Precision medicine in type 2 diabetes: Using individualised prediction models to optimise selection of treatment

Author: John M Dennis *

Author details:

1. John M Dennis (PhD). University of Exeter Medical School. Address: Institute of Biomedical & Clinical Science, RILD Building, Royal Devon & Exeter Hospital, Barrack Road, Exeter EX2 5DW, UK

*Correspondence to:
John Dennis: Email j.dennis@exeter.ac.uk, phone +44 7734 940921

Word count: 3960

Figures: 7

Tweet:

NEW ARTICLE: Using simple clinical features to optimise selection of #Type2Diabetes glucose-lowering treatment. A framework for a practical #precisionmedicine approach using existing routine clinical and trial data @john_den_ @ExeterMed #ExeterDiabetes

Please include Figure 6B to summarise the article – just the Figure without the advantages and disadvantages table

Abstract

Despite the known heterogeneity of type 2 diabetes, and variable response to glucose lowering medications, current evidence on optimal treatment is predominantly based on average effects in clinical trials rather than individual-level characteristics. A precision medicine approach based on treatment response would aim to improve on this by identifying predictors of differential drug response for people based on their characteristics, and then using this information to select optimal treatment. Recent research has demonstrated robust and clinically relevant differential drug response with all non-insulin treatments after metformin (sulfonylureas, thiazolidinediones, DPP4 inhibitors, GLP-1 receptor agonists and SGLT2 inhibitors) using routinely available clinical features. This *Perspective* reviews this current evidence, and discusses how differences in drug response could inform selection of optimal type 2 diabetes treatment in the near future. It presents a novel framework for developing and testing precision medicine based strategies to optimise treatment, harnessing existing routine clinical and trial data sources. This framework was recently applied to demonstrate that 'subtype' approaches, in which people are classified into subgroups based on features reflecting underlying pathophysiology, are likely to have less clinical utility compared to approaches that combine the same features as continuous measures in probabilistic 'individualised prediction' models.

Introduction

Type 2 diabetes is a complex disease, characterised by hyperglycaemia associated with varying degrees of insulin resistance and impaired insulin secretion, and influenced by non-genetic and genetic factors. Despite this, glucose-lowering treatment is similar for most people. Current type 2 diabetes guidelines recommend the choice between glucose-lowering treatment options is based on clinical characteristics,(1) an approach in-line with the central goal of precision medicine, the tailoring of medical treatment to an individual. After initial metformin, the most recent guidelines recommend glucagon-like peptide-1 receptor agonists (GLP-1 RA) or sodium-glucose cotransporter-2 inhibitors (SGLT2i) in people with established atherosclerotic cardiovascular disease, heart failure, or chronic kidney disease, but this stratification only applies to up to 15-20% of people with type 2 diabetes.(2, 3) For the remaining majority, evidence of benefit beyond glucose-lowering with these drug classes has not been robustly demonstrated, and the optimal treatment pathway is not clear.(1) Evidence on the key considerations, notably glucose-lowering efficacy, tolerability, and side-effects is mainly derived from average treatment effects from clinical trials. This means there is little information available on whether a specific person in the clinic is more or less likely than the average trial participant to respond well to a particular treatment, or develop side-effects. Given this knowledge gap, there is currently great interest in developing approaches that can characterise people beyond the standard type 2 diabetes phenotype, and use this heterogeneity to optimise the selection of glucose-lowering treatment.

Any successful implementation of precision medicine in type 2 diabetes is likely to be very different from the most successful examples of precision medicine to-date. These have been in cancer and single gene diseases such as monogenic diabetes, where expensive genetic testing defines the aetiology, and the specific aetiology helps to determine treatment.(4, 5) In type 2 diabetes, unlike cancer, tissue is not available, and unlike rare forms of diabetes,

current genetic testing does not allow clear definition of the underlying pathophysiology.(6) This makes identification of discrete, non-overlapping subtypes of type 2 diabetes much less likely.(7)

In this perspective, I focus on a fundamental aim of precision medicine, the selection of optimal type 2 diabetes treatment based on likely differences in drug effect (henceforth, heterogeneity of treatment effect [HTE]). I provide an overview of the evidence from recent studies of HTE in type 2 diabetes and present a framework for using existing routine clinical and trial data sources to develop and test precision medicine based strategies to optimise treatment. The focus is on glycaemic response as nearly all current evidence of HTE for diabetes drugs is for differences in HbA1c. However, the framework outlined can easily be extended to evaluate HTE for non-glycaemic endpoints, including microvascular and macrovascular complications. Type 2 diabetes is a highly prevalent condition with relatively inexpensive treatment, meaning precision medicine approaches based on inexpensive markers have greatest potential to translate into clinical practice in the near future. As a result this article concentrates on the use of routinely available clinical features to select optimal treatment, although the principles discussed equally apply to the use of genomic or non-routine biomarkers.(6) Recent reviews of the pharmacogenomics of type 2 diabetes drug response are available elsewhere.(8, 9)

Why type 2 diabetes glucose-lowering treatment is an excellent candidate for a precision medicine approach

Type 2 diabetes treatment is an excellent candidate for a precision medicine approach for the following reasons: 1) There are many different drug classes available after metformin with different mechanisms of action but the same principal aim: to lower blood glucose; 2) At the individual-level, glucose-lowering response to each drug appears to vary greatly (**Figure 1**); 3) There is not a clear 'best' overall treatment outside a small proportion of individuals with specific complications. For the remainder, current treatment guidelines do not provide information on which drug class is best for lowering blood glucose, for which people(1); 4) There is great heterogeneity in the clinical phenotype of type 2 diabetes, making it plausible people with different underlying pathophysiology will vary in response to the different drug classes, depending on the mechanism of action of the drug.

