	<0.001
Lipid profile, mmol/l				
HDL-cholesterol	1.3 ± 0.4	1.2 ± 0.3	1.4 ± 0.4	<0.001
LDL-cholesterol	2.8 ± 0.8	3.3 ± 0.9	2.8 ± 0.8	<0.001
Triacylglycerol	1.4 (1.1–1.9)	1.8 (1.3–2.4)	1.2 (0.9–1.6)	<0.001
C-reactive protein (µg/ml)	2.7 (1.1–5.8)	3.8 (2.0–7.9)	1.5 (0.5–3.5)	<0.001
Medication usage				
Metformin	243 (79.2)	118 (90.8)	204 (81.6)	0.02
Sulfonylurea	127 (41.4)	63 (48.5)	100 (40.0)	0.27
DPP4 inhibitor	7 (2.3)	10 (7.7)	6 (2.4)	0.01
Insulin	14 (4.6)	25 (19.2)	25 (10.0)	<0.001
Statins	232 (75.6)	92 (70.8)	183 (73.2)	0.49
RAS blocker	141 (45.9)	65 (50.0)	116 (46.4)	0.79

Data are presented as means ± SD, median (IQR) or *n* (%)

ACR, albumin/creatinine ratio; DPP4, dipeptidyl peptidase 4; RAS, renin–angiotensin system

Table 2 Association of PRS with subgroup membership

PRS	Beta cell dysfunction coefficient (95% CI)	<i>p</i> value	Insulin resistance coefficient (95% CI)	<i>p</i> value
MOD	Reference		Reference	
SIRD-RII	1.00 (–0.31, 2.31)	0.14	–0.15 (–0.85, 0.56)	0.68
MARD-II	1.39 (0.32, 2.47)	0.01	–0.11 (–0.69, 0.46)	0.70
Male sex	Reference		Reference	
Female sex	1.30 (0.34, 2.25)	0.01	0.28 (–0.23, 0.79)	0.28

For the linear regression models, the PRS was the dependent variable and cluster membership (MOD subgroup as reference), sex and GWAS principal components 1–3 were used as covariates

