diabetesresearchandclinicalpractice190(2022)110012

Contents lists available at ScienceDirect

Diabetes Research and Clinical Practice

journal homepage: www.journals.elsevier.com/diabetes-research-and-clinical-practice

Distribution and characteristics of newly-defined subgroups of type 2 diabetes in randomised clinical trials: *Post hoc* cluster assignment analysis of over 12,000 study participants

Wolfgang Landgraf[a,*], Gregory Bigot[b], Sibylle Hess[a], Olof Asplund[c], Leif Groop[d], Emma Ahlqvist[c], Annemari Käräjämäki[e], David R. Owens[f], Brian M. Frier[g], Geremia B. Bolli[h]

[a] *Sanofi, Frankfurt, Germany*
[b] *IVIDATA Life Sciences, Levallois-Perret, France*
[c] *Lund University Diabetes Centre, Department of Clinical Sciences, Skåne University Hospital, Malmö, Sweden*
[d] *Institute of Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland*
[e] *Department of Primary Health Care, Vaasa Central Hospital, and Diabetes Center, Vaasa Health Care Center, Vaasa, Finland*
[f] *Swansea University, Diabetes Research Group Cymru, College of Medicine, Swansea, UK*
[g] *The Queen's Medical Research Institute, University of Edinburgh, Edinburgh, UK*
[h] *University of Perugia School of Medicine, Department of Medicine, Section of Endocrinology and Metabolism, Perugia, Italy*

ARTICLE INFO

Keywords:
Cluster
C-peptide
Type 2 diabetes
Randomised clinical trial
Real-world studies

ABSTRACT

Aims: Newly-defined subgroups of type 2 diabetes mellitus (T2DM) have been reported from real-world cohorts but not in detail from randomised clinical trials (RCTs).
Methods: T2DM participants, uncontrolled on different pre-study therapies (n = 12.738; 82 % Caucasian; 44 % with diabetes duration > 10 years) from 14 RCTs, were assigned to new subgroups according to age at onset of diabetes, HbA1c, BMI, and fasting C-peptide using the nearest centroid approach. Subgroup distribution, characteristics and influencing factors were analysed.
Results: In both, pooled and single RCTs, "*mild-obesity related diabetes*" predominated (45 %) with mean BMI of 35 kg/m^2. "*Severe insulin-resistant diabetes*" was found least often (4.6 %) and prevalence of "*mild age-related diabetes*" (23.9 %) was mainly influenced by age at onset of diabetes and age cut-offs. Subgroup characteristics were widely comparable to those from real-world cohorts, but all subgroups showed higher frequencies of diabetes-related complications which were associated with longer diabetes duration. A high proportion of "*severe insulin-deficient diabetes*" (25.4 %) was identified with poor pre-study glycaemic control.
Conclusions: Classification of RCT participants into newly-defined diabetes subgroups revealed the existence of a heterogeneous population of T2DM. For future RCTs, subgroup-based randomisation of T2DM will better define the target population and relevance of the outcomes by avoiding clinical heterogeneity.

1. Introduction

Diabetes mellitus is a chronic and heterogeneous disease affecting 537 million people worldwide in 2021, most of them presenting with type 2 diabetes mellitus (T2DM) [1,2]. As T2DM varies considerably in clinical presentation, disease progression, and development of complications [3–6] approaches have been made to improve the subclassification of disorders within the blanket term of T2DM by applying k-means clustering of simple clinical variables [7–16]. Four new subgroups of T2DM have been derived from real-world cohorts which differ significantly in age at onset of diabetes, HbA1c, residual β-cell function, presence of obesity or insulin resistance, risk of developing diabetes-related complications and the need for insulin. The subgroups are categorised as follows: "*severe insulin-deficient diabetes (SIDD)*", "*severe insulin-resistant diabetes (SIRD)*", "*mild obesity-related diabetes (MOD)*", and "*mild age-related diabetes (MARD)*" [7]. A fifth subgroup with "*severe autoimmune diabetes (SAID)*" has been added characterised by the presence of glutamic acid decarboxylase-65 autoantibodies (GADA) and comprised of people with type 1 diabetes mellitus (T1DM) and latent autoimmune diabetes in adults (LADA).

* Corresponding author at: Sanofi, General Medicines, Medical Department Diabetes Franchise, c/o Oskar Helene Park 33, 14195 Berlin, Germany.
E-mail address: wolfgang.landgraf@sanofi.com (W. Landgraf).

https://doi.org/10.1016/j.diabres.2022.110012
Received 9 June 2022; Received in revised form 14 July 2022; Accepted 15 July 2022
Available online 18 July 2022

Traditionally the efficacy and safety of glucose-lowering drugs, including oral medications, insulins, and glucagon-like peptide-1 receptor agonists (GLP-1 RA) have been tested in randomised clinical trials (RCTs) involving both T1DM and T2DM. In phase 3/4 T2DM studies, in particular, recruitment of participants relied on broad clinical characteristics for inclusion/exclusion criteria and regimens to be tested, rather than considering disease-related markers such as degree of insulin deficiency/resistance, age at onset diabetes or diabetes duration. Therefore, conducting a post hoc classification of study participants into the newly-defined T2DM subgroups should reveal the level of patient heterogeneity and allow re-analyses of subgroup-specific responses to treatments tested in RCTs. A few cardiovascular outcome trials in T2DM have now assessed MACE and other outcomes according to the new T2DM subgroups [17–20]. However, underlying methodology and variables applied for subgroup assignments were not consistent across these RCTs or between real-world cohorts making it difficult to compare outcomes related to subgroup characteristics [21]. Therefore, the aim of the present analysis was to examine the distribution of the proposed newly-defined T2DM subgroups within a broad set of RCTs by applying the methodology of Ahlqvist et al [7]. The chosen studies involved participants diagnosed conventionally as having T2DM with either short (<5 years) or long-standing (>10 years) duration of disease, inadequately controlled on different pre-study treatments (e.g., oral glucose lowering drugs, basal insulin, or GLP-1 RAs). Specifically, we were interested in how clinical factors such as age at onset of T2DM, BMI, C-peptide levels, and diabetes duration, influenced subgroup distributions and characteristics. In addition, RCT T2DM subgroups were compared with those originally identified in the Swedish All New Diabetics in Scania (ANDIS) and the Diabetes Registry Vaasa (DIREVA) real-world cohorts [7] involving people with new-onset diabetes (ANDIS) and long-standing diabetes (DIREVA).

2. Methods

2.1. RCT populations

We used participant-level data from a large clinical database composed of T2DM RCTs that had been conducted by Sanofi or predecessor companies between 2000 and 2019 to which full access was granted for screening of baseline cluster variables. A total of 14 T2DM RCTs [22–33] were identified of which 12,738 study participants had complete data available for assignment to subgroups (**Fig. S1**, **Supplementary data**). Original enrolment of these individuals was based solely on clinical parameters and inclusion criteria across studies (**Table S1**). All baseline participant-level data including demographics, clinical and laboratory parameters, vital signs, and medical history of diabetes-related complications (retinopathy, neuropathy, nephropathy) were collected from each study and used for characterisation of diabetes subgroups. Of note, macrovascular co-morbidities (e.g., myocardial infarction, stroke) were not assessed at baseline as people with those events were excluded in 13 out of 14 RCTs. Chronic kidney disease (CKD) was defined by the diagnosis of impaired renal function (eGFR < 60 ml/min/m^2). Missing baseline values (others than the cluster variables) were not replaced except for eGFR which was calculated from serum creatinine levels by using the CKD-*Epi* formula in those participants where baseline eGFR data were absent [34].

