Diabetologia (2026) 69:2672–2684
https://doi.org/10.1007/s00125-026-06802-6

REVIEW

Type 2 diabetes subtypes for precision medicine: methodological challenges and alternative prediction-based approaches

Tim Mori[1,2] · Christian Herder[2,3,4] · Pedro Cardoso[5] · John M. Dennis[5] · Oliver Kuß[1,2,6]

Received: 5 January 2026 / Accepted: 1 June 2026 / Published online: 25 July 2026
© The Author(s) 2026

Abstract

Growing interest in precision diagnostics has stimulated efforts to classify type 2 diabetes into more granular subtypes. Recent research has focused on four pathophysiology-based subtypes: severe insulin-deficient diabetes (SIDD); severe insulin-resistant diabetes (SIRD); moderate obesity-related diabetes (MOD); and moderate age-related diabetes (MARD). While these subtypes provide an attractive framework for precision medicine, they also come with methodological challenges. This narrative review highlights three key issues: (1) uncertainty about the existence of four discrete subtypes; (2) low certainty in subtype assignment; and (3) limited evidence for prognostic and therapeutic value compared with other approaches. As an alternative to subtyping, we discuss individualised prediction models, which provide personalised treatment recommendations based on routine clinical features. These models avoid some of the limitations of discrete subtyping but they do not provide insights into a person's underlying pathophysiology. Thus, these two approaches can be viewed as complementary, each serving a different purpose in precision medicine.

Keywords Classification uncertainty · Clusters · Limitations · Methodology · Precision medicine · Prediction models · Review · Subtypes · Type 2 diabetes mellitus

Abbreviations

ALT	Alanine aminotransferase
ANDIS	All New Diabetics in Scania
c-index	Concordance index
CPRD	Clinical Practice Research Datalink
GLP1-RA	Glucagon-like peptide-1 receptor agonist
KDIGO	Kidney Disease: Improving Global Outcomes
MARD	Moderate age-related diabetes
MOD	Moderate obesity-related diabetes
NRE	Normalised relative entropy
PMDI	Precision Medicine in Diabetes Initiative
SAID	Severe autoimmune diabetes
SCORE	Systematic Coronary Risk Evaluation
SGLT2i	Sodium–glucose cotransporter 2 inhibitor
SIDD	Severe insulin-deficient diabetes
SIRD	Severe insulin-resistant diabetes

✉ Tim Mori
tim.mori@ddz.de

✉ Oliver Kuß
Oliver.Kuss@ddz.de

1 Institute for Biometrics and Epidemiology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany

2 German Center for Diabetes Research (DZD), Partner Düsseldorf, München-Neuherberg, Germany

3 Institute for Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University, Düsseldorf, Germany

4 Department of Endocrinology and Diabetology, Medical Faculty and University Hospital Düsseldorf, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

5 Institute of Clinical & Biomedical Sciences, University of Exeter Medical School, Exeter, UK

6 Centre for Health and Society, Faculty of Medicine, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

Precision medicine and type 2 diabetes subtypes

Type 2 diabetes is a heterogeneous disease and its clinical presentation differs considerably between individuals [1–3]. For example, it is well established that some individuals with type 2 diabetes are living with obesity and are insulin resistant, whereas others are rather lean and insulin deficient [4]. Despite such differences, current guidelines for the management of type 2 diabetes

typically employ a one-size-fits-all approach with few personalised recommendations [5]. Given the growing interest in precision medicine (see Text box, 'Precision medicine'), recent research has focused on classifying diabetes into subtypes (or clusters, phenotypes or endotypes) that are more granular than the current type 1 and type 2 distinction [1, 6–9].

Precision medicine

According to the ADA/EASD Precision Medicine in Diabetes Initiative (PMDI), 'Precision medicine focuses on minimizing errors and improving accuracy in medical decisions and health recommendations. (…) Precision medicine emphasizes tailoring diagnostics or therapeutics (prevention or treatment) to subgroups of populations sharing similar characteristics' [10]. The PMDI defines four key pillars of precision medicine as follows [7]:

- **Precision diagnosis.** 'Involves refining the characterisation of the diabetes diagnosis for therapeutic optimisation or to improve prognostic clarity using information about a person's unique biology, environment and/or context'.
- **Precision prognostics.** 'Focuses on improving the precision and accuracy with which a patient's disease-related outcomes are predicted (…)'.
- **Precision prevention.** 'Includes using information (…) to determine [a person's] likely responses to health interventions and risk factors and/or to monitor progression towards disease'.
- **Precision treatment.** 'Involves using information (…) to guide the choice of an efficacious therapy to achieve the desired therapeutic goal or outcome, while reducing unnecessary side effects'.

In their seminal paper published in 2018, Ahlqvist et al [8] used an exploratory clustering approach to show that newly diagnosed adult-onset diabetes (including both type 1 and type 2 diabetes) could be classified into five different subtypes (or subgroups). Their precision diagnosis approach was based on data-driven hard clustering using six clinical features: GADA, HbA_{1c}, BMI, age at onset of diabetes, HOMA2-B and HOMA2-IR. HOMA2-B and HOMA2-IR are proxy measures of insulin secretion and insulin resistance, which are computed from fasting glucose and fasting insulin or C-peptide concentrations [11]. The five subtypes were labelled and described as follows [8, 12]:

1. severe autoimmune diabetes (SAID): GADA positive; early disease onset; low BMI; insulin deficient; high HbA_{1c}
2. severe insulin-deficient diabetes (SIDD): GADA negative; otherwise similar to SAID
3. severe insulin-resistant diabetes (SIRD): GADA negative; late disease onset; high BMI; high insulin resistance
4. moderate obesity-related diabetes (MOD): GADA negative; early disease onset; high BMI; low insulin resistance (also known as 'mild' obesity-related diabetes)
5. moderate age-related diabetes (MARD): GADA negative; late disease onset; low BMI; low insulin resistance (also known as 'mild' age-related diabetes)

In this framework, adult-onset type 1 diabetes (typically GADA positive [1]) corresponds to SAID, whereas type 2 diabetes is subdivided into SIDD, SIRD, MOD and MARD. Thus, the key innovation lies in the more granular subclassification of type 2 diabetes (see Fig. 1) [6, 7, 13]. Regarding precision prognostics, several studies have since shown that the subtypes differ in their risk of glycaemic, microvascular and macrovascular outcomes [6, 8, 9, 14]. For example, individuals with SIRD may have the highest risk of chronic kidney disease, whereas those with SIDD may have the highest risk of retinopathy.

The proposed diabetes subtypes provide an attractive framework for type 2 diabetes precision medicine. However, discrete subtyping (i.e. hard clustering) also comes with methodological limitations that are sometimes overlooked. The aim of this narrative review is to raise awareness of these limitations by highlighting three key issues: (1) uncertainty about the existence of four discrete aetiological subtypes; (2) low certainty in subtype assignment; and (3) the limited evidence for prognostic and therapeutic value compared with other approaches. As an alternative to subtyping, we discuss recently proposed individualised prediction models, which aim to inform on precision treatment by providing personalised treatment recommendations based on individual patient-level clinical features.