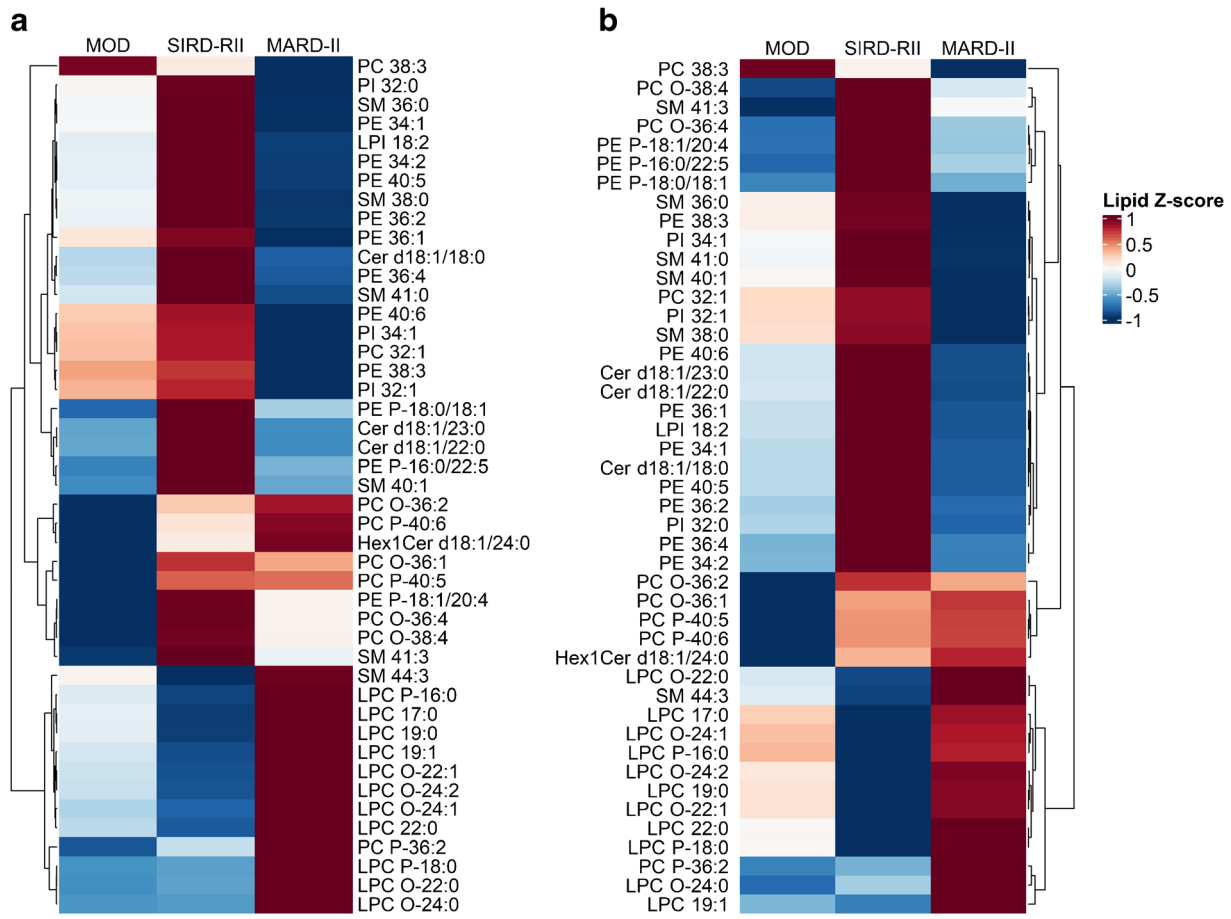


Fig. 1 Plasma lipid species that differed significantly across the three subgroups in the discovery cohort (a) and the validation cohort (b). MOD, moderate albuminuria; SIRD-RII, severe insulin resistance and dyslipidemia; MARD-II, moderate albuminuria and dyslipidemia; PC, phosphatidylcholine; LPI, lysophosphatidylinositol; PE, phosphatidylethanolamine; PI, phosphatidylinositol; Cer, ceramide; Hex1Cer, monohexosylceramide; LPC, lysophosphatidylcholine

Table 3 Association of novel subgroup with adverse cardiovascular risk during the follow-up period by Cox proportional hazard regression

Outcome	Unadjusted HR (95% CI)	p value	Adjusted HR ^a (95% CI)	p value
Progressive CKD				
MOD	Reference		Reference	
SIRD-RII	2.33 (1.05, 5.18)	0.04	3.67 (1.53, 8.8)	0.004
MARD-II	1.84 (0.88, 3.85)	0.11	2.64 (1.20, 5.82)	0.02
Heart failure				
MOD	Reference		Reference	
SIRD-RII	2.30 (1.08, 4.89)	0.03	5.23 (2.35, 11.6)	<0.001
MARD-II	0.99 (0.45, 2.17)	0.97	0.87 (0.39, 1.98)	0.75
MACE				
MOD	Reference		Reference	
SIRD-RII	0.93 (0.39, 2.22)	0.87	1.92 (0.75, 4.89)	0.17
MARD-II	0.69 (0.32, 1.50)	0.35	0.61 (0.27, 1.36)	0.23
All-cause mortality				
MOD	Reference		Reference	
SIRD-RII	1.12 (0.48, 2.59)	0.80	2.99 (1.22, 7.30)	0.02
MARD-II	1.11 (0.56, 2.23)	0.76	0.78 (0.38, 1.62)	0.51

^a Age, sex and ethnicity were adjusted for outcomes of all-cause mortality, incident heart failure and MACE in Cox proportional hazard regression models. Baseline eGFR was also adjusted for outcome of progressive CKD

MARD-II subgroup had a significantly higher PRS for beta cell dysfunction after adjustment for sex and GWAS principal components 1–3 (Table 2). There was no significant difference in the PRS for beta cell dysfunction between the SIRD-RII and MOD subgroups, and no significant difference in the PRS for insulin resistance among the three subgroups.

Distinct lipidomic patterns across the three subgroups The clinical profiles of the discovery and validation cohorts were comparable (ESM Tables 3 and 4). Of the 315 lipid species included in the discovery study, 75 differed across the three subgroups (p value $<1.59 \times 10^{-4}$), and 45 of them also differed significantly across the three subgroups in the validation cohort (nominal $p < 0.05$, ESM Fig. 3). The SIRD-RII subgroup had high levels of glycerophospholipids, mainly phosphatidylethanolamine, phosphatidylcholine and phosphatidylinositol subspecies, but lower levels of lysophosphatidylcholine (LPC), including subspecies with alkyl ether and plasmalogen bonds. They also had remarkably high level of sphingolipids (sphingomyelins and

ceramides). In contrast, the MARD-II subgroup had low levels of glycerophospholipids and sphingolipids but higher levels of LPC (Fig. 1). The subsequent between-group comparisons identified 17 lipid species, mainly phosphatidylethanolamine, phosphatidylcholine, ceramides and sphingomyelins, that differed between the SIRD-RII and MOD subgroups. The phosphatidylethanolamine, phosphatidylinositol, phosphatidylcholine and sphingomyelin levels were lower in the MARD-II subgroup compared with the MOD subgroup, and levels of LPC subspecies were higher (ESM Tables 5 and 6).