2.2. Assignment of study participants to diabetes subgroups

Classification of T2DM individuals into the newly-defined diabetes subgroups was performed as reported by Ahlqvist et al [7]. Because GADA status was not available in most study participants a two-step approach was chosen to identify T1DM or LADA participants who constituted the SAID subgroup. In a first step, RCT participants who had GADA positive measurements but were erroneously included in these T2DM RCTs, were assigned to SAID (n = 64). In a second step, individuals with no GADA measurements but having both FCP < 0.25 nmol/L and no previous use of sulfonylurea at baseline (n = 440) were deemed as having either T1DM (undiagnosed or undetected) or LADA and therefore were assigned to SAID subgroup (**Fig. S1**). All remaining study participants (n = 12.234) were classified as having T2DM; a sex-specific nearest centroid approach was used, adapted for the use of FCP alone, rather than calculating homeostasis model assessment (HOMA2) estimates of β-cell function (HOMA2-B) and insulin resistance (HOMA2-IR) [7] for the assignation of participants into the four T2DM subgroups SIDD, SIRD, MOD, and MARD. Using FCP instead of HOMA2 parameters has been proven to identify identical subgroups as validated in the ANDIS cohort and shown in the ORIGIN trial population [17]. Subgroup variables, including age at onset of diabetes, HbA1c, BMI and FCP levels for each participant were scaled and centred. Participants were then assigned to one of the four subgroups (clusters) using the smallest Euclidean distance to cluster centroids, derived from ANDIS coordinates. Distribution of values of the four subgroup (cluster) variables across the single and pooled T2DM RCTs are represented in **Fig. S2 and Fig. S3**. Notably, FCP values were not normally distributed because some extreme outliers at high values were observed in a few RCTs [22,24,25]. The distributions of diabetes subgroups were determined for each T2DM RCT separately. For investigating the impact of various diabetes durations (<5, ≥10 years), different pre-study medications (OADs, insulin ± OAD, GLP-1RA ± OAD), and race/ethnicity on subgroup distribution and characteristics, the single subgroups from all RCTs were pooled and used for comparisons. Pooled subgroups from RCTs were also used for comparison with those from the real-world ANDIS and DIREVA cohorts [7].

2.3. Statistical analysis

Characterisation of diabetes subgroups and comparisons between subgroups are shown descriptively. Data are presented as the mean (SD), median (range), or proportion (%). Mapping and pooling of databases were performed using SAS software (version 9.4; SAS Institute, Cary, NC, USA).

3. Results

3.1. Characteristics of study participants before classification

Demographics, clinical inclusion criteria, medical history, and pre-study medications across eligible RCTs are shown in **Table S1 and Table S2**. Overall, 68 % (n = 8.702) of participants were insulin-naïve and 3.9 % (n = 501) were GLP-1 RA pre-treated. Metformin was allowed as a pre-study medication in all RCTs and 81 % received this first-line medication prior to enrolment. Sulfonylureas and thiazolidinediones (TZDs) were allowed in at least 11 RCTs and 55 % and 8 % of participants received these OADs as pre-study medications, respectively. Around 36 % of all participants had a history of at least one diabetes-related complication at baseline. Study participants had a median (range) age of 60 (19–93) years with a median diabetes duration of 8.5 (0–58) years. However, duration of diabetes varied considerably from short-term (<5 years, 27 %), to long-standing (≥10 years; 44 %). Mean (SD) HbA1c was 8.23 (1.23) % [66.5 mmol/mol], BMI 31.2 (5.9) kg/m^2, and FCP 0.86 (0.61) nmol/L (**Table S2**). Most study participants were of Caucasian (white) origin (n = 10.394; 81.6 %), a small proportion were Asian/Oriental (n = 869, 6.8 %) or Black/Afro-Americans (n = 602, 4.7 %). A total of n = 2.218 (17.4 %) participants were of Hispanic/Latino ethnicity.

3.2. Distribution of study participants into the newly-defined diabetes subgroups

A small proportion of RCT participants (4.0 %; n = 504) was assigned to SAID, most of them (3.3 %; n = 415) originated from the large ELIXA

trial [33] in which GADA status was not available and assignment was made as described under Methods. Distribution of the remaining RCT participants (n = 12,234) into the T2DM subgroups for each RCT is shown in Fig. 1. Across all RCTs, most study participants were assigned to the MOD subgroup (n = 5.734; 45.0 %), followed by MARD (n = 3.040; 23.9 %) and SIDD (n = 2.877; 22.5 %), whereas SIRD (n = 583; 4.6 %) represented the smallest group of participants. Distribution of the four subgroup-determining variables across the pooled T2DM RCTs showed that baseline mean (SD) BMI was markedly greater in SIRD (32.3 [4.9] kg/m^2) and MOD (35.0 [5.6] kg/m^2) whereas SIDD (27.5 [3.6] kg/m^2) and MARD (27.5 [3.4] kg/m^2) had the lowest BMI (Fig. 2, **Table S2**). Mean age at onset of diabetes was higher in the MARD and SIRD subgroups (57–59 years) than in the other subgroups (45–48 years). Highest mean FCP levels were found in the SIRD subgroup (2.51 nmol/L) compared to other T2DM subgroups (0.72–0.90 nmol/L), and lowest FCP levels were characteristic for SAID (0.16 nmol/L). Glycaemic control (mean HbA1c at baseline) was poorest in the SIDD subgroup (9.5 %; 80 mmol/mol) while the lowest mean HbA1c of 7.3 % (56 mmol/mol) was observed in MARD. No relevant differences in blood chemistry, electrolytes, and liver enzymes were observed between diabetes subgroups (**Table S2**).

The distribution of diabetes subgroups by race/ethnicity is illustrated in **Fig. S4**. The small Asian/Oriental, Black/Afro-American and Hispanic/Latino subpopulations differed in diabetes subgroup distributions compared to the largest subpopulation of Caucasians who dominated the T2DM RCT pool. SIDD and MARD subgroups were largest in Asian/Oriental (37/35 %), whereas MOD was the largest subgroup in Black/Afro-Americans (58 %) and Hispanic/Latino people (38 %).

3.3. Diabetes duration across RCT diabetes subgroups

Study participants showed a wide range of median diabetes duration from 6 to 15 years across RCTs. The SAID subgroup had the longest mean diabetes duration overall, followed by MOD and SIDD subgroups, and the shortest mean diabetes duration was observed in SIRD and MARD at randomisation (Fig. 3**A**). High proportions of long-standing diabetes (>10 years) were found in SAID, SIDD and MOD subgroups, whereas SIRD and MARD subgroups were characterised by individuals with rather short-term diabetes (<5 years) (Fig. 3**B**).