Methodological challenges of type 2 diabetes subtyping

Uncertainty regarding the existence of four discrete subtypes A frequently cited argument in favour of the four type 2 diabetes subtypes is that they have been replicated in several independent studies across diverse populations [6, 9]. Initially, Ahlqvist et al [8] replicated the diabetes subtypes in three Scandinavian cohorts. According to the Precision

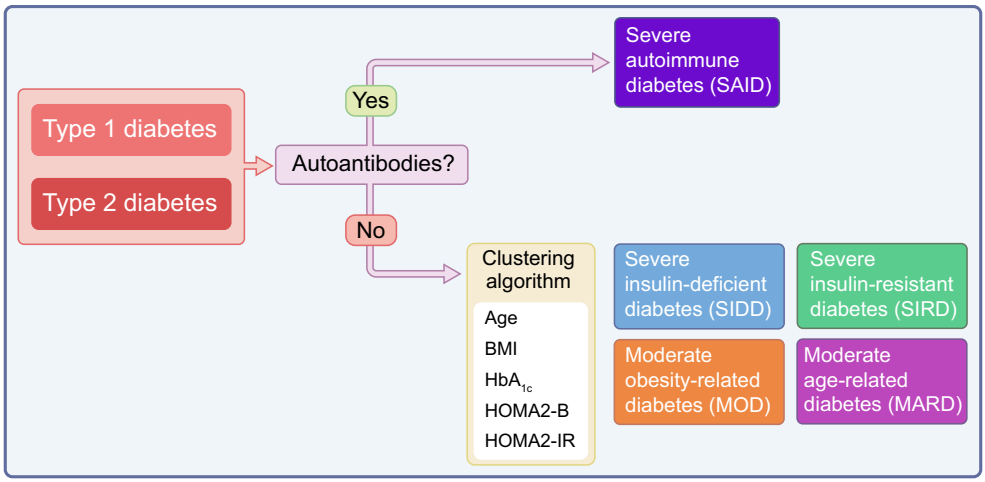


Fig. 1 Overview of the precision diagnostic approach of Ahlqvist et al [8]. SAID corresponds to adult-onset type 1 diabetes (due to the presence of autoantibodies), whereas SIDD, SIRD, MOD and MARD correspond to subtypes of type 2 diabetes derived using a data-driven clustering algorithm. Note that the clustering algorithm (*k*-means) was applied only to individuals with type 2 diabetes (i.e. those without autoantibodies). This figure is available as part of a downloadable slideset

Medicine in Diabetes Initiative (PMDI) systematic review on subclassification of type 2 diabetes [6], the subtypes have since been directly replicated in at least 22 studies that included around 88,000 individuals from diverse ethnic and ancestral origins (e.g. in German [14], North American [15], Chinese [15] and Japanese [16] cohorts).

However, there is substantial methodological variability in how these different studies aimed to replicate the four type 2 diabetes subtypes. Varghese et al [17] distinguish between three different levels of replication. The first level directly uses the cluster coordinates identified in the original Swedish cohort (All New Diabetics in Scania [ANDIS]) and applies them to another population (e.g. [14, 18]). This has also been referred to as the 'nearest centroid approach' [14]. The second level involves de novo clustering in a different population but with the a priori assumption that there are exactly four discrete subtypes (e.g. [15, 19]). In *k*-means cluster analysis, this corresponds to having a predefined *k*-value that determines the number of clusters (i.e. k=4). The third level involves de novo clustering without such an a priori assumption, therefore allowing a different number of clusters to emerge from the analysis (e.g. [16, 20]).

The first two approaches automatically presume the existence of four discrete subtypes and therefore do not replicate the cluster analysis as a whole. The most convincing evidence would stem from the third approach, with confirmation in independent studies that four clusters provide an ideal subclassification of type 2 diabetes. However, in their literature review, Varghese et al [17] found that only 18 of 43 replication studies (42%) followed the third approach. Of these 18 studies, only seven (39%) were able to successfully replicate the four type 2 diabetes subtypes and the remaining 11 studies either identified different subtypes or failed to identify all of the subtypes proposed by Ahlqvist et al [8]. For example, Christensen et al [20] found only three instead of four type 2 diabetes subtypes when applying the same clustering procedure (i.e. *k*-means clustering with the same variables) in a Danish cohort.

However, it should be noted that not all of these replication studies had access to all five continuous clustering variables used by Ahlqvist et al [8]. In particular, HOMA2 measures were unavailable in four out of 18 studies (22%) [17]. Indeed, there has been considerable interest in replicating the diabetes subtypes with more readily available clinical features (i.e. without HOMA2 measures) [6, 21–24]. For example, Lugner et al [21] performed *k*-means clustering on data from the National Diabetes Register of Sweden using a different set of variables. Instead of HOMA2 measures they included six other routinely measured clinical features (systolic BP, diastolic BP, HDL-cholesterol, LDL-cholesterol, triacylglycerols and eGFR). However, a recent study suggests that such modified sets of clustering variables may not be suitable for identifying the four subtypes using de novo clustering [23]. In particular, HOMA2 measures appear to be indispensable for identifying individuals with SIRD [23, 25]. Therefore, replication studies with modified sets of clustering variables cannot provide definitive evidence for or against the four diabetes subtypes proposed by Ahlqvist et al [8].

Nevertheless, they illustrate that the results of diabetes cluster analyses are not unequivocal but are dependent on the specific input variables, clustering algorithm and population [6].

From a methodological perspective, a challenge of the subtypes is that they were derived from a data-driven cluster analysis (analogous to unsupervised learning), meaning that no 'ground truth' was available [26]. In other words, there is no way of knowing whether the proposed diabetes clusters correspond to true biological subgroups of diabetes, or whether they arise because a set of interdependent variables (e.g. BMI, insulin resistance, HbA_{1c}) is entered into the clustering algorithm. In this sense, clustering can be seen as another way of capturing correlations between clinical features. However, unlike correlation analysis, it does not provide inferential tools like CIs or *p* values [27, 28]. Instead, cluster analysis relies on heuristic methods such as elbow plots or gap statistics to determine the optimal number of clusters from the data [27]. In other words, while cluster analysis is a useful tool for grouping similar data points together, its utility for discovering biologically distinct diabetes subtypes remains unclear.

To assess whether the proposed subtypes could be driven by correlations between clinical features, rather than by an underlying subgroup structure, Aoki et al [29] carried out a simulation study. They simulated a virtual patient dataset without any subgroups but preserving the correlation structure between the clustering variables. The assumed correlations were based on empirical data from the multinational SAVOR-TIMI 53 cardiovascular outcome trial [30]. For example, they included a strong positive correlation (r=0.5) between HbA_{1c} and fasting glucose (used for HOMA2 calculation) and a moderate negative correlation (r=−0.25) between HbA_{1c} and age at diabetes onset [29]. When they applied *k*-means clustering to these simulated data, the algorithm still reproduced the four subtypes (SIDD, SIRD, MOD and MARD), even though no discrete subgroups were present. A limitation of these findings is that the empirical basis for the virtual patient dataset (the SAVOR-TIMI 53 trial) might not be entirely suitable for studying the diabetes subtypes. In particular, the simulation was based on a multi-ancestry population with long-standing diabetes (mean of 8.7 years) and established CVD. Moreover, as the findings are based on a simulation study, they cannot be interpreted as definitive evidence against the existence of discrete subtypes. However, they illustrate that data-driven diabetes clusters could, in theory, be driven by correlations between clinical features rather than by discrete biological subgroups.

Taken together, the findings presented above suggest that there is considerable uncertainty regarding the existence of four discrete subtypes of type 2 diabetes. Nevertheless, the subtypes still appear to capture meaningful and clinically relevant differences in diabetes pathophysiology [8, 14, 31]. Aetiological differences between subtypes are considered further in the discussion section.

Low certainty in subtype assignment As well as the general uncertainty regarding the existence of four discrete subtypes, there is also uncertainty at the individual level when assigning people to these subtypes. While each individual with type 2 diabetes can technically be assigned to a subtype using the clustering algorithm, not every individual will match with that subtype equally well [2, 32]. Even in the original publication of Ahlqvist et al [8], a close examination of the distribution of clinical features reveals large variability within each subtype. For example, not every individual with SIRD shows severe insulin resistance and not every individual with SIDD exhibits reduced insulin secretion. In line with this, Pearson [2] argues that the phenotypic variability of type 2 diabetes may be difficult to capture with just four categories and that individual classification probabilities may be low.