Risks for cardio-renal complications in the three subgroups during follow-up The median follow-up duration was 7.3 years (IQR 6.7–7.7). The crude incident rates for progressive CKD, incident heart failure, MACE and all-cause mortality are shown in ESM Table 7. The incident rate for heart failure in the SIRD-RII subgroup (14.5 per 1000 person-years; 95% CI 7.7, 24.8) was twofold higher than that in the other subgroups: 6.4 for the MOD subgroup (95% CI 3.5, 10.7) and 6.3 for the MARD-II

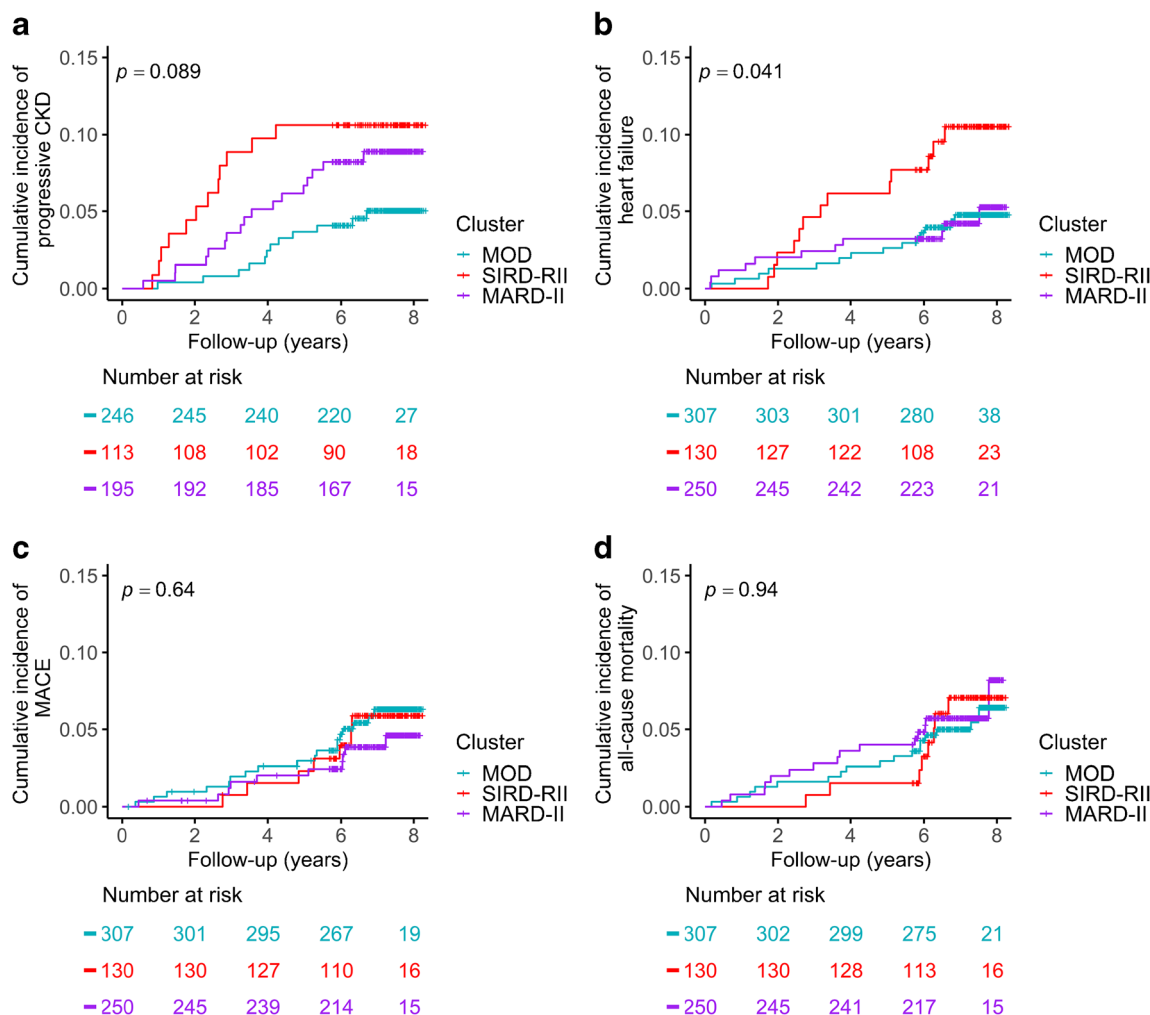


Fig. 2 Cumulative incidence of adverse clinical outcomes by diabetes subgroups: (a) progressive CKD, (b) incident heart failure, (c) MACE, and (d) all-cause mortality. Only participants with baseline eGFR above $60 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$ were included in the analysis on progressive CKD

subgroup (95% CI 3.1, 11.2). Additionally, the SIRD-RII subgroup had the highest risk for progressive CKD, followed by the MARD-II and MOD subgroups: 15.9 (95% CI 8.2, 27.8), 12.7 (95% CI 7.4, 20.3) and 6.8 (95% CI 3.5, 11.9) per 1000 person-years, respectively.