The prevalence of diabetes subgroups changed when study participants were analysed by categories of diabetes duration (<5 vs ≥ 10 years). Long-standing diabetes increased the prevalence of SAID (6.3 vs 2.1 %), SIDD (25.6 vs 16.6 %) and MOD (49.9 vs 37.9 %) whereas short-term duration of diabetes increased the prevalence of SIRD (6.6 vs 2.5 %) and MARD (36.8 vs 15.7 %) more than twice (Fig. 5).

3.4. Diabetes-related complications in RCT diabetes subgroups

Diabetes-related complications including retinopathy, neuropathy, nephropathy, and CKD (eGFR < 60 ml/min/1.73 m^3) have been documented as part of the medical history and prevalence at randomisation across RCTs is presented in **Table S1**. The highest prevalence of complications was found in the EDITION-1 and EDITION-2 studies [33,34] which had participants with the longest mean duration of diabetes compared to other studies. In those studies in which participants had similar diabetes duration, the prevalence of complications was comparable (**Table S1**). After classification in diabetes subgroups, retinopathy was more prevalent in SAID, SIDD and MOD compared to SIRD and MARD across all study participants and was found to be 2 to 5-fold higher in those with long-standing compared to short-term diabetes (Fig. 4, **Table S2**). Diabetes-related neuropathy was most prevalent (24 %) in the SIRD subgroup with short-term diabetes compared to other diabetes subtypes (10–16 %) and was equal in all subgroups with long-standing diabetes (26–35 %). CKD was most prevalent in SAID, SIRD and MARD subgroups, both in participants with short-term and long-standing diabetes.

Diabetes-related nephropathy was slightly more frequently diagnosed in those participants with short-term diabetes and assigned to SIRD (10 %) compared to other subtypes (4–7 %) but was found most prevalent in SAID (23 %), MARD (17 %) and MOD (16 %) subgroups with long-standing diabetes compared to SIDD and SIRD (each 12 %) (data not shown).

3.5. Pre-study medications in RCT diabetes subgroups

Subgroup distributions by pre-study medications (OADs, insulins,

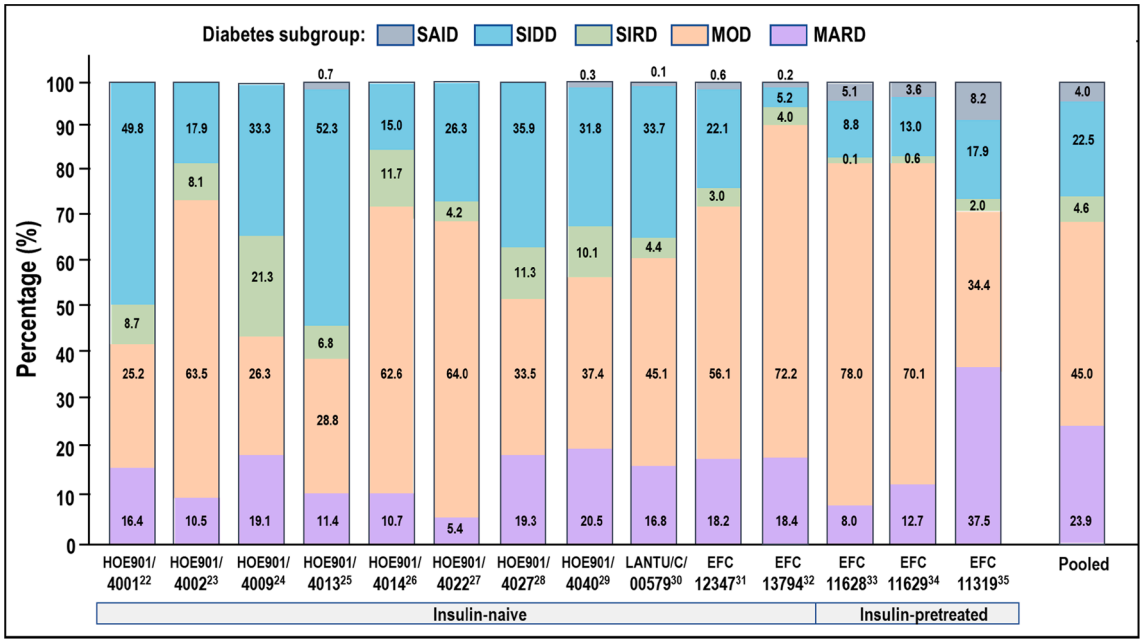

Fig. 1. Participant distributions into newly-defined diabetes subgroups (clusters) across fourteen eligible T2DM RCTs Diabetes subgroups were determined based on age at onset of diabetes, HbA1c, BMI and fasting C-peptide (see under Methods). SAID = severe autoimmune diabetes; SIDD = severe insulin-deficient diabetes; SIRD = severe insulin-resistant diabetes; MOD = mild obesity-related diabetes; MARD = mild age-related diabetes.

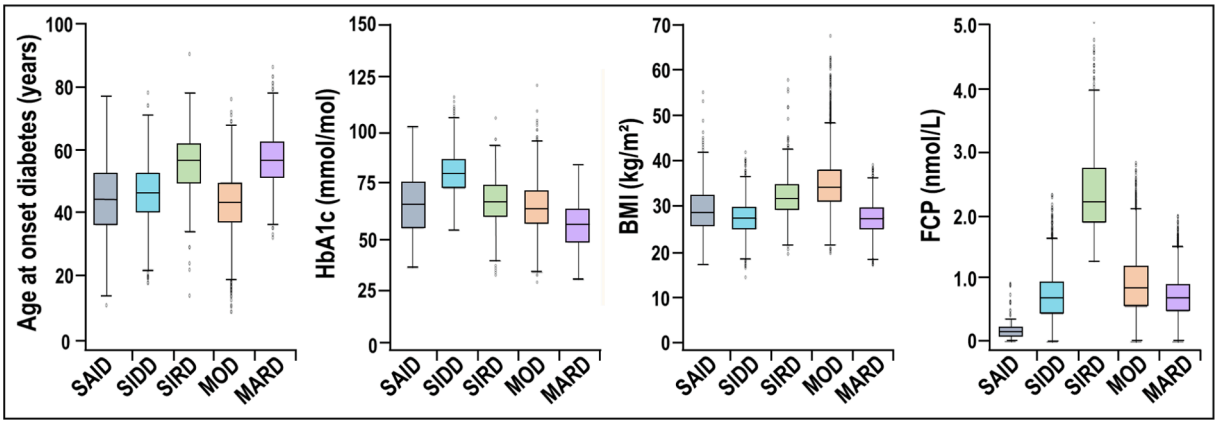

Fig. 2. Distributions of age at diagnosis, HbA1c, body mass index, and fasting C-peptide across replicated pooled new diabetes subgroups from fourteen T2DM RCTs (n = 12,738). SAID = severe autoimmune diabetes; SIDD = severe insulin-deficient diabetes; SIRD = severe insulin-resistant diabetes; MOD = mild obesity-related diabetes; MARD = mild age-related diabetes; FCP, fasting C-peptide; BMI, body mass index; boxes are the median, and 25th and 75th percentiles; whiskers are the 1st and 99th percentiles. Values outside these percentiles are represented by open circles.