In practice, this means that an individual's clinical profile may lie between two subtypes (e.g. sharing similarities with both MOD and MARD) [25, 33] or not match any of the four subtypes [2]. For an illustrative example of borderline cases falling between two subtypes, consider the two individuals shown in Table 1. The clinical profiles of the two individuals are almost

Table 1 Example of two individuals who lie between two type 2 diabetes subtypes

Variable	Individual A	Individual B
GADA	Negative	Negative
Sex	Female	Female
HbA_{1c} (mmol/mol)	65	65
HbA_{1c} (%)	8.1	8.1
BMI (kg/m^2)	30.0	30.5
Age at onset (years)	58	58
HOMA2-B	121	121
HOMA2-IR	2.83	2.83
Subtype assignment	MARD	MOD

While recent evidence shows some differences in complications between the MOD and MARD subtypes [12], the focus is usually more on the severe SIDD and SIRD subtypes, which are further apart in terms of Euclidean distances

identical, except for a 0.5 kg/m^2 difference in BMI. Yet, the clustering algorithm (nearest centroid approach based on the cluster coordinates from the ANDIS cohort [8]) assigns individual A to the MARD subtype and individual B to the MOD subtype. For such individuals, the subtype assignment is unstable, as small changes in clinical variables may lead to a different subtype. Consequently, their classification certainty is low and subtype-specific risk predictions may be less reliable. Moreover, measurement error as well as longitudinal biological changes (e.g. in HbA_{1c} or BMI) may lead to changes in subtype assignment [34]. For example, in the German Diabetes Study almost one out of four individuals changed their subtype within 5 years after diagnosis [14]. Besides natural disease progression, subtype assignment may also change in response to treatment, as some of the cluster variables represent therapeutic targets (e.g. HbA_{1c} and BMI) [12]. For example, an individual with a BMI of 35 kg/m^2 at diagnosis might change their subtype after successful weight loss following treatment with glucagon-like peptide-1 receptor agonists (GLP1-RAs).

In response to this problem, the second PMDI consensus report [10] recommends that diabetes subtype assignments should always be accompanied by a measure of classification uncertainty. In principle, there are several measures that can be used to assess an individual's classification uncertainty [35], ranging from graphical displays (e.g. spider plots) to statistical measures (e.g. Euclidean distance [8], distance ratio [22], classification probability [36, 37] and normalised relative entropy [NRE] [35]) (see Text box, 'Measures of

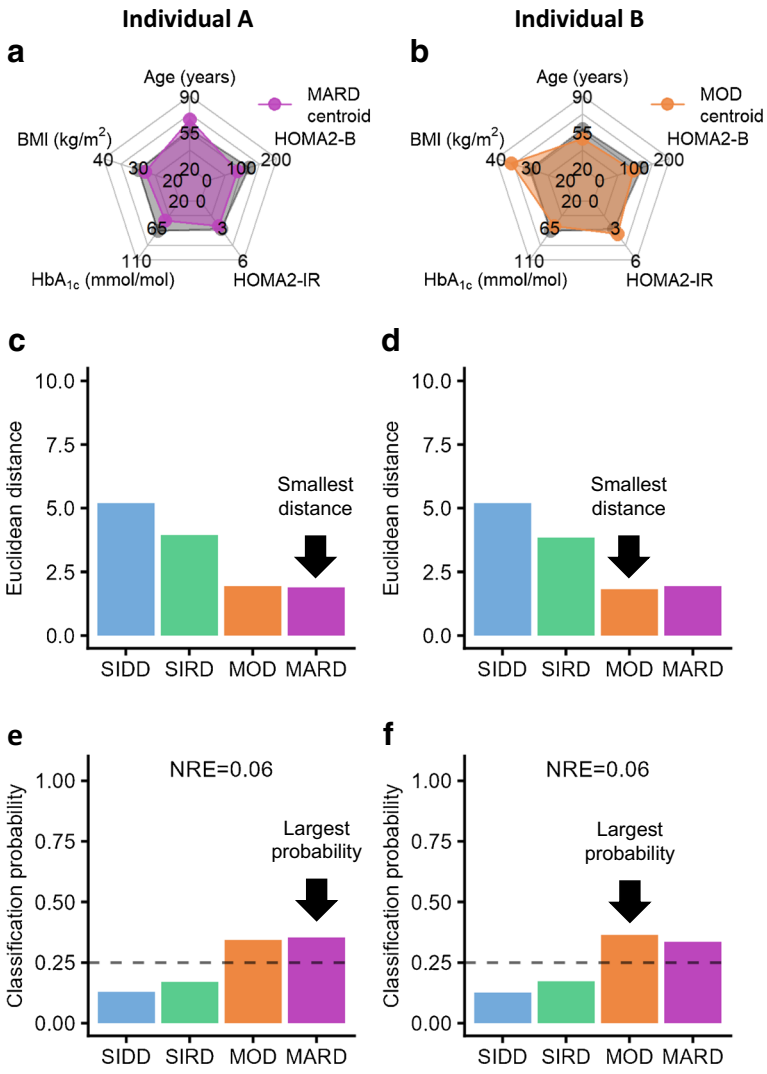


Fig. 2 Illustrative example of different measures of classification uncertainty applied to individual A (**a**, **c**, **e**) and individual B (**b**, **d**, **f**) from Table 1. (**a**, **b**) The spider plots show that both individuals share a similar clinical profile, which does not match entirely with either MARD or MOD. (**c**, **d**) Because of their higher BMI by 0.5 kg/m^2, individual B has a minimally smaller Euclidean distance to the MOD centroid (i.e. the typical MOD profile). (**e**, **f**) As a result, the classification probability for MOD is slightly higher than that for MARD. Because both individuals lie at the border between MOD and MARD, they have a low classification certainty, as indicated by the low classification probabilities and NRE values close to zero. This figure is available as part of a downloadable slideset

classification uncertainty'). Figure 2 shows an illustrative example of different measures of classification uncertainty applied to individuals A and B from Table 1.

Measures of classification uncertainty

- A **spider plot** visualises the clinical profile of an individual and compares it with the typical profile of their assigned subtype.
- The **Euclidean distance** measures the degree of dissimilarity between an individual and a subtype centroid (i.e. the typical subtype profile).
- The **distance ratio** ranges from 0 to 1 and compares the smallest Euclidean distance (to the assigned subtype) with the second smallest (to the next closest subtype).
- The **classification probability** ranges from 0% to 100% and measures the probability that an individual belongs to a subtype.
- The **NRE** ranges from 0 (complete uncertainty) to 1 (complete certainty) and measures the overall classification certainty of an individual based on the distribution of the four classification probabilities.

In practice, however, measures of classification uncertainty have rarely been reported in studies of the type 2 diabetes subtypes [22, 35–37]. A recent study proposed using the NRE statistic to quantify an individual's classification uncertainty on a scale from 0 (complete uncertainty) to 1 (complete certainty) [35]. Application of this method to a cohort of 859 individuals from the German Diabetes Study revealed a low median NRE of 0.13 (95% CI 0.12, 0.14), indicating substantial classification uncertainty. While the degree of uncertainty might vary depending on the cohort and the chosen measure, current evidence shows that a certain level of classification uncertainty is inevitable [22, 35–37]. It is therefore essential to acknowledge this limitation of discrete subtyping and to routinely report uncertainty measures.

A general problem of hard subtype assignment is that, by definition, it reduces the information available in a system. Even when the classification probability for one subtype is clearly higher than the probabilities for the other subtypes, the relative magnitude of the classification probabilities still carries information about an individual's clinical profile. Indeed, it has been suggested that the classification probabilities or Euclidean distances should be directly related to clinical outcomes, rather than the hard subtype assignment [26]. Alternatively, measures of classification uncertainty can be used to mitigate some of the information loss that occurs when comparing clinical outcomes between hard subtype assignments [35]. For example, individuals can be weighted by their NRE, such that those who are less representative of their subtype contribute less to the analysis. Another option is to use alternative prediction-based approaches for precision medicine instead of diabetes subtypes, as discussed in the next section.

Limited prognostic and therapeutic value of subtypes The third challenge of the subtyping approach is that the information loss due to hard subtype assignments is disadvantageous for predicting clinical outcomes (i.e. precision prognostics) [26, 34]. Instead, it would be more informative to use statistical models that directly incorporate continuous clinical features as predictors. This issue has been investigated in several studies comparing the performance of type 2 diabetes subtypes with conventional prediction models [19, 29, 38].