Cumulative incidences were plotted for visualisation by the Kaplan–Meier approach (Fig. 2). Cox proportional hazard regression models suggested that the SIRD-RII subgroup had a 2.30-fold unadjusted risk (95% CI 1.08, 4.89) for heart failure compared with the MOD subgroup. Adjustment for index age, sex and ethnicity strengthened the association (adjusted HR 5.23; 95% CI 2.35, 11.60). The SIRD-RII subgroup had a 2.33-fold unadjusted risk (95% CI 1.05, 5.18) and a 3.67-fold adjusted risk (95% CI 1.53, 8.80) for progressive CKD, with adjusted and unadjusted hazard ratios of 1.84 (95% CI 0.88, 3.85) and 2.64 (95% CI 1.20, 5.82) in the MARD-II subgroup compared with the MOD subgroup. Additionally, the SIRD-RII subgroup had the same unadjusted risks for MACE and all-cause mortality as the other subgroups, despite being more than 10 years younger and having similar diabetes duration. Further analysis suggested that they had a 2.99-fold age-adjusted risk (95% CI 1.22, 7.30) for all-cause mortality compared with the MOD subgroup (Table 3).

Additional analyses To assess whether the participant's sex affects cluster analysis, we regressed sex on clinical variables and used regression residuals as the new classifiers [37]. This new analysis also partitioned participants into three clusters, and the cluster membership showed high agreement with that in the primary analysis (approximately 90% concordance, ESM Table 8). We also clustered participants into four subgroups according to centroids derived from the ANDIS cohort [8]. As shown in ESM Figs 4 and 5, 67% of the participants in the severe insulin-deficient diabetes (SIDD) subgroup were from the SIRD-RII subgroup although they did not have a significantly lower level of fasting C-peptide. In the follow-up period, the SIDD subgroup had the highest risk for progressive CKD (ESM Table 9). This finding was different from that of previous studies, which showed that the SIRD group had the highest risk for progressive CKD [8, 10, 12, 23]. Participants with a type 1 diabetes PRS in the top five percentiles had a slightly lower HOMA2-B (57% vs 69%, $p=0.02$) compared with those in the remaining 95 percentiles. However, fasting C-peptide did not differ between the two groups ($p=0.19$). As shown in ESM Table 10, 9% and 7% participants in the SIRD-RII and MARD-II subgroups, respectively, were classified as having a high type 1 diabetes PRS.

Discussion

By applying the k-means algorithm on the same clinical variables as proposed by the previous study in a European

population [8], we identified three novel subgroups in participants with recent-onset type 2 diabetes in our South East Asian cohort. The largest subgroup (MOD, 45% of participants) is characterised by mild obesity, insulin resistance and preserved insulin secretion. The second largest subgroup (MARD-II, 36% of participants) is characterised by a slightly older age of diabetes onset and low beta cell secretion with no overt insulin resistance. The third subgroup (SIRD-RII, 19% of participants) is characterised by severe insulin resistance, poor glycaemic control and relative insulin insufficiency as indicated by preserved insulin secretion but a low HOMA2-B. Our genetics and lipidomics studies suggest a significant difference in genetic risks for diabetes aetiology and distinct pathophysiological features in the three subgroups. Importantly, we demonstrate that the clinical variable-based cluster analysis may potentially stratify patients by risk for cardio-renal complications after diabetes onset.

As hypothesised, we identified diabetes subgroups with overlapping but distinct characteristics in our South East Asian population compared with patients of European descent. In the landmark study by Ahlqvist et al [8], and also in subsequent replication studies in the European and US populations [10, 12, 21], the largest subgroup is MARD (approximately 40%), followed by the MOD subgroup (approximately 20%) and the SIRD subgroup (approximately 20%). In the present study, the largest subgroup is MOD (45%), followed by the MARD-II subgroup (36%). The dominance of obesity-related diabetes in this Asian study population may be attributable to the socioeconomic transformation and concurrent rapid increase in the prevalence of obesity over recent decades in this population. Our study indicates that de novo cluster analysis is warranted to subtype heterogeneous type 2 diabetes patients in various ethnic populations. As shown in ESM Fig. 5, patients in our SIDD subgroup derived from the ANDIS cohort centroids do indeed have a low HOMA2-B. However, their C-peptide level is close to the mean value of the full cohort, suggesting that these participants do not have insulin deficiency.