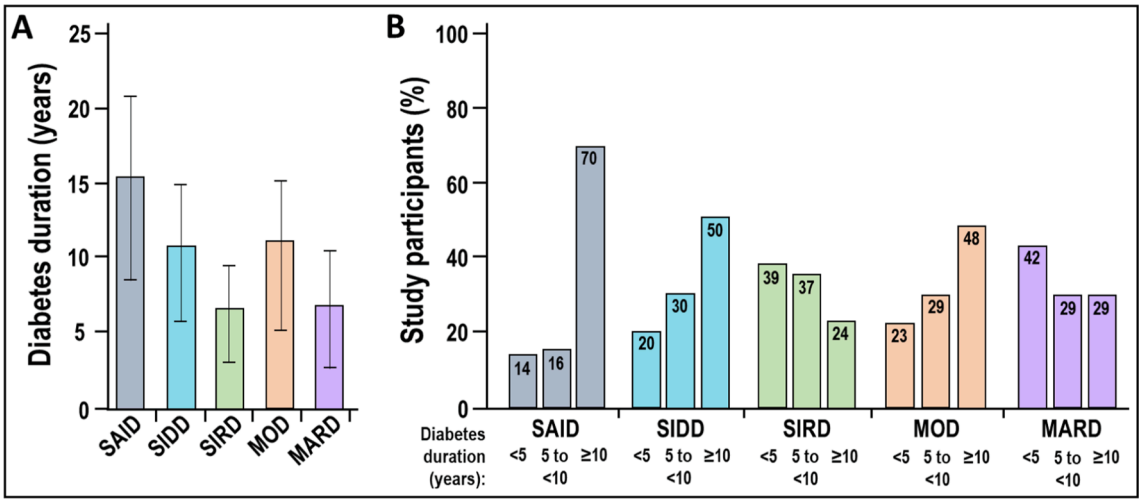

Fig. 3. Mean diabetes duration of pooled RCT participants at baseline according to newly-defined diabetes subgroups (A) and distribution of different categories of diabetes duration across diabetes subgroups (B). SAID = severe autoimmune diabetes; SIDD = severe insulin-deficient diabetes; SIRD = severe insulin-resistant diabetes; MOD = mild obesity-related diabetes; MARD = mild age-related diabetes. Whiskers represent 25th and 75th percentiles.

GLP-1 RA) are shown in Fig. S4. MOD was found to be the dominating subgroup regardless of pre-study treatments. Interestingly, in OAD-pretreated (insulin-naïve) participants 25.8 % contributed to the SIDD subgroup.

Metformin, sulfonylureas, and TZDs were permitted as pre-study glucose-lowering medications in at least 11 or more RCTs (Table S1). First-line recommended metformin was widely used across all T2DM subgroups (73–85 %) (Fig. S5). Sulfonylurea was more often prescribed in the SIRD (82 %) and SIDD (73 %) subgroups compared to MARD (54 %) and MOD (49 %). Among the few pre-study prescriptions of TZDs, the SIDD and MOD subgroups showed about twice the frequency (8–9 %) than SIRD (5 %) and MARD (5 %).

Lipid-lowering therapy was allowed in all T2DM individuals enrolled in the RCTs. After classification of RCT participants into diabetes subgroups only about 50 % or less of T2DM individuals in SIDD and SIRD subgroups received lipid-lowering treatment, and the SIRD subtype showed the worst lipid profile and highest TG/HDL-C ratio compared to other subgroups (Fig. S6, Table S2).

3.6. RCT diabetes subgroups compared to real-world subgroups

Compared to the Scandinavian ANDIS and DIREVA real-world cohorts, distributions of T2DM subgroups in RCTs differed. The most striking differences were found in the prevalences of MOD, SIRD and MARD subgroups between cohorts. In RCTs, MOD was the greatest and SIRD the smallest subgroup, whereas in ANDIS/DIREVA MARD was the predominant subgroup and SIRD was also more prevalent in the real-world studies irrespective of diabetes duration (Fig. 1, Fig. 5).

Key characteristics of T2DM subgroups from pooled RCTs and ANDIS/DIREVA real-world cohorts are summarised in Table S3. The most conformable subgroup characteristic was mean BMI irrespective of diabetes duration, whereas the most striking difference was found in the prevalence of diabetes-related complications between RCT and real-world (ANDIS) participants. Differences in age at onset, FCP and HbA1c were also observed between subgroups from RCTs and real-world cohorts.

4. Discussion

While the new diabetes subgroups have been described in Scandinavian real-world cohorts, their contribution in randomised clinical trials has not been explored. Therefore, data of 12,738 individuals with T2DM who had participated in 14 RCTs between 2000 and 2019 were analysed to determine the distribution of these proposed new diabetes