Aoki et al [29] compared the predictive performance of the subtypes to that of established clinical risk scores using data from the SAVOR-TIMI 53 trial, which included individuals with established CVD and long-standing diabetes (mean of 8.7 years). They employed the widely used Systematic Coronary Risk Evaluation (SCORE) [39] and the Kidney Disease: Improving Global Outcomes (KDIGO) classification [40] to predict CVD and kidney outcomes, respectively. Even though the CVD SCORE was treated as a broad ordinal variable (10 year CVD risk of ≤5%, 5–10%, 10–15%), its prognostic performance was similar to that of the subtypes (concordance index [c-index] 0.57 and 0.56 [95% CI not reported], respectively). For kidney events, prediction using the KDIGO classification clearly outperformed the subtype approach (c-index 0.62 and 0.54 [95% CI not reported], respectively). Similar results were found by Bjarkø et al [38], who assessed the predictive performance of the subtypes in a Norwegian population-based study. Using data from the Trøndelag Health Study (HUNT Study), they found that individual established risk factors (e.g. HbA_{1c}, eGFR) were at least as good as the subtypes for predicting vascular complications (e.g. retinopathy, chronic kidney disease). In a re-analysis of the ADOPT [41] trial (median follow-up time of 4 years), Dennis et al [19] found that age at diagnosis alone was as informative as the subtype classification for predicting glycaemic disease progression (R^2=0.09 and 0.08, respectively).

Beyond precision prognostics, Dennis et al [19] critically examined the subtypes' potential for precision treatment. They found that the subtypes differed modestly in their response to glucose-lowering drugs in the ADOPT trial, which compared metformin, sulfonylurea and thiazolidinedione treatment. However, a model that directly used the clustering variables as predictors, even without HOMA2

measures, explained individual treatment response substantially better. When predicting response to sulfonylureas, the model explained 33% of the variability in treatment response, compared with 20% for the subtype-based approach. Similar results were found when predicting response to thiazolidinedione treatment (R^2_{model}=0.32 vs $R^2_{subtypes}$=0.17) and metformin treatment (R^2_{model}=0.35 vs $R^2_{subtypes}$=0.15). To further compare these two approaches, Dennis et al [19] performed an external validation using the RECORD trial [42] (mean follow-up time of 5.5 years), which investigated the same medications. They found that treatment recommendations based on the prediction model yielded greater benefits than those based on subtypes, with an additional HbA_{1c} reduction of 3.9 mmol/mol (0.36 percentage points) compared with 1.8 mmol/mol (0.16 percentage points).

A limitation of these findings is that they are based on selective populations from randomised clinical trials, which might not be representative of the general type 2 diabetes population. The ADOPT trial [41] included recently diagnosed individuals without prior pharmacological therapy and with fasting plasma glucose levels of 7.0–10.0 mmol/l. The RECORD trial [43] included individuals on metformin or sulfonylurea monotherapy and with HbA_{1c} levels of 53–75 mmol/mol (7–9%). As a result, the average HbA_{1c} value in the identified SIDD subtype was substantially lower (median: ADOPT, 67 mmol/mol [8.3%]; RECORD, 72 mmol/mol [8.7%]) [19] than in the original study by Ahlqvist et al [8] (mean: ANDIS, 102 mmol/mol [11.5%]). Moreover, differences in the risk of kidney disease, repeatedly shown in other cohorts [8, 14, 16], could not be replicated clearly in these cohorts [19]. A possible explanation is that these cohorts excluded individuals with baseline renal impairment, resulting in a relatively high median eGFR at baseline even for the SIRD subtype (ADOPT, 90 ml/min per 1.73 m^2; RECORD, 97 ml/min per 1.73 m^2) [19]. Finally, both ADOPT [41] and RECORD [43] excluded individuals with comorbidities such as heart failure and did not investigate more-recent therapies such as sodium–glucose cotransporter 2 inhibitors (SGLT2i) and GLP1-RAs.

An overarching limitation of the above studies is that they involved retrospective analyses of cohorts that were not originally designed to investigate the type 2 diabetes subtypes [44]. To our knowledge, there has been only one prospective clinical trial that was explicitly designed to assess differences in treatment response among the diabetes subtypes [45]. In this trial, Dwibedi et al [45] assessed whether individuals with SIDD (n=126) and SIRD (n=113) respond differently to glucose-lowering treatment with GLP1-RA and SGLT2i. The study was designed to detect a treatment effect difference (i.e. a treatment-by-subgroup interaction) of 3.5 mmol/mol (0.32 percentage points) in 6 month HbA_{1c} with a statistical power of 80%. However, their results showed no differential treatment response (i.e. no statistically significant interaction), with GLP1-RA being more effective than SGLT2i in both SIDD and SIRD (HbA_{1c} benefit of 8.6 mmol/mol [0.79 percentage points] and 7.8 mmol/mol [0.71 percentage points], respectively). Instead, they found a statistically significant interaction between treatment and baseline HbA_{1c}, suggesting that GLP1-RAs are increasingly more effective than SGLT2i at higher baseline HbA_{1c} levels. In other words, their results again suggested that continuous pathophysiological variables might be more useful than diabetes subtypes for predicting glucose-lowering treatment response. However, there are two major limitations to this clinical trial. First, the strict eligibility criteria (e.g. requirement for metformin monotherapy) excluded individuals with more progressive diabetes and resulted in SIDD and SIRD groups that were less distinct from each other. For example, the mean baseline HbA_{1c} level of individuals with SIDD and SIRD was 57.5 mmol/mol (7.4%) and 53.4 mmol/mol (7.0%), respectively. Second, the trial only had a follow-up time of 6 months and focused on glycaemic outcomes. However, proponents of subtyping typically focus on the importance of SGLT2i and GLP1-RA treatment for reducing cardio-renal complications, rather than glucose, in individuals with SIRD [12, 46]. More detailed considerations of potential treatment options for the diabetes subtypes based on their pathophysiology have been provided in other reviews [9, 46].

From a methodological perspective, the aforementioned findings regarding the potentially limited prognostic and therapeutic value of subtypes are not surprising, given that clustering methods (analogous to unsupervised learning) are not designed to achieve the best possible prediction of future clinical outcomes [34]. Instead, they focus on cross-sectional data to identify homogenous groups of individuals with distinct clinical features (i.e. precision diagnosis). If interest lies in precision prognostics or precision treatment, then dedicated prediction methods (analogous to supervised learning) might be preferable, as those are designed to achieve an optimal prediction of future clinical outcomes.

Implications for type 2 diabetes precision medicine

Individualised prediction models as an alternative to subtyping Building on the work above [19], Dennis and colleagues continued to develop and apply the 'individualised prediction' framework for precision treatment of type 2 diabetes [47–51]. In this framework [47], the aim is to use simple clinical features to directly model treatment response (i.e. precision treatment) rather than to identify distinct aetiologies of diabetes (i.e. precision diagnosis). The goal is to

identify the optimal treatment for an individual based on their clinical features at a particular point in time, without requiring an explicit subtype assignment. To ensure applicability in clinical practice, Dennis [47] emphasised the benefit of using routinely available clinical features, such as age, BMI and eGFR, to provide a low-cost equitable precision medicine approach. So far, applications of the individualised prediction approach have focused on treatment selection models (or 'treatment algorithms') for second- and third-line glucose-lowering drugs [48, 49, 51].

A recent major advance within this framework is the development of a 'five-drug-class model' (available at www.diabetesgenes.org/t2-treatment/) [51]. This model allows simultaneous prediction of an individual's 12 month HbA_{1c} response to any of the five glucose-lowering drug classes commonly prescribed after metformin in primary care: SGLT2i, GLP1-RAs, dipeptidyl peptidase-4 inhibitors, sulfonylureas and thiazolidinediones. The model was developed using observational primary care data from the UK Clinical Practice Research Datalink (CPRD), comprising more than 100,000 drug initiations. Nine routine clinical features (age, duration of diabetes, sex, HbA_{1c}, BMI, eGFR, HDL-cholesterol, total cholesterol, alanine aminotransferase [ALT]) were included as predictors of differential treatment response. An illustrative example of the five-drug-class model applied to individual A from Table 1 is shown in Fig. 3.