Elucidating the aetiology of diabetes may shed light on strategies for diabetes prevention and treatment. Our data suggest that obesity and the related insulin resistance are the main driving factor for diabetes pathogenesis in this Asian study population. This highlights the importance for prevention and treatment of obesity to slow down the rising prevalence of diabetes in this region [20]. However, the MARD-II subgroup has neither overt obesity nor insulin resistance. Instead, they are characterised by a 40% lower fasting C-peptide and a low HOMA2-B. These features are different in patients of European descent, in whom the HOMA2-B index in this subgroup remained high [8]. Patients in this subgroup have a higher PRS for beta cell dysfunction, which suggests the presence of impaired beta cell function determined by genetic background. This is in agreement with a large study

showing that patients with type 2 diabetes who have a high number of beta cell dysfunction-related genetic loci have a reduced fasting C-peptide level [3].

Participants in the SIRD-RII subgroup had the highest HOMA2-IR and the worst glycaemic control. The high levels of BMI, triacylglycerol and C-reactive protein, low level of HDL-cholesterol, and the markedly increased levels of sphingo- and glycerophospholipids also support the presence of severe insulin resistance. The low HOMA2-B in this subgroup should be interpreted in the presence of high HOMA2-IR and uncontrolled hyperglycaemia. We reasoned that participants in this subgroup do not have absolute insulin deficiency because their fasting C-peptide was at a level comparable to that in the MOD subgroup and 40% higher than that in the MARD-II subgroup. Instead, the low level of HOMA2-B suggests that the beta cell secretion capacity is unable to adequately compensate for severe insulin resistance [2]. Hence, we designated them as having ‘relative insulin insufficiency’. This is in agreement with our genetic study, which did not find a higher PRS for beta cell dysfunction in this subgroup. We postulate that the relatively insufficient insulin secretion in this subgroup may be partly attributable to glucotoxicity and lipotoxicity, given the uncontrolled hyperglycaemia, overt dyslipidaemia and abnormal lipidomics profile. Intriguingly, we did not observe a significant difference in the PRS for insulin resistance across the three subgroups. This may suggest that genetic risk is not the main determinant for insulin resistance in the SIRD-RII subgroup. However, it may be more reasonable to attribute the null analytical outcome to the relatively small sample size in the current study.

Plasma lipidomic signatures have been associated not only with the risk for type 2 diabetes pathogenesis but also the risk for diabetic complications [38]. Compared with the MOD subgroup, the SIRD-RII subgroup is characterised by activation of the ceramide/sphingomyelin pathway and remodelling of glycerophospholipid metabolism, but the levels of these two classes of lipids were lower in the MARD-II subgroup. Both sphingolipid and glycerophospholipid metabolism have been associated with insulin resistance [39, 40], supporting the consistency between the lipidomic signature and clinical phenotype in the current study. On the other hand, LPC level was higher in the SIRD-RII subgroup but lower in the MARD-II subgroup. The pathophysiological mechanisms underlying the contrasting pattern of LPC between these two subgroups remain unknown.

In agreement with previous studies, the clinical variable-derived subgroups also demonstrated distinct cardio-renal risks in our cohort [1]. The SIRD-RII subgroup shows the highest cardio-renal risks, as manifested by a significantly higher risk for heart failure and progressive CKD. Our finding is consistent with data from the ANDIS cohort, the German Diabetes Study and a retrospective study in a Japanese

population, which showed that the SIRD subgroup was at increased risk for progressive CKD [8, 11, 12]. We extended these previous studies by showing that heart failure may be another important adverse clinical outcome associated with SIRD that warrants further studies. The excessive cardio-renal risk in the SIRD-RII subgroup is attributable to the more severe clinical risk factors, including poor glycaemic control and dyslipidaemia, as well as novel risk factors such as activation of the ceramide pathway and increased inflammation tone [40]. Interestingly, we also observed that the MARD-II subgroup had a moderately elevated risk for progressive CKD. The mean age in this subgroup is only 4 years older than that in the MOD subgroup. Hence, ageing is unlikely to account for the difference in CKD risk. It is possible that diabetes may remain undiagnosed for a longer period in the MARD-II subgroup due to the less pronounced metabolic risk profile, but this is speculative.

Our study adds evidence to support that data-driven cluster analysis upon clinical variables may set the foundation for precision medicine. In addition to diabetes prevention and healthcare resource allocation, it may have implications for precision of medication treatment. The MARD-II subgroup has a low insulin secretion and high PRS for beta cell dysfunction. Patients in this subgroup may better respond to insulin secretagogues, as shown in the ADOPT trial [21]. On the other hand, patients in the SIRD-RII and MOD subgroups may respond better to medications that improve insulin sensitivity beyond and above interventions for weight loss. The SIRD-RII subgroup may benefit from early administration of sodium–glucose cotransporter 2 inhibitors and glucagon-like peptide 1 receptor agonists, given the high cardio-renal risk in this subgroup [41].