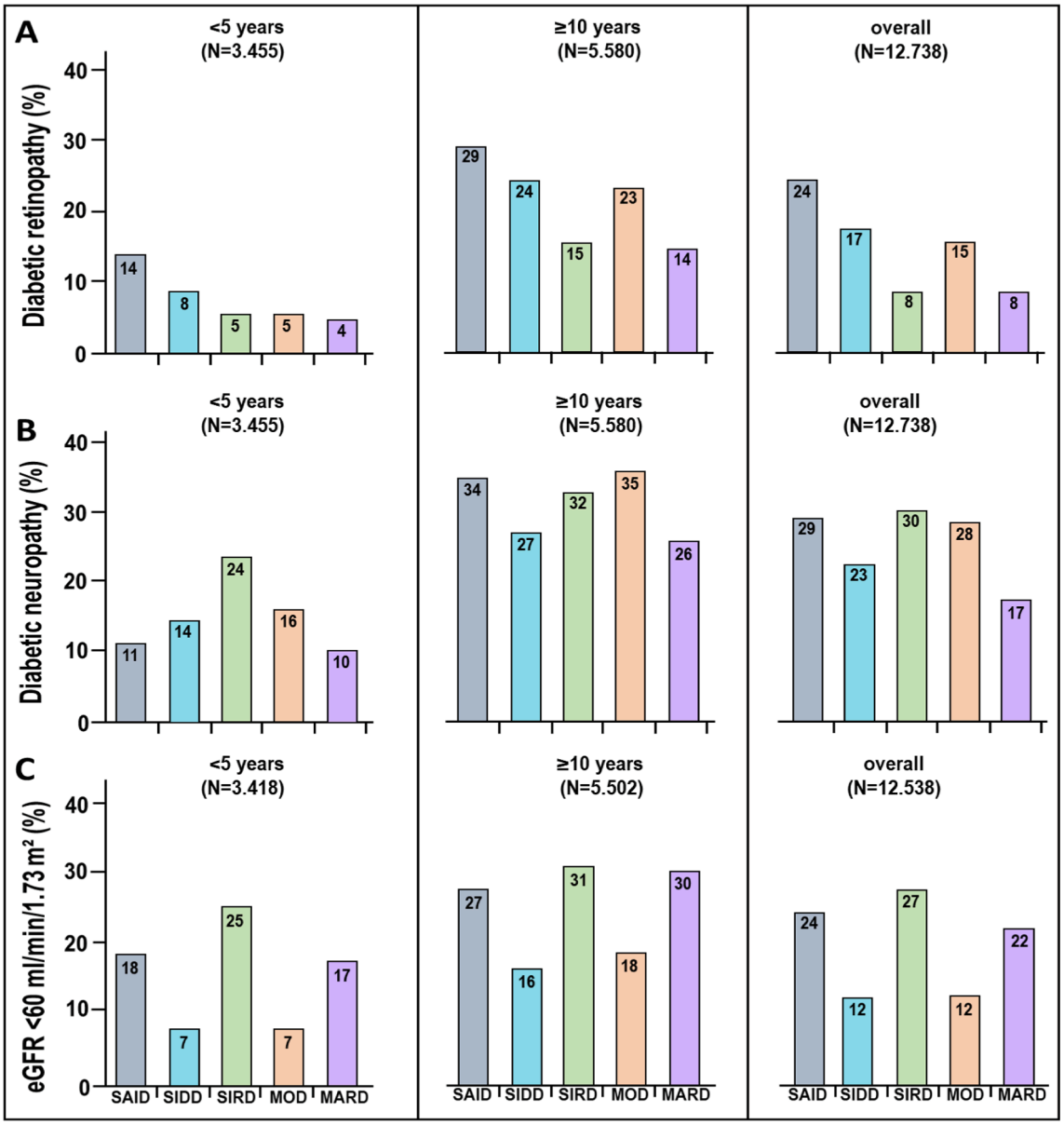

Fig. 4. Prevalence of (A) diabetes-related retinopathy, (B) neuropathy, and (C) chronic kidney disease (eGFR < 60 ml/min/1.73 m^2) in pooled diabetes subgroups from T2DM RCTs (overall and stratified by diabetes duration at baseline). SAID = severe autoimmune diabetes; SIDD = severe insulin-deficient diabetes; SIRD = severe insulin-resistant diabetes; MOD = mild obesity-related diabetes; MARD = mild age-related diabetes.

subgroups. In the present cluster assignment analysis that was based on four clinical variables (HbA1c, BMI, FCP, age at onset) all five newly-defined subgroups have been replicated across the RCTs. Most T2DM RCT participants were assigned to the MOD and the least to the SIRD subgroup across the studies. In addition, a smaller number of participants with confirmed or potential T1DM/LADA (4.0 %) overall were identified. It is in part likely that the SAID subgroup in T2DM RCTs was included because GADA had not been measured at screening. Large cardiovascular outcome trials without GADA measuring have been a source for enrolment of SAID patients as demonstrated by the findings from the ELIXA [35] and ORIGIN trials [17,37] where 8.2 % and 3.4 % of participants respectively were assigned retrospectively to the SAID subgroup. Therefore, in order to avoid individuals with T1DM/LADA to be recruited in T2DM trials it would be desirable to include mandatory measurements of GADA when screening potential participants to avoid diagnostic failures.

Unlike real-world registries, which seldom impose the strict inclusion criteria of RCTs, it is not unexpected that distributions of T2DM subgroups in RCTs differ from real-world populations and mainly depend on thresholds applied to inclusion criteria. For example, the lower prevalence of the SIRD and MARD subgroups observed in the present group of RCTs, both characterised by a higher mean age of onset diabetes, could be explained by the age cut-off (<75 years) applied in most, resulting in a younger population with a lower mean age of onset of diabetes and longer diabetes duration. However, another explanation for the lower prevalence of SIRD and MARD subgroups may exist that older people are enrolled less frequently into RCTs, as supported by the observations from studies with no age cut-off [31–35] where the mean age of the participants was comparable to studies which had an age cut-off. The SIRD subgroup appears to increase only through the number of participants with high FCP levels, and the MARD subgroup with higher mean age of onset of diabetes (or shorter diabetes duration). In contrast, the predominant MOD and SIDD subgroups in our RCTs appeared to be strongly determined by diabetes duration and the presence of overweight/obesity, and to a certain degree by (high) baseline HbA1c levels of participants. Our data suggest that the proportion of MOD is lower in

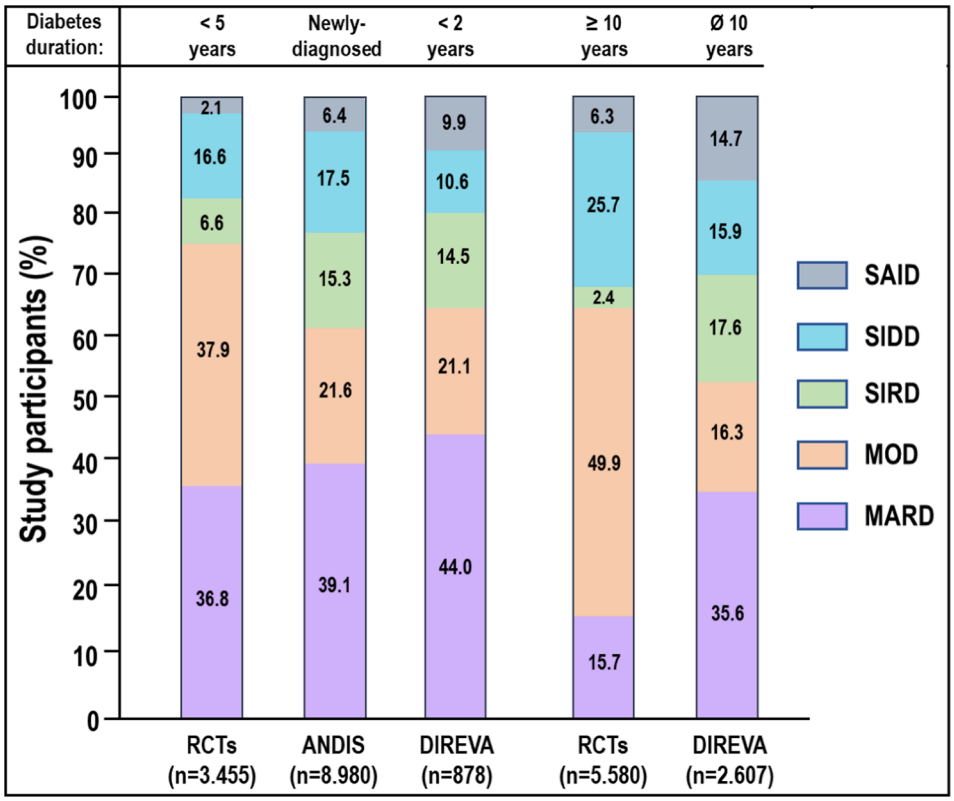

Fig. 5. Distribution of newly-defined diabetes subgroups in pooled RCT participants and in the Scandinavian ANDIS/DIREVA real-world cohorts stratified by diabetes duration categories at baseline or sampling. AID = severe autoimmune diabetes; SIDD = severe insulin-deficient diabetes; SIRD = severe insulin-resistant diabetes; MOD = mild obesity-related diabetes; MARD = mild age-related diabetes; ANDIS, All New Diabetics in Scania; DIREVA, Diabetes Registry in Vaasa.

participants with shorter diabetes duration but remains higher than those generally found in real world populations [7–12,15], presumably as a result of preferred inclusion criteria that were widely used across the selected RCTs or by an unknown selection bias applied to RCTs. In addition, a higher prevalence of MOD, in parallel to a lower one of SIDD, was observed at higher mean BMI whereas higher mean HbA1c suggested a reduction of MOD and an increase of SIDD in RCTs.