Model validation was performed in an independent CPRD validation cohort as well as in three clinical trials, including a three-way crossover trial [52]. In the CPRD cohort, applying a clinically relevant threshold of ≥3 mmol/mol (0.27 percentage points), the model predicted optimal therapies for the majority of initiations, with 54.9% (n=116,525) having one or two drug classes predicted to provide a greater HbA_{1c} reduction than other drug classes. The most frequently predicted optimal therapies for HbA_{1c} reduction were GLP1-RAs (optimal in 33.4% of treatment initiations), SGLT2i (28.9%) or sulfonylureas (27.6%) [51]. Treatment in line with these recommendations was associated with a mean additional HbA_{1c} reduction of 5.0–5.3 mmol/mol (0.46–0.48 percentage points) compared with non-optimal treatment. The model-predicted HbA_{1c} benefits showed good calibration in external validation. Absolute observed HbA_{1c} benefits were smaller in the more selective clinical trial populations (mean additional HbA_{1c} reductions of 1.9–2.4 mmol/mol [0.17–0.22 percentage points]) where mean baseline HbA_{1c} was lower than in the real-world datasets. CPRD analysis suggested following the model recommendations was also associated with a 38% lower risk of glycaemic failure (reaching an HbA_{1c} ≥69 mmol/mol [8.5%]) up to 5 years post therapy initiation, although long-term benefits remain to be confirmed in external data. Overall, these initial findings position the individualised prediction approach as a promising alternative to subtype-based stratification for precision treatment selection.

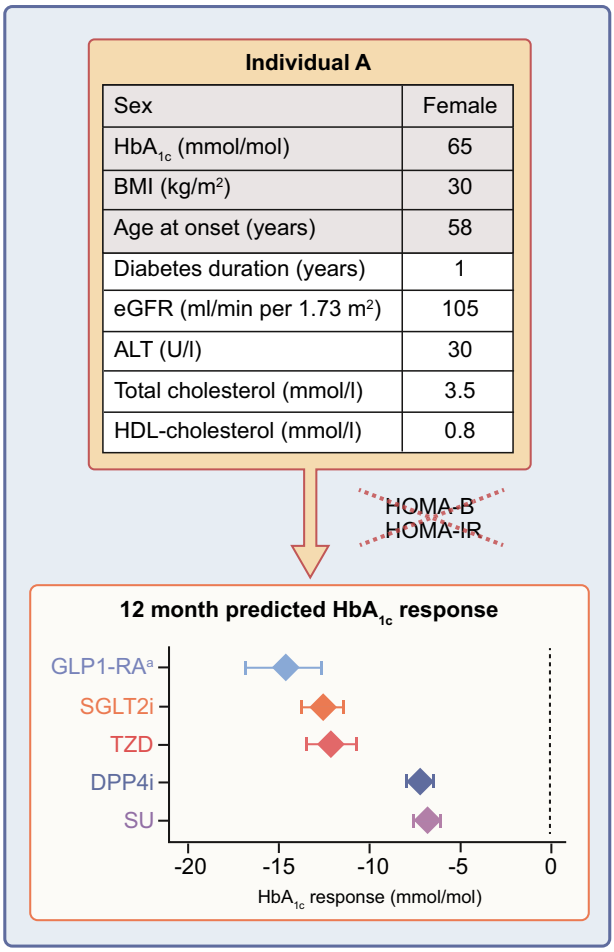


Fig. 3 Illustrative example of the five-drug-class model [51] applied to individual A from Table 1. Similar to the subtyping approach, the model requires as input an individual's sex, HbA_{1c}, BMI and age at treatment decision (grey features in table). While the model does not require HOMA2-B and HOMA2-IR, it depends on five additional routine clinical features (white features in table): diabetes duration, eGFR, ALT, total cholesterol and HDL-cholesterol. For individual A, the model predicts that treatment with GLP1-RA, SGLT2i or thiazolidinedione (TZD) will achieve the highest glucose-lowering response. While a slightly larger benefit is predicted for GLP1-RA, the overlapping CIs suggest that SGLT2i and TZD may be similarly effective. In such a setting of clinical equipoise, a shared decision-making process that incorporates patient preferences (e.g. oral medication vs injectable) could support treatment selection [5]. Note that, because the clinical variables are modelled continuously, small-value changes (e.g. 0.5 kg/m^2 higher BMI of individual B) have little effect and lead to the same therapies being predicted as optimal (data not shown). [a]Model development excluded semaglutide. DPP4i, dipeptidyl peptidase 4 inhibitor; SU, sulfonylurea. This figure is available as part of a downloadable slideset

Nevertheless, there are also limitations to this individualised prediction approach. First, the focus lies on optimising treatment selection for lowering glucose rather than preventing cardio-renal complications [53]. However, validation of

the five-drug-class model in observational CPRD data also showed an associated lower risk of cardiovascular and renal complications after 5 years, likely reflective of the frequent recommendation of GLP1-RAs and SGLT2i [51]. Second, the frequent recommendation of these novel medications (62.4% of treatment initiations in CPRD were recommended GLP1-RA or SGLT2i) could lead to increased healthcare costs. However, treatment recommendations were often not limited to a single drug class (76.4% of initiations in CPRD had two or more optimal drug classes) and the model allows for flexible deployment by omitting treatments that might not be available (e.g. due to costs or contraindications) [51]. Third, prediction models developed from observational data come with their own methodological challenges such as overfitting, confounding and incomplete data [54, 55]. While initial results seem reassuring [51, 56], careful validation [57] and calibration [58] of these models remains crucial to support decision-making in clinical practice.

The role of subtypes in type 2 diabetes precision medicine One main advantage of diabetes subtypes over individualised prediction models is that they may provide insights into the underlying pathophysiology of the disease [31]. Ahlqvist et al [59] emphasised that, even though they employed data-driven clustering methods, their subtyping approach was not hypothesis free. Their variable selection reflected major pathways in the development of type 2 diabetes (e.g. insulin resistance) and captured disease heterogeneity in a clinically meaningful way. In principle, clustering methods can help identify subgroups with distinct clinical features that would otherwise be clouded out by the majority. This can then prompt investigation into which pathophysiological mechanisms may have caused this subgroup to emerge. For example, a recent study found that individuals with SIRD had a higher intrahepatic lipid content than the other subtypes; this, in turn, was associated with a higher cardiovascular risk [60]. The notion that the subtypes differ in terms of their aetiology is also supported by genetic evidence [9, 25, 61]. For example, unlike the other subtypes, SIRD is not associated with polygenic risk scores for impaired insulin secretion [62]. Given such pathophysiological insights, Ahlqvist et al [31] have argued that the subtypes provide 'a more holistic view of type 2 diabetes' than individualised prediction models.

In terms of its underlying assumptions, the subtype framework differs from the 'palette model' of type 2 diabetes heterogeneity [3]. The latter emphasises that disease heterogeneity arises from multiple component pathways that jointly contribute to an individual's phenotype, often resulting in a 'mixed aetiology'. While Ahlqvist et al [25] acknowledge that diabetes is caused by many overlapping mechanisms, they make the assumption that most individuals are predominantly affected by one pathway. A key advantage of the subtyping approach is that discrete subtypes are easy to communicate and can provide a broad range of clinical information [9, 47]. Rather than having to use a different prediction model for each clinical outcome, the subtype assignment can feasibly provide an overview based on established subtype-specific risks of complications. That said, including a measure of classification uncertainty is advisable to gauge how applicable subtype-specific predictions are for a given individual [35].

When discussing the feasibility of the subtyping approach, it is important to also consider logistical obstacles to implementation in clinical practice. While a user-friendly web application has recently been developed [63], subtype assignment requires HOMA2 indices of insulin resistance and insulin secretion [8], particularly for correctly identifying individuals with SIRD [25]. Calculating these HOMA2 values introduces additional computational complexity, as it requires a separate (although freely available) calculator program [11]. The more fundamental challenge is that the calculation requires fasting insulin or C-peptide measurements as input. However, these are not routinely collected in clinical practice, making a widespread implementation challenging at present [7, 21, 22, 24]. In addition, access to these measurements may be more restricted in low-income and middle-income countries with resource-constrained healthcare systems [64]. However, as C-peptide measurements are relatively inexpensive (e.g. compared with novel GLP1-RA medications), a broader rollout could be possible in the future [12].