The strengths of the current study include a well-characterised cohort with a long follow-up. We included only participants with recent-onset type 2 diabetes to partly mitigate confounding, and used the same clustering algorithm on the same clinical variables to enable reasonable comparison between this Asian study population and patients of European descent. Nevertheless, several important weaknesses must be highlighted. First, the sample size is moderate, and thus our study on the insulin resistance PRS may be underpowered. Second, although we have excluded type 1 diabetes on the basis of clinical criteria, we did not measure GAD antibody. As shown in our analyses using the type 1 diabetes PRS, we cannot exclude the possibility that a small proportion of participants may have autoimmune-related diabetes. Third, we used the triacylglycerol/HDL-cholesterol ratio instead of HOMA indices to cluster the validation cohort for the lipidomics study. Although this may be considered a reasonable approach, as shown by a comparable cardio-renal risk profile (ESM Fig. 6), the concordance of subgroup assignments derived from two sets of classifying variables is moderate (Cohen’s kappa 0.72, ESM Table 11), especially for the

MOD and MARD-II subgroups. Finally, we have not performed an external validation. Hence, the generalisability of our findings should be assessed in future studies.

In summary, cluster analysis identified three subgroups of patients with recent-onset type 2 diabetes in our South East Asian cohort. These subgroups demonstrate not only distinct clinical phenotypes but also differences in genetic aetiology, pathophysiology features and cardio-renal risks. Together with previous studies in other populations, our data suggest that cluster analysis on clinically available variables may be used as a starting point to stratify the heterogeneous diabetic population into subgroups for precision medicine.

Supplementary Information The online version contains peer-reviewed but unedited supplementary material available at <https://doi.org/10.1007/s00125-022-05741-2>.

Acknowledgements We thank participants in the SMART2D cohort and all staff in the clinical research unit at Khoo Teck Puat Hospital, Singapore, for their contributions to the study.

Data availability The datasets generated and analysed during the current study are not publicly available. However, anonymised data are available from the corresponding author upon reasonable request.

Funding The work was supported by Singapore National Medical Research Council grants (CSA-INV/0020/2017, CS-IRG MOH-000066, OFLCG/001/2017), the Singapore Alexandra Health Research Program, and Science–Translational and Applied Research (STAR) grants 18203 and 20201. The funding bodies played no roles in study design, data collection and interpretation, and the decision to submit the research for publication.

Authors' relationships and activities The authors declare that there are no relationships or activities that might bias, or be perceived to bias, their work.

Contribution statement SCL, JW and J-JL designed the study. JW, J-JL and RLG researched the data. SL, JL, YM, KA, YMS, JIT, PIB, FT, MRW, ST, WET, CFS and SCL collected data and contributed important intellectual knowledge. J-JL and JW drafted the manuscript, and all other co-authors revised the manuscript critically for important intellectual content, approved publication of the manuscript, and account for all aspects of the work. SCL is the guarantor of this work, and, as such, had full access to all the data in the study and takes responsibility for integrity of data and accuracy of data analysis.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Chung WK, Erion K, Florez JC et al (2020) Precision medicine in diabetes: a Consensus Report from the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetologia* 63(9):1671–1693. <https://doi.org/10.1007/s00125-020-05181-w>
2. Kahn SE, Chen YC, Esser N et al (2021) The β Cell in Diabetes: Integrating Biomarkers With Functional Measures. *Endocr Rev* 42(5):528–583. <https://doi.org/10.1210/endo/bnab021>
3. Udler MS, Kim J, von Grotthuss M et al (2018) Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: A soft clustering analysis. *PLoS Med* 15(9):e1002654. <https://doi.org/10.1371/journal.pmed.1002654>
4. Ahlqvist E, Prasad RB, Groop L (2020) Subtypes of Type 2 Diabetes Determined From Clinical Parameters. *Diabetes* 69(10):2086–2093. <https://doi.org/10.2337/dbi20-0001>
5. Bancks MP, Chen H, Balasubramanyam A et al (2021) Type 2 Diabetes Subgroups, Risk for Complications, and Differential Effects Due to an Intensive Lifestyle Intervention. *Diabetes Care* 44(5):1203–1210. <https://doi.org/10.2337/dc20-2372>
6. Eckel RH, Bornfeldt KE, Goldberg IJ (2021) Cardiovascular disease in diabetes, beyond glucose. *Cell Metab* 33(8):1519–1545. <https://doi.org/10.1016/j.cmet.2021.07.001>
7. McCarthy MI (2017) Painting a new picture of personalised medicine for diabetes. *Diabetologia* 60(5):793–799. <https://doi.org/10.1007/s00125-017-4210-x>
8. 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](https://doi.org/10.1016/S2213-8587(18)30051-2)
9. Anjana RM, Baskar V, Nair ATN et al (2020) Novel subgroups of type 2 diabetes and their association with microvascular outcomes in an Asian Indian population: a data-driven cluster analysis: the INSPIRED study. *BMJ Open Diabetes Res Care* 8(1):1506. <https://doi.org/10.1136/bmjdr-2020-001506>
10. Slieker RC, Donnelly LA, Fitipaldi H et al (2021) Replication and cross-validation of type 2 diabetes subtypes based on clinical variables: an IMI-RHAPSODY study. *Diabetologia* 64(9):1982–1989. <https://doi.org/10.1007/s00125-021-05490-8>
11. Tanabe H, Saito H, Kudo A et al (2020) Factors Associated with Risk of Diabetic Complications in Novel Cluster-Based Diabetes Subgroups: A Japanese Retrospective Cohort Study. *J Clin Med* 9(7):2083. <https://doi.org/10.3390/jcm9072083>
12. Zaharia OP, Strassburger K, Strom A et al (2019) Risk of diabetes-associated diseases in subgroups of patients with recent-onset diabetes: a 5-year follow-up study. *Lancet Diabetes Endocrinol* 7(9):684–694. [https://doi.org/10.1016/S2213-8587\(19\)30187-1](https://doi.org/10.1016/S2213-8587(19)30187-1)
13. Zou X, Zhou X, Zhu Z, Ji L (2019) Novel subgroups of patients with adult-onset diabetes in Chinese and US populations. *Lancet Diabetes Endocrinol* 7(1):9–11. [https://doi.org/10.1016/S2213-8587\(18\)30316-4](https://doi.org/10.1016/S2213-8587(18)30316-4)
14. Lugner M, Gudbjornsdottir S, Sattar N et al (2021) Comparison between data-driven clusters and models based on clinical features to predict outcomes in type 2 diabetes: nationwide observational study. *Diabetologia* 64(9):1973–1981. <https://doi.org/10.1007/s00125-021-05485-5>
15. Herder C, Maalmi H, Strassburger K et al (2021) Differences in Biomarkers of Inflammation Between Novel Subgroups of Recent-Onset Diabetes. *Diabetes* 70(5):1198–1208. <https://doi.org/10.2337/db20-1054>
16. Slieker RC, Donnelly LA, Fitipaldi H et al (2021) Distinct Molecular Signatures of Clinical Clusters in People With Type 2