Interestingly, in two studies [31,32] involving insulin-pretreated T2DM individuals with long-standing diabetes (mean duration 13–16 years) and low average FCP levels, pretreated with either a basal-bolus or high dose of basal insulin regimen, the proportion of the SIDD subgroup was surprisingly low in both studies compared to other RCTs. The reason for this observation is unclear, but high average BMI values > 34 kg/m^2 in these T2DM individuals have probably assigned many of them to the MOD rather than to the SIDD subgroup despite low FCP levels. Similar high proportions (≥63 %) of the MOD subgroup have been observed in a few more of our RCTs [23,26,27,32] with high average BMIs of > 32 kg/m^2. In contrast, in studies where average BMI was lower, the proportion of SIDD was always higher demonstrating an inverse relationship between these two subgroup variables. These findings from the present group of RCTs raise the question as to whether k-means clustering coordinates and centroids derived from ANDIS individuals with new-onset diabetes will need some adjustment when applied in individuals with long-standing diabetes.

Another important finding from the present analysis is that diabetes duration of RCT participants was a significant modulating factor for the distributions of diabetes subgroups and the prevalence of diabetes-related comorbidities. "*Severe autoimmune diabetes*" (SAID, 6 % vs 2 %) and "*severe insulin-deficient diabetes*" (SIDD, 26 % vs 17 %) subgroups were markedly more prevalent in people with long-term (≥10 years) as compared to those with shorter-term (<5 years) diabetes. Similar findings were reported from a real-world cohort involving more than 2.000 individuals with established diabetes where SAID (11 % vs 6 %) and SIDD (25 % vs 14 %) subgroups were twice as prevalent in patients with long-term (mean: 11 years) as compared to those with new-onset diabetes [17], and also found for the SIDD subgroup (16 % vs 11 %) in the Finnish DIREVA real-world cohort. The prevalence of diabetes-related complications in all five diabetes subgroups was consistently higher in study participants with long-standing diabetes compared to those with shorter-term duration of diabetes. Not surprisingly, study participants with long-standing diabetes had a higher prevalence of all types of diabetes-related complications than has been reported in real-world cohorts composed of people who were newly-diagnosed [7,15].

The replicated RCT T2DM subgroups largely differed in their mean residual beta cell function as expressed by FCP levels, with SIDD and MARD subgroups characterised by severe to moderate insulin deficiency but obviously also containing individuals with a preserved beta cell function. A similar distribution of FCP levels across the diabetes subgroups was reported in the ORIGIN trial and these results also showed that the SIDD subgroup with the most advanced insulin deficiency derived the greatest benefit from the use of insulin (glargine) instead of standard-of-care therapy [17]. Moreover, previous findings from RCTs that enrolled only insulin-naïve T2DM RCT participants who were stratified by different FCP levels further suggest that T2DM subgroups may not only differ substantially in the need for timely insulin therapy (SIDD) and higher insulin dose requirements (SIRD), but also in hypoglycaemia risk when being treated with basal insulin [36].

Interestingly, in the analysed RCTs that enrolled only insulin-naïve T2DM participants of whom most were pretreated with metformin and/or sulfonylurea for many years, 22–50 % have been classified to the SIDD subgroup, which had the poorest glycaemic control at baseline compared to other subgroups. These findings may imply that the SIDD subgroup represents T2DM individuals who are difficult to diagnose in routine practice and the highest HbA1c levels observed in this subgroup further suggest either a failure to appreciate the urgency for insulin replacement therapy or proper insulin titration and/or timely intensification of insulin treatment for those individuals as recommended by international guidelines [37,38].

The possibility may exist that the SIDD subgroup is barely distinguishable from the MOD or MARD subgroup at the time of diagnosis,

despite an expected difference in BMI, which is a useful differentiating factor, at least from the MOD subgroup. As individuals assigned to the MOD and MARD subgroups are sufficiently controlled with glucose-lowering drugs during the first years after diagnosis, this may also explain why the proportions of the SIDD, MOD and MARD subgroups in OAD-pretreated (insulin-naïve) and insulin-pretreated participants were similar in the present RCTs, and the SIDD subgroup, as expected, was neither more prevalent nor being the dominating subgroup in insulin-pretreated RCT participants. Moreover, high prescription rates of metformin and, to a lesser degree of sulfonylurea, in these relatively lean people in the SIDD subgroup, who had lower BMI levels compared to MOD and MARD, further supports these assumptions. The "*severe insulin-resistant* (SIRD)" subgroup with substantial prescriptions of metformin and sulfonylurea and relatively low use of TZDs, is another T2DM population which requires more effort to improve glycaemic control and lipid-lowering profiles. Awareness of the need to measure FCP levels along with glycaemic parameters and BMI as main characteristics of this T2DM subgroups will also assist clinical decision-making to implement a personalised and optimal treatment strategy for those individuals.

4.1. Study limitations

We recognise that this *post hoc* cluster assignment analysis has some limitations as full access to patient-level data was available only for this clinical database. Therefore, the findings of this analysis are restricted to study participants of the database. Moreover, as one large study accounted for 40 % of the pooled study population and some studies had no upper limits for BMI as an inclusion criterion, this heterogeneity between studies may have introduced a bias in the interpretation of the pooled subgroup distributions. Another confounding factor influencing subgroup distributions in RCTs could be an unknown selection bias induced by investigator-based decisions to include only people who are generally compliant with advised medical care, availability of time, and with a high willingness to participate in RCTs. Those people might differ from the routine clinic populations in characteristics that could have impacted subgrouping. However, as characteristics of RCT and real-world subgroups did not substantially differ out of diabetes-related complications this impact should be considered as small. The unavailability of GADA status in most study participants may have introduced some uncertainty about the robustness of definition of the SAID subgroup. Therefore, it cannot be excluded that some participants with undetected or undiagnosed T1D or LADA have been assigned to subgroups other than SAID, based on our approach of using FCP thresholds of 0.25 nmol/L and previous sulfonylurea intake as decision criteria. Moreover, a few T2DM participants with FCP < 0.25 nmol/L may have been assigned to SAID. However, as this uncertainty refers to a very small number of participants overall, the impact on distributions of participants to T2DM subgroups can be considered to be low and clinical characteristics of the assigned SAID subgroup compare very well with those reported in ANDIS. As FCP values across the 14 RCTs were assayed at different laboratories using different methodologies (ELISA, RIA, Luminex), and study participants had different pre-study treatments (sulfonylurea, insulin) which potentially could have affected baseline FCP levels, the absolute FCP values might have varied between RCTs. However, distribution of FCP values and mean values across studies (Table S1, Fig. S2) showed comparable patterns of FCP suggesting no, or only a minimal effect on the subgroup assignment. It can also be dismissed that FCP levels might have been affected by reduced renal function (eGFR < 60 ml/min/1.73 m^2) as C-peptide is mainly excreted via the kidney. Correlation of FCP levels with reduced kidney function did not indicate any relationship between those variables in the RCT population.