Conclusions and future directions

Type 2 diabetes subtyping has made important contributions towards precision medicine but it also comes with several methodological limitations. It is important to be mindful of challenges such as low classification certainty and practical implementation barriers due to use of non-routine biomarkers. Individualised prediction models offer an actionable alternative by providing personalised treatment recommendations based on routine clinical features. However, they do not provide insights into the underlying pathophysiology of the disease and thus might not paint a complete picture of diabetes heterogeneity. As such, these two approaches can be considered as complementary, each serving different purposes in precision medicine [59, 61, 65].

From a methodological perspective [66], clustering methods (analogous to unsupervised learning) aim to delineate groups with distinct clinical features by minimising within-cluster differences. In contrast, prediction methods (analogous to supervised learning) are designed to achieve the best possible prediction of future outcomes by minimising

prediction errors. In other words, these two approaches optimise for different parameters, which naturally leads to different use cases [34]: for precision treatment, individualised prediction models might be preferable, as advocated by the second PMDI consensus report [10]; for precision diagnostics, diabetes subtyping can be employed to gain insights into patterns of phenotypic variability and pathophysiology; and for precision prognostics, diabetes subtyping can be supplemented by outcome-specific prediction models, such as SCORE2-Diabetes for cardiovascular outcomes [67].

For precision prevention, a distinct set of subtypes among people at risk of type 2 diabetes has been proposed [68–70]. The methodological limitations of type 2 diabetes subtypes outlined above directly transfer to these subtypes of people at risk of type 2 diabetes and, more broadly, to any data-driven clustering approach. By raising awareness of these challenges through this narrative review, we hope to enable a more careful application of diabetes subtyping in research and clinical practice.

Regarding the challenge of classification uncertainty, the argument can be extended to any type of diabetes classification that uses a threshold. For example, an individual with an HbA_{1c} of 48 mmol/mol (6.5%) would be classified as having diabetes, whereas an identical individual with an HbA_{1c} of 46 mmol/mol (6.4%) would not [1]. Similarly, a very small difference in GADA titre could determine whether a person is 'autoantibody positive' and thus considered to have type 1 or type 2 diabetes [71]. For a detailed discussion of the challenges associated with delineating type 1 and type 2 diabetes, see Gale [71], which provides a useful parallel to our discussion of type 2 diabetes subtypes.

As for the individualised prediction approach, future research could consider models with multivariate outcomes to allow for simultaneous prediction of several outcomes (e.g. HbA_{1c}, kidney function, cardiovascular risk). In the future, incorporation of prediction models into decision support systems coupled to electronic medical records could provide a feasible implementation for clinical practice. At the same time, the predicted optimal treatment may not necessarily be consistent across different outcomes, making it potentially difficult to derive treatment recommendations for an individual in such a scenario.

Reflecting on precision medicine more broadly, the second PMDI consensus report [10] raises concerns about a lack of adequate validation studies. In particular, there is a lack of prospectively designed clinical trials to validate proposed precision medicine approaches. One exception is the clinical trial mentioned above, which was explicitly designed to assess differential treatment responses of individuals classified as SIDD and SIRD [45]. Another notable example is the TriMaster study [52], a three-way crossover trial that demonstrated differential responses to three glucose-lowering medications depending on individuals' kidney function and BMI. However, such precision medicine trials remain scarce, and the PMDI consensus report concludes that the clinical translation of precision medicine in diabetes care is, at present, 'largely aspirational' [10].

Consequently, an important objective for future research is to generate high-quality evidence from appropriately designed clinical trials [72–74]. For example, repeated crossover trials (or '*N*-of-1-trials'), in which each individual receives at least one treatment twice, are particularly useful to address precision medicine questions [75, 76]. This trial design is uniquely suited to demonstrate treatment effect heterogeneity. In other words, it enables researchers to determine whether, and to what extent, individuals truly differ in their responses to different glucose-lowering medications [72, 77]. While no such trials have yet been conducted for glucose-lowering medications, a successful example has recently been published for blood-pressure-lowering treatment [78].

Beyond establishing treatment heterogeneity, the second prerequisite for precision medicine is the identification of clinical features that alter treatment response [72]. For example, the five-drug-class model suggests that routinely collected features such as BMI, eGFR and cholesterol can identify people who would benefit more from one treatment than another [51]. To evaluate such precision medicine approaches, a pragmatic strategy would be a randomised trial in which one group of healthcare providers uses the treatment algorithm and the other group does not [47]. Such a design would enable a comparison of the precision medicine approach with current standards of care in a routine clinical practice setting [47, 79]. Regardless of the study design, future research should adhere to the BePRECISE reporting guidelines [80] to facilitate evaluation of the clinical utility of proposed precision medicine strategies.

In conclusion, this narrative review highlights three key methodological challenges of type 2 diabetes subtypes: (1) uncertainty about the existence of four discrete subtypes; (2) low classification in subtype assignment; and (3) limited evidence for prognostic and therapeutic value compared with other approaches. As an alternative, we discuss individualised prediction models that avoid some of the limitations of discrete subtyping and provide personalised treatment recommendations based on routinely collected clinical features.

Supplementary Information The online version contains a slide-set of the figures for download available at https://doi.org/10.1007/s00125-026-06802-6.

Acknowledgements Parts of this work are based on TM's PhD thesis [81]. Figure 1 has been adapted and parts of the introduction and discussion section of the PhD thesis have been reused.

Funding Open Access funding enabled and organized by Projekt DEAL. The German Diabetes Center is funded by the German Federal Ministry of Health and the Ministry of Culture and Science of the state

of North Rhine-Westphalia. The German Center for Diabetes Research is funded by the German Federal Ministry of Education and Research. JMD and PC are supported by a Wellcome Trust Early-Career Award (227070/Z/23/Z) and the NIHR Exeter Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. TM is supported by the EASD mentorship programme. CH was supported by a grant from Horizon Europe (Intercept-T2D, 101095433). Neither of these funders had any role in the design, execution, analyses, interpretation of data or decision to submit results in this study. In particular, the research reported here received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. The sole responsibility for the content of this publication lies with the authors.

Authors' relationships and activities CH is Reviews Editor for *Diabetologia*. JMD received an honorarium from Abbott. OK received honoraria for biostatistical education from Berlin-Chemie. The authors declare that there are no other relationships or activities that might bias, or be perceived to bias, their work.

Contribution statement TM conceptualised the review, prepared the visualisations, wrote the first draft and edited and reviewed the manuscript. CH, PC, and JMD contributed substantially to the conceptualisation of the review and to data interpretation, and reviewed it critically for important intellectual content. OK contributed substantially to the conceptualisation of the review and to data interpretation, reviewed it critically for important intellectual content, and provided supervision. All authors gave final approval of this version to be published. TM is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