- Diabetes: An IMI-RHAPSODY Study. *Diabetes* 70(11):2683–2693. <https://doi.org/10.2337/db20-1281>
17. Zaharia OP, Strassburger K, Knebel B et al (2020) Role of Patatin-Like Phospholipase Domain-Containing 3 Gene for Hepatic Lipid Content and Insulin Resistance in Diabetes. *Diabetes Care* 43(9):2161–2168. <https://doi.org/10.2337/dc20-0329>
 18. Kong AP, Xu G, Brown N, So WY, Ma RC, Chan JC (2013) Diabetes and its comorbidities—where East meets West. *Nat Rev Endocrinol* 9(9):537–547. <https://doi.org/10.1038/nrendo.2013.102>
 19. Ma RCW (2018) Epidemiology of diabetes and diabetic complications in China. *Diabetologia* 61(6):1249–1260. <https://doi.org/10.1007/s00125-018-4557-7>
 20. International Diabetes Federation (2017) IDF Diabetes Atlas, 8th edn. Available from <https://diabetesatlas.org/atlas/eighth-edition/>, accessed 22 March 2018
 21. Dennis JM, Shields BM, Henley WE, Jones AG, Hattersley AT (2019) Disease progression and treatment response in data-driven subgroups of type 2 diabetes compared with models based on simple clinical features: an analysis using clinical trial data. *Lancet Diabetes Endocrinol* 7(6):442–451. [https://doi.org/10.1016/S2213-8587\(19\)30087-7](https://doi.org/10.1016/S2213-8587(19)30087-7)
 22. Kahkoska AR, Geybels MS, Klein KR et al (2020) Validation of distinct type 2 diabetes clusters and their association with diabetes complications in the DEVOTE, LEADER and SUSTAIN-6 cardiovascular outcomes trials. *Diabetes Obes Metab* 22(9):1537–1547. <https://doi.org/10.1111/dom.14063>
 23. Pigeyre M, Hess S, Gomez MF et al (2022) Validation of the classification for type 2 diabetes into five subgroups: a report from the ORIGIN trial. *Diabetologia* 65(1):206–215. <https://doi.org/10.1007/s00125-021-05567-4>
 24. Pek SL, Tavintharan S, Wang X et al (2015) Elevation of a novel angiogenic factor, leucine-rich- α 2-glycoprotein (LRG1), is associated with arterial stiffness, endothelial dysfunction, and peripheral arterial disease in patients with type 2 diabetes. *J Clin Endocrinol Metab* 100(4):1586–1593. <https://doi.org/10.1210/jc.2014-3855>
 25. Liu JJ, Liu S, Wang J et al (2022) Risk of Incident Heart Failure in Individuals With Early-Onset Type 2 Diabetes. *J Clin Endocrinol Metab* 107(1):e178–e187. <https://doi.org/10.1210/clinem/dgab620>
 26. Levey AS, Stevens LA, Schmid CH et al (2009) A new equation to estimate glomerular filtration rate. *Ann Intern Med* 150(9):604–612. <https://doi.org/10.7326/0003-4819-150-9-200905050-00006>
 27. Henning C (2007) Cluster-wise assessment of cluster stability. *Computational Statistics & Data Analysis* 52(1):258–271
 28. Lim SC, Dorajoo R, Zhang X et al (2017) Genetic variants in the receptor for advanced glycation end products (RAGE) gene were associated with circulating soluble RAGE level but not with renal function among Asians with type 2 diabetes: a genome-wide association study. *Nephrol Dial Transplant* 32(10):1697–1704. <https://doi.org/10.1093/ndt/gfw263>
 29. 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>