5. Conclusions

This cluster assignment analysis which has been applied to more than 12,000 study participants of mainly Caucasian origin from 14 RCTs has identified five diabetes subgroups that were described previously in newly-diagnosed people from real-world populations. The present analysis indicates that RCT subgroups of varying diabetes duration and who were receiving different glucose-lowering therapies prior to enrolment do not differ markedly in their clinical characteristics (HbA1c, BMI, FCP, age at diagnosis) from real-world Scandinavian subgroups (ANDIS/DIREVA). However, distributions of RCT participants to the five subgroups and prevalence of diabetes-related complications moderately change when stratified by diabetes duration of the participants. The most important clinical observation of the present analysis is that people classified into either "*severe-insulin deficient diabetes (SIDD)*" or "*severe insulin-resistant diabetes (SIRD)*" represent two T2DM subgroups in which both glycaemic and lipid levels are not adequately controlled, and the prescribed pre-study glucose-lowering therapy did not adhere to current international guidelines. These two vulnerable subgroups, which are also associated with an increased risk of developing microvascular complications, could benefit most from individualised diabetes and lipid management. The availability in the future of a "*subgroup calculator/classifier*" may facilitate the way towards precision medicine in diabetes by identifying T2DM subgroups based on four simple clinical parameters and thereby assist clinicians to facilitate treatment decisions.

Through recognition of the different diabetes subgroups, study participants for future comparative T2DM RCTs could be selected according to these newly-defined diabetes subgroups by applying prior cluster assignment, in preference to using traditional clinical parameters alone. It is also proposed that T2DM study protocols should include mandatory GADA and FCP measurements at screening, to limit both clinical heterogeneity and to exclude recruitment of people with T1DM or LADA.

CRediT authorship contribution statement

Wolfgang Landgraf: Conceptualization, Data curation, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Gregory Bigot:** Conceptualization, Formal analysis, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **Sibylle Hess:** Conceptualization, Validation, Writing – review & editing. **Olof Asplund:** Methodology, Software, Validation, Writing – review & editing. **Leif Groop:** Methodology, Writing – review & editing. **Emma Ahlqvist:** Methodology, Validation, Writing – review & editing. **Annemari Käräjämäki:** Data curation, Writing – review & editing. **David R. Owens:** Validation, Writing – original draft, Writing – review & editing. **Brian M. Frier:** Validation, Writing – original draft, Writing – review & editing. **Geremia B. Bolli:** Validation, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: W.L. and S.H. are employees of Sanofi, Germany, and Sanofi shareholders. G.B.is an IVIDATA employee, contracted to Sanofi. E.A. O.A, A. K. and L.G. declare no competing interests. D.R.O. has received honoraria for lecturing and consulting from Boehringer Ingelheim, Eli Lilly, Roche Diagnostics, Sanofi, and Takeda. B.M.F. has served on advisory panels for Eli Lilly and Zucara Therapeutics, and on the speakers' bureau for Eli Lilly, Novo Nordisk, Sanofi, and Abbott. G.B.B. is a consultant for Eli Lilly and Sanofi; has received research support from Sanofi; and is on the speakers' bureau for Eli Lilly, Menarini, and Sanofi.

Acknowledgments

This work is dedicated to the memory of Prof. Jürgen Sandow, a passionate pharmacologist and endocrinologist. The authors would like

to thank Tiinamaija Tuomi, University of Helsinki, and Mathieu Coudert, Keni Lee, and Jukka Westerbacka, all of Sanofi, for input and/or critical review of this manuscript. The DIREVA Study has been supported by the Vaasa Hospital District, State Research funding via the Turku University Hospital, Vaasa Central Hospital, Jakobstadsnejdens Heart Foundation, the Medical Foundation of Vaasa, and the Finnish Medical Foundation.

Role of the funding source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The underlying diabetes RCTs were funded by Sanofi.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.diabres.2022.110012.