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. American Diabetes Association Professional Practice Committee (2025) 2. Diagnosis and classification of diabetes: standards of care in diabetes—2025. Diabetes Care 48(Suppl 1):S27–S49. https://doi.org/10.2337/dc25-S002
2. Pearson ER (2019) Type 2 diabetes: a multifaceted disease. Diabetologia 62(7):1107–1112. https://doi.org/10.1007/s00125-019-4909-y
3. 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
4. Gale EAM (2013) Is type 2 diabetes a category error? Lancet 381(9881):1956–1957. https://doi.org/10.1016/S0140-6736(12)62207-7
5. Davies MJ, Aroda VR, Collins BS et al (2022) Management of hyperglycemia in type 2 diabetes, 2022. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetologia 65:1925–1966. https://doi.org/10.2337/dci22-0034
6. Misra S, Wagner R, Ozkan B et al (2023) Precision subclassification of type 2 diabetes: a systematic review. Commun Med 3(1):138. https://doi.org/10.1038/s43856-023-00360-3
7. 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 63:1671–1693. https://doi.org/10.1007/s00125-020-05181-w
8. Ahlqvist E, Storm P, Käräjämäki 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
9. Herder C, Roden M (2022) A novel diabetes typology: towards precision diabetology from pathogenesis to treatment. Diabetologia 65(11):1770–1781. https://doi.org/10.1007/s00125-021-05625-x
10. Tobias DK, Merino J, Ahmad A et al (2023) Second international consensus report on gaps and opportunities for the clinical translation of precision diabetes medicine. Nat Med 29(10):2438–2457. https://doi.org/10.1038/s41591-023-02502-5
11. Wallace TM, Levy JC, Matthews DR (2004) Use and abuse of HOMA modeling. Diabetes Care 27(6):1487–1495. https://doi.org/10.2337/diacare.27.6.1487
12. Asplund O, Thangam M, Prasad RB et al (2026) Comorbidities and mortality in subgroups of adults with diabetes with up to 14 years follow-up: a prospective cohort study in Sweden. Lancet Diabetes Endocrinol 14(1):29–40. https://doi.org/10.1016/S2213-8587(25)00283-9
13. Franks PW, Cefalu WT, Dennis J et al (2023) Precision medicine for cardiometabolic disease: a framework for clinical translation. Lancet Diabetes Endocrinol 11(11):822–835. https://doi.org/10.1016/S2213-8587(23)00165-1
14. 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
15. 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
16. 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
17. Varghese JS, Carrillo-Larco RM, Narayan KMV (2023) Achieving replicable subphenotypes of adult-onset diabetes. Lancet Diabetes Endocrinol 11(9):635–636. https://doi.org/10.1016/S2213-8587(23)00195-X
18. 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
19. 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
20. Christensen DH, Nicolaisen SK, Ahlqvist E et al (2022) Type 2 diabetes classification: a data-driven cluster study of the Danish Centre for Strategic Research in Type 2 Diabetes (DD2) cohort. BMJ Open Diabetes Res Care 10(2):e002731. https://doi.org/10.1136/bmjdrc-2021-002731

21. Lugner M, Gudbjörnsdottir 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
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. Li Z, Liu S, Ho JC, Narayan KMV, Ali MK, Varghese JS (2026) Subtypes of newly diagnosed type 2 diabetes and risk of complications: analysis of electronic health records in the USA. Diabetologia 69(6):1545–1556. https://doi.org/10.1007/s00125-026-06687-5
24. 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
25. Ahlqvist E, Prasad RB, Groop L (2022) 100 YEARS OF INSULIN: towards improved precision and a new classification of diabetes mellitus. J Endocrinol 252(3):R59–R70. https://doi.org/10.1530/joe-20-0596
26. Van Smeden M, Harrell FE, Dahly DL (2018) Novel diabetes subgroups. Lancet Diabetes Endocrinol 6(6):439–440. https://doi.org/10.1016/S2213-8587(18)30124-4
27. Hastie T, Tibshirani R, Friedman J (2009) Cluster analysis. In: The elements of statistical learning: data mining, inference, and prediction. Vol 2. Springer, New York, pp 501–527
28. Toffalini E, Gambarota F, Perugini A et al (2024) Clusters that are not there: an R tutorial and a Shiny app to quantify a priori inferential risks when using clustering methods. Int J Psychol 59(6):1183–1198. https://doi.org/10.1002/ijop.13246
29. Aoki Y, Hamrén B, Clegg LE et al (2021) Assessing reproducibility and utility of clustering of patients with type 2 diabetes and established CV disease (SAVOR-TIMI 53 trial). PLoS One 16(11):e0259372. https://doi.org/10.1371/journal.pone.0259372
30. Scirica BM, Bhatt DL, Braunwald E et al (2013) Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus. N Engl J Med 369(14):1317–1326. https://doi.org/10.1056/NEJMoa1307684
31. Ahlqvist E, Tuomi T, Groop L (2019) Clusters provide a better holistic view of type 2 diabetes than simple clinical features. Lancet Diabetes Endocrinol 7(9):668–669. https://doi.org/10.1016/S2213-8587(19)30257-8
32. Franks PW, Sargent JL (2024) Diabetes and obesity: leveraging heterogeneity for precision medicine. Eur Heart J 45(48):5146–5155. https://doi.org/10.1093/eurheartj/ehae746
33. Leslie RD, Ma RCW, Franks PW, Nadeau KJ, Pearson ER, Redondo MJ (2023) Understanding diabetes heterogeneity: key steps towards precision medicine in diabetes. Lancet Diabetes Endocrinol 11(11):848–860. https://doi.org/10.1016/S2213-8587(23)00159-6
34. Coral DE, Sargent JL, Carrasco-Zanini J et al (2026) Five tenets for advancing evidence-based precision medicine. Nat Med 32(4):1211–1222. https://doi.org/10.1038/s41591-026-04309-6
35. Mori T, Zaharia OP, Straßburger K et al (2025) Recognising, quantifying and accounting for classification uncertainty in type 2 diabetes subtypes. Diabetologia 68(10):2139–2150. https://doi.org/10.1007/s00125-025-06486-4
36. Tanabe H, Sato M, Miyake A et al (2024) Machine learning-based reproducible prediction of type 2 diabetes subtypes. Diabetologia 67(11):2446–2458. https://doi.org/10.1007/s00125-024-06248-8
37. Wesolowska-Andersen A, Brorsson CA, Bizzotto R et al (2022) Four groups of type 2 diabetes contribute to the etiological and clinical heterogeneity in newly diagnosed individuals: an IMI DIRECT study. Cell Rep Med 3(1):100477. https://doi.org/10.1016/j.xcrm.2021.100477
38. Bjarkø VV, Haug EB, Langhammer A et al (2024) Clinical utility of novel diabetes subgroups in predicting vascular complications and mortality: up to 25 years of follow-up of the HUNT Study. BMJ Open Diabetes Res Care 12(6):e004493. https://doi.org/10.1136/bmjdrc-2024-004493
39. Conroy RM, Pyörälä K, Fitzgerald AP et al (2003) Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. Eur Heart J 24(11):987–1003. https://doi.org/10.1016/S0195-668X(03)00114-3
40. Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group (2012) KDIGO clinical practice guideline for acute kidney injury. Kidney Int Suppl 2:1–138. https://doi.org/10.1159/000339789
41. Kahn SE, Haffner SM, Heise MA et al (2006) Glycemic durability of rosiglitazone, metformin, or glyburide monotherapy. N Engl J Med 355(23):2427–2443. https://doi.org/10.1056/NEJMoa066224
42. Home PD, Pocock SJ, Beck-Nielsen H et al (2009) Rosiglitazone evaluated for cardiovascular outcomes in oral agent combination therapy for type 2 diabetes (RECORD): a multicentre, randomised, open-label trial. Lancet 373(9681):2125–2135. https://doi.org/10.1016/S0140-6736(09)60953-3
43. Home PD, Pocock SJ, Beck-Nielsen H et al (2005) Rosiglitazone Evaluated for Cardiac Outcomes and Regulation of Glycaemia in Diabetes (RECORD): study design and protocol. Diabetologia 48(9):1726–1735. https://doi.org/10.1007/s00125-005-1869-1
44. Roden M (2026) The long and winding road to precision diabetology. Lancet Diabetes Endocrinol 14(1):3–5. https://doi.org/10.1016/S2213-8587(25)00303-1
45. Dwibedi C, Ekström O, Brandt J et al (2024) Randomized open-label trial of semaglutide and dapagliflozin in patients with type 2 diabetes of different pathophysiology. Nat Metab 6(1):50–60. https://doi.org/10.1038/s42255-023-00943-3
46. Tanabe H, Masuzaki H, Shimabukuro M (2021) Novel strategies for glycaemic control and preventing diabetic complications applying the clustering-based classification of adult-onset diabetes mellitus: a perspective. Diabetes Res Clin Pract 180:109067. https://doi.org/10.1016/j.diabres.2021.109067
47. Dennis JM (2020) Precision medicine in type 2 diabetes: using individualized prediction models to optimize selection of treatment. Diabetes 69(10):2075–2085. https://doi.org/10.2337/dbi20-0002
48. Dennis JM, Young KG, McGovern AP et al (2022) Development of a treatment selection algorithm for SGLT2 and DPP-4 inhibitor therapies in people with type 2 diabetes: a retrospective cohort study. Lancet Digital Health 4(12):e873–e883. https://doi.org/10.1016/S2589-7500(22)00174-1
49. Cardoso P, Young KG, Nair AT et al (2024) Phenotype-based targeted treatment of SGLT2 inhibitors and GLP-1 receptor agonists in type 2 diabetes. Diabetologia 67(5):822–836. https://doi.org/10.1007/s00125-024-06099-3
50. Venkatasubramaniam A, Mateen BA, Shields BM et al (2023) Comparison of causal forest and regression-based approaches to evaluate treatment effect heterogeneity: an application for type 2 diabetes precision medicine. BMC Med Inform Decis Mak 23(1):110. https://doi.org/10.1186/s12911-023-02207-2
51. Dennis JM, Young KG, Cardoso P et al (2025) A five-drug class model using routinely available clinical features to optimise prescribing in type 2 diabetes: a prediction model development and validation study. Lancet 405(10480):701–714. https://doi.org/10.1016/S0140-6736(24)02617-5
52. Shields BM, Dennis JM, Angwin CD et al (2023) Patient stratification for determining optimal second-line and third-line therapy