 30. Harrison JW, Tallapragada DSP, Baptist A et al (2020) Type 1 diabetes genetic risk score is discriminative of diabetes in non-Europeans: evidence from a study in India. *Sci Rep* 10(1):9450. <https://doi.org/10.1038/s41598-020-65317-1>
 31. Prasad RB, Asplund O, Shukla SR et al (2022) Subgroups of patients with young-onset type 2 diabetes in India reveal insulin deficiency as a major driver. *Diabetologia* 65(1):65–78. <https://doi.org/10.1007/s00125-021-05543-y>
 32. Liu JJ, Liu S, Saulnier PJ et al (2020) Association of Urine Haptoglobin With Risk of All-Cause and Cause-Specific Mortality in Individuals With Type 2 Diabetes: A Transethnic Collaborative Work. *Diabetes Care* 43(3):625–633. <https://doi.org/10.2337/dc19-1295>
 33. McLaughlin T, Abbasi F, Cheal K, Chu J, Lamendola C, Reaven G (2003) Use of metabolic markers to identify overweight individuals who are insulin resistant. *Ann Intern Med* 139(10):802–809. <https://doi.org/10.7326/0003-4819-139-10-200311180-00007>
 34. Liu JJ, Pek SLT, Liu S et al (2021) Association of Plasma Leucine-Rich α -2 Glycoprotein 1 (LRG1) with All-Cause and Cause-Specific Mortality in Individuals With Type 2 Diabetes. *Clin Chem* 67(12):1640–1649. <https://doi.org/10.1093/clinchem/hvab172>
 35. Liu JJ, Pek SLT, Wang J et al (2021) Association of Plasma Leucine-Rich α -2 Glycoprotein 1, a Modulator of Transforming Growth Factor- β Signaling Pathway, With Incident Heart Failure in Individuals With Type 2 Diabetes. *Diabetes Care* 44(2):571–577. <https://doi.org/10.2337/dc20-2065>
 36. Levey AS, Gansevoort RT, Coresh J et al (2020) Change in Albuminuria and GFR as End Points for Clinical Trials in Early Stages of CKD: A Scientific Workshop Sponsored by the National Kidney Foundation in Collaboration With the US Food and Drug Administration and European Medicines Agency. *Am J Kidney Dis* 75(1):84–104. <https://doi.org/10.1053/j.ajkd.2019.06.009>
 37. Bancks MP, Bertoni AG, Carnethon M et al (2021) Association of Diabetes Subgroups With Race/Ethnicity, Risk Factor Burden and Complications: The MASALA and MESA Studies. *J Clin Endocrinol Metab* 106(5):e2106–e2115. <https://doi.org/10.1210/clinem/dgaa962>
 38. Alshehry ZH, Mundra PA, Barlow CK et al (2016) Plasma Lipidomic Profiles Improve on Traditional Risk Factors for the Prediction of Cardiovascular Events in Type 2 Diabetes Mellitus. *Circulation* 134(21):1637–1650. <https://doi.org/10.1161/CIRCULATIONAHA.116.023233>
 39. Chang W, Hatch GM, Wang Y, Yu F, Wang M (2019) The relationship between phospholipids and insulin resistance: From clinical to experimental studies. *J Cell Mol Med* 23(2):702–710. <https://doi.org/10.1111/jcmm.13984>
 40. Green CD, Maceyka M, Cowart LA, Spiegel S (2021) Sphingolipids in metabolic disease: The good, the bad, and the unknown. *Cell Metab* 33(7):1293–1306. <https://doi.org/10.1016/j.cmet.2021.06.006>
 41. Tuttle KR, Brosius FC 3rd, Cavender MA et al (2021) SGLT2 Inhibition for CKD and Cardiovascular Disease in Type 2 Diabetes: Report of a Scientific Workshop Sponsored by the National Kidney Foundation. *Am J Kidney Dis* 77(1):94–109. <https://doi.org/10.1053/j.ajkd.2020.08.003>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.