References

[1] Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF diabetes Atlas: Global, regional, and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diab Res Clin Pract 2022;183:109119.
[2] American Diabetes Association. 2. Classification and diagnosis of diabetes: Standards of Medical Care in Diabetes - 2021. Diabetes Care 2021;44 (Suppl. 1): S133-52.
[3] Pajvani UB, Accili D. The new biology of diabetes. Diabetologia 2015;58(11): 2459–68.
[4] Tuomi T, Santoro N, Caprio S, Cai M, Weng J, Groop L. The many faces of diabetes: a disease with increasing heterogeneity. The Lancet 2014;383(9922):1084–94.
[5] Faerch K, Hulmán A, Solomon TPJ. Heterogeneity of pre-diabetes and type 2 diabetes: implications for prediction, prevention, and treatment responsiveness. Curr Diabetes Rev 2016;12(1):30–41.
[6] Cefalu WT, Andersen DK, Arreaza-Rubín G, Pin C, et al. Heterogeneity of Diabetes: β-Cells, Phenotypes, and Precision Medicine: Proceedings of an International Symposium of the Canadian Institutes of Health Research's Institute of Nutrition, Metabolism and Diabetes and the US National Institutes of Health's National Institute of Diabetes and Digestive and Kidney Diseases. Diabetes 2022;71(1):1-22.
[7] Ahlqvist E, Storm P, Käräjämäki A, Martinell M, Dorkhan M, Carlsson A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. Lancet Diabetes Endocrinol 2018;6(5): 361–9.
[8] Dennis JM, Shields BM, Henley WE, Jones AG, Hattersley AT. 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 2019;7(6):442–51.
[9] Zaharia OP, Strassburger K, Strom A, Bönhof GJ, Karusheva Y, Antoniou S, et al. Risk of diabetes-associated diseases in subgroups of patients with recent-onset diabetes: a 5-year follow-up study. Lancet Diabetes Endocrinol 2019;7(9):684–94.
[10] Zou X, Zhou X, Zhu Z, Ji L. Novel subgroups of patients with adult-onset diabetes in Chinese and US populations. Lancet Diabetes Endocrinol 2019;7(1):9–11.
[11] Safai N, Ali A, Rossing P, Ridderstråle M. Stratification of type 2 diabetes based on routine clinical markers. Diabetes Res Clin Pract 2018;141:275–83.
[12] Slieker RC, Donnelly LA, Fitipaldi H, Bouland GA, Giordano GN, Åkerlund M, et al. Replication and cross-validation of type 2 diabetes subtypes based on clinical variables: an IMI-RHAPSODY study. Diabetologia 2021;64(9):1982–9.
[13] Mansour Aly D, Dwivedi OP, Prasad RB, Käräjämäki A, Hjort R, Thangam M, et al. Genome-wide association analyses highlight etiological differences underlying newly defined subtypes of diabetes. Nat Genet 2021;53(11):1534–42.
[14] Prasad RB, Asplund O, Shukla SR, Wagh R, Kunte P, Bhat D, et al. Subgroups of patients with young-onset type 2 diabetes in India reveal insulin deficiency as a major driver. Diabetologia 2022;65(1):65–78.
[15] Fedotkina O, Sulaieva O, Ozgumus T, Cherviakova L, et al. Novel reclassification of adult diabetes is useful to distinguish stages of b-cell function linked to the risk of vascular complications: the DOLCE study from Northern Ukraine. Front Genet 2021;12:637945.
[16] Christensen DH, Nicolaisen SK, Ahlqvist E, Stidsen JV, Nielsen JS, Hojlund K, et al. Type 2 diabetes classification: a data-driven cluster study of the Danish Centre for Strategic Research in Type 2 Diabetes (DD2) cohort. BMJ Open Diab Res Care 2022;10(2). https://doi.org/10.1136/bmjdrc-2021-002731.
[17] Pigeyre M, Hess S, Gomez MF, Asplund O, Groop L, Paré G, et al. Validation of the classification for type 2 diabetes into five subgroups: A report from the ORIGIN trial. Diabetologia 2022;65(1):206–15.
[18] Kahkoska AR, Geybels MS, Klein KR, Kreiner FF, Marx N, Nauck MA, et al. 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 2020;22(9):1537–47.
[19] Sharma A, Zheng Y, Ezekowitz JA, Westerhout CM, Udell JA, Goodman SG, et al. Cluster analysis of cardiovascular phenotypes in patients with type 2 diabetes and established atherosclerotic cardiovascular disease: a potential approach to precision medicine. Diabetes Care 2022;45(1):204–12.
[20] Aoki Y, Hamrén B, Clegg LE, Stahre C, Bhatt DL, Raz I, et al. Assessing reproducibility and utility of clustering of patients with type 2 diabetes and established CV disease (SAVOR -TIMI 53 trial). PLoS ONE 2021;16(11).
[21] Herder C, Roden M. A novel diabetes typology: towards precision diabetology from pathogenesis to treatment. Diabetologia 2022:1–13. https://doi.org/10.1007/s00125-021-05625-x.
[22] Fritsche A, Schweitzer MA, Häring HU. Glimepiride combined with morning insulin glargine, bedtime neutral protamine hagedorn insulin, or bedtime insulin glargine in patients with type 2 diabetes: a randomized, controlled trial. Ann Intern Med 2003;138(12):952–9.
[23] Riddle MC, Rosenstock J, Gerich J. The Treat-to-Target Trial: randomized addition of glargine or human NPH insulin to oral therapy of type 2 diabetic patients. Diabetes Care 2003;26:3080–6.
[24] Standl E, Maxeiner S, Raptis S. Once−daily insulin glargine administration in the morning compared to bedtime in combination with morning glimepiride in patients with type 2 diabetes: an assessment of treatment flexibility. Horm Metab Res 2006;38(3):172–7.
[25] Eliaschewitz FG, Calvo C, Valbuena H, Ruiz M, Aschner P, Villena J, et al. Therapy in type 2 diabetes: insulin glargine vs. NPH insulin both in combination with glimepiride. Archives Med Res 2006;37(4):495–501.
[26] Rosenstock J, Sugimoto D, Strange P, Stewart JA, Soltes-Rak E, Dailey G. Triple therapy in type 2 diabetes. Insulin glargine or rosiglitazone added to combination therapy of sulfonylurea plus metformin in insulin-naïve patients. Diabetes Care 2006;29(3):554–9.
[27] Hollander P, Sugimoto D, Vlajnic A, Kilo C. Combination therapy with insulin glargine plus metformin but not insulin glargine plus sulfonylurea provides similar glycemic control to triple oral combination therapy in patients with type 2 diabetes uncontrolled with dual oral agent therapy. J Diabetes Complications 2015;29(8): 1266–71.
[28] Janka HU, Plewe G, Riddle MC, Kliebe-Frisch C, Schweitzer MA, Yki-Järvinen H. Comparison of basal insulin added to oral agents versus twice-daily premixed insulin as initial insulin therapy for type 2 diabetes. Diabetes Care 2005;28(2): 254–9.
[29] Bretzel RG, Nuber U, Landgraf W, Owens DR, Bradley C, Linn T. Once-daily basal insulin glargine versus thrice-daily prandial insulin lispro in people with type 2 diabetes on oral hypoglycaemic agents (APOLLO): an open randomised controlled trial. Lancet 2008;371(9618):1073–84.
[30] Swinnen SG, Dain M-P, Aronson R, Davies M, et al. 24-Week, randomized, treat-to-target trial comparing initiation of insulin glargine once-daily with insulin detemir twice-daily in patients with type 2 diabetes inadequately controlled on oral glucose-lowering drugs. Diabetes Care 2010;33:1176–8.
[31] Bolli GB, Riddle MC, Bergenstal RM, et al. New insulin glargine 300 U/mL compared with glargine 100 U/mL in insulin-naive people with type 2 diabetes on oral glucose-lowering drugs: a randomized controlled trial (EDITION 3). Diabetes Obes Metab 2015;17(4):386–94.
[32] Blonde L, Rosenstock J, Del Prato S, Henry R, Shehadeh N, Frias J, et al. Switching to IGlarLixi versus continuing daily or weekly GLP-1 RA in type 2 diabetes inadequately controlled by GLP-1 RA and oral antihyperglycemic therapy: The LixiLan-G randomized clinical trial. Diabetes Care 2019;42(11):2108–16.
[33] Pfeffer MA, Claggett B, Diaz R, Dickstein K, Gerstein HC, Køber LV, et al. Lixisenatide in patients with type 2 diabetes and acute coronary syndrome. New Eng J Med 2015;373(23):2247–57.
[34] Levey AS, Stevens LA, Schmid CH, Zhang Y, et al. New equation to estimate glomerular filtration rate. Ann Intern Med 2009;50:604–12.
[35] Birkeland KI, Grill V, Wium C, McQueen MJ, Lopez-Jaramillo P, Lee SF, et al. The association of basal insulin treatment versus standard care with outcomes in anti-GAD positive and negative subjects: A post-hoc analysis of the ORIGIN trial. Diabetes Obes Metab 2019;21(2):429–33.
[36] Landgraf W, Owens DR, Frier BM, Zhang M, Bolli GB. Fasting C-peptide, a biomarker for hypoglycaemia risk in insulin-naïve people with type 2 diabetes initiating basal insulin glargine 100 U/mL. Diabetes Obes Metab 2020;22(3): 315–23.
[37] Davies MJ, D'Alessio DA, Fradkin J, Kernan WN, Mathieu C, Mingrone G, et al. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetologia 2018;61(12):2461–98.
[38] Garber AJ, Abrahamson MJ, Barzilay JI, Blonde L, Bloomgarden ZT, Bush MA, et al. Consensus Statement by the American Association of Clinical Endocrinologists and American College of Endocrinology on the Comprehensive Type 2 Diabetes Management Algorithm – 2018 Executive Summary. Endocrine Pract 2018;24(1):91–121.