for type 2 diabetes: the TriMaster study. Nat Med 29(2):376–383. https://doi.org/10.1038/s41591-022-02120-7
53. Seidu S, Ikhile D (2025) Personalised type 2 diabetes therapy in primary care. Lancet 405(10480):672–674. https://doi.org/10.1016/S0140-6736(24)02844-7
54. Van Calster B, van Smeden M, van Amsterdam W, Coemans M, Wynants L, Steyerberg EW (2026) The enemies of reliable and useful clinical prediction models: a review of statistical and scientific challenges. Annu Rev Stat Appl 13:465–492. https://doi.org/10.1146/annurev-statistics-042324-123749
55. Wilkinson J, Arnold KF, Murray EJ et al (2020) Time to reality check the promises of machine learning-powered precision medicine. Lancet Digital Health 2(12):e677–e680. https://doi.org/10.1016/S2589-7500(20)30200-4
56. Güdemann LM, Young KG, Cardoso P et al (2026) Validation of an algorithm for selection of SGLT2 and DPP4 inhibitor therapies in people with type 2 diabetes across major UK ethnicity groups: a retrospective cohort study. Lancet Reg Health – Europe 61. https://doi.org/10.1016/j.lanepe.2025.101547
57. Van Calster B, Steyerberg EW, Wynants L, van Smeden M (2023) There is no such thing as a validated prediction model. BMC Med 21(1):70. https://doi.org/10.1186/s12916-023-02779-w
58. Van Calster B, McLernon DJ, van Smeden M et al (2019) Calibration: the Achilles heel of predictive analytics. BMC Med 17(1):230. https://doi.org/10.1186/s12916-019-1466-7
59. Ahlqvist E, Prasad RB, Tuomi T, Rosengren A, Groop L (2018) Novel diabetes subgroups - authors' reply. Lancet Diabetes Endocrinol 6(6):440–441. https://doi.org/10.1016/S2213-8587(18)30139-6
60. Zaharia O-P, Kupriyanova Y, Bobrov P et al (2025) Role of adipose tissue compartments for cardiovascular risk in diabetes endotypes. Diabetes Care 48(12):2031–2035. https://doi.org/10.2337/dc24-2023
61. Deutsch AJ, Ahlqvist E, Udler MS (2022) Phenotypic and genetic classification of diabetes. Diabetologia 65(11):1758–1769. https://doi.org/10.1007/s00125-022-05769-4
62. Mansour Aly D, Dwivedi OP, Prasad RB et al (2021) Genome-wide association analyses highlight etiological differences underlying newly defined subtypes of diabetes. Nat Genet 53(11):1534–1542. https://doi.org/10.1038/s41588-021-00948-2
63. Mori T, Prystupa K, Straßburger K et al (2025) A web-based application for diabetes subtyping: The DDZ Diabetes-Cluster-Tool. Acta Diabetol 62(2):281–283. https://doi.org/10.1007/s00592-024-02436-5
64. 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):e001506. https://doi.org/10.1136/bmjdrc-2020-001506
65. Dennis JM, Shields BM, Henley WE, Jones AG, Hattersley AT (2019) Clusters provide a better holistic view of type 2 diabetes than simple clinical features - authors' reply. Lancet Diabetes Endocrinol 7(9):669. https://doi.org/10.1016/S2213-8587(19)30250-5
66. Hastie T, Tibshirani R, Friedman J (2009) Unsupervised learning. In: The elements of statistical learning: data mining, inference, and prediction. Vol 2. Springer, New York, pp 485–586
67. SCORE2-Diabetes Working Group and the ESC Cardiovascular Risk Collaboration (2023) SCORE2-Diabetes: 10-year cardiovascular risk estimation in type 2 diabetes in Europe. Eur Heart J 44(28):2544–2556. https://doi.org/10.1093/eurheartj/ehad260
68. Wagner R, Heni M, Tabak AG et al (2021) Pathophysiology-based subphenotyping of individuals at elevated risk for type 2 diabetes. Nat Med 27(1):49–57. https://doi.org/10.1038/s41591-020-1116-9
69. Wagner R, Selvin E, Sehgal R et al (2026) Beyond glucose—rethinking prediabetes for precision prevention. Diabetes Care 49(2):226–235. https://doi.org/10.2337/dci25-0054
70. Huemer M-T, Spagnuolo MC, Maalmi H et al (2025) Phenotype-based clusters, inflammation and cardiometabolic complications in older people before the diagnosis of type 2 diabetes: KORA F4/FF4 cohort study. Cardiovasc Diabetol 24(1):83. https://doi.org/10.1186/s12933-025-02617-8
71. Gale EAM (2006) Declassifying diabetes. Diabetologia 49(9):1989–1995. https://doi.org/10.1007/s00125-006-0348-7
72. Kuss O, Opitz ME, Brandstetter LV, Schlesinger S, Roden M, Hoyer A (2023) How amenable is type 2 diabetes treatment for precision diabetology? A meta-regression of glycaemic control data from 174 randomised trials. Diabetologia 66(9):1622–1632. https://doi.org/10.1007/s00125-023-05951-2
73. Gewandter JS, McDermott MP, He H et al (2019) Demonstrating heterogeneity of treatment effects among patients: an overlooked but important step toward precision medicine. Clin Pharmacol Ther 106(1):204–210. https://doi.org/10.1002/cpt.1372
74. Dworkin RH, McDermott MP, Farrar JT, O'Connor AB, Senn S (2014) Interpreting patient treatment response in analgesic clinical trials: implications for genotyping, phenotyping, and personalized pain treatment. Pain 155(3):457–460. https://doi.org/10.1016/j.pain.2013.09.019
75. Senn S (2016) Mastering variation: variance components and personalised medicine. Stat Med 35(7):966–977. https://doi.org/10.1002/sim.6739
76. Senn S (2018) Statistical pitfalls of personalized medicine. Nature 563(7733):619–621. https://doi.org/10.1038/d41586-018-07535-2
77. Kuss O, Roden M, Schlesinger S, Hoyer A (2024) The potential of precision diabetology for type 2 diabetes treatment—evidence from a meta-regression for all-cause mortality from large cardiovascular outcome trials. Acta Diabetol 62(7):1073–1080. https://doi.org/10.1007/s00592-024-02425-8
78. Sundström J, Lind L, Nowrouzi S et al (2023) Heterogeneity in blood pressure response to 4 antihypertensive drugs: a randomized clinical trial. JAMA 329(14):1160–1169. https://doi.org/10.1001/jama.2023.3322
79. Griffin S (2022) Diabetes precision medicine: plenty of potential, pitfalls and perils but not yet ready for prime time. Diabetologia 65(11):1913–1921. https://doi.org/10.1007/s00125-022-05782-7
80. Lim SS, Semnani-Azad Z, Morieri ML et al (2024) Reporting guidelines for precision medicine research of clinical relevance: the BePRECISE checklist. Nat Med 30(7):1874–1881. https://doi.org/10.1038/s41591-024-03033-3
81. Mori T (2025) Diabetes subtyping for precision medicine – from practical implementation to methodological challenges (Doctoral dissertation). Available from https://katalog.ulb.hhu.de/Record/9947786537706443. Accessed 8 April 2026

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