Defining the treatment selection approach in type 2 diabetes

Despite the large biological noise in HbA1c, the majority of people appear to respond when initiated on a glucose-lowering drug (**Figure 1**), and it is unlikely many who appear not to responds are true 'non responders.'(10) Therefore the aim of precision medicine in type 2 diabetes is not to identify people who will and will not respond (which can only be achieved through repeated crossover trial designs(11, 12)), but instead to identify people who are likely to have a greater relative benefit on one drug class over another. This means that the necessary first step is to identify whether there are markers robustly predictive of greater or lesser response to each drug class to a clinically significant degree. In the absence of single markers with huge effect sizes, which have not been found to date, the second step is to optimally use multiple markers in combination to select treatment for individuals.

Identifying robust predictors of type 2 diabetes treatment response using routine and trial data

A focus on identifying routine clinical markers means HTE can be evaluated using existing observational and trial datasets that capture information on the drug response of people initiating type 2 diabetes treatment. The conventional approach is to examine HTE in clinical trials using "one-at-a-time" subgroup analysis in which participants are subcategorised based on a set of single characteristics in turn, such as sex and age (old vs. young). This approach does not provide credible evidence of differences in response due to low statistical power, lack of multivariable adjustment, and the risk of false-negative and false-positive findings.(13) This means very few 'discovered' positive subgroups are subsequently replicated.(14, 15)

While subgroup analysis of trials is limited, a combination of large observational routine clinical datasets and trial data (increasingly available(16, 17)) provides a powerful starting point to robustly evaluate HTE. Large anonymized routine clinical electronic health record databases, such as the UK's Clinical Practice Research Datalink,(18) provide a rich source of 'real-world' information on demographics, clinical features, diagnoses, laboratory tests, and prescriptions. One two-step approach to 'triangulate' routine and trial data sources is shown in **Figure 2**, on the basis that the best evidence for robust HTE is replication of effect in multiple independent datasets with differing strengths and weaknesses. In Step 1, due to the large sample size and availability of head-to-head data for all drug classes, routine clinical data are used for 'discovery' analysis, with assessment of drug-by-marker interactions to identify candidate features associated with differential response across drug classes. As in these observational data drug selection is not random and there are likely to be large

differences in baseline clinical features between treatment groups, careful identification of confounders and statistical adjustment is required. To further reduce bias, the use of causal inference methods such as inverse probability of treatment weighting,(19) or target trial approaches where studies are set up to emulate the design of an 'ideal' randomised trial, should be considered.(20) Nonetheless, unmeasured confounding may still bias findings, meaning a second step of external validation is required confirm findings. In Step 2, specific markers associated with potentially clinically relevant differences in drug response can be tested for reproducibility as pre-specified hypotheses in clinical trial datasets where treatment allocation is randomised and blinded, and there is systematic baseline assessment and follow-up, meaning the risk of confounding is much lower.(21) This two-step approach takes advantage of the larger, more heterogeneous, population in routine care datasets for feature discovery, whilst minimising the risk of data-mining in the smaller, richer, trial datasets.

What clinical features alter type 2 diabetes treatment response?

Recent studies have demonstrated clinically relevant differences in response by clinical features for all non-insulin glucose-lowering drug classes commonly used after metformin. Studies that do not adjust for baseline HbA1c are not reported here, given the demonstrated risk of false associations in such analysis.(22)

Sulfonylurea (SU) and thiazolidinedione (TZD) treatment: The first robust demonstration of HTE for type 2 diabetes therapy using the routine and trial data framework previous described. Observational data from United Kingdom (UK) primary care data was used as a discovery dataset, in which it was demonstrated that males without obesity (BMI <30) have on average a greater glucose lowering response with SU compared to TZD treatment, while, conversely, females with obesity (BMI ≥30) have a greater response to TZD than SU

treatment.(21) Differences in response in these subgroups were then validated, and confirmed to hold for long-term response, in randomised trial replication data, with differences in effect size within these subgroups equivalent to the addition of another glucose-lowering treatment (**Figure 3**).

Dipeptidyl peptidase 4 inhibitors (DPP4i) and GLP-1 receptor agonists: With DPP4i, the prospective PRIBA study demonstrated that markers of higher insulin resistance are consistently associated with lesser glucose-lowering response in the non-insulin treated.(23) Differences were clinically relevant; a subgroup defined by obesity (BMI≥30) and high triglycerides (≥2.3mmol/L) [31% of participants] had a response less than half that of a non-obese, low triglyceride (<2.3mmol/L) subgroup [22% of participants] (6 month response -5.3 mmol/mol [-0.5%] and -11.3 mmol/mol [-1.0%], respectively). Conversely, there was no evidence of an association between markers of insulin resistance and glucose-lowering response for non-insulin treated people initiating GLP-1 RA (**Figure 4**). Results were replicated in UK primary care data. Interestingly, in insulin-treated people but not non-insulin treated people, the same study found that with GLP-1 RA clinical markers of low beta cell function such as lower C-peptide and longer duration of diabetes were associated with reduced glucose-lowering efficacy.(24) With DPP4i, several other studies support the association between lower BMI, lower insulin resistance and greater response, and also suggest a benefit in glucose-lowering for people of Asian ethnicity.(25, 26)

SGLT2 inhibitors: Analysis of trial data have reported markedly greater relative benefit with SGLT2i at higher baseline HbA1c levels, compared to DPP4i or SU treatment.(27, 28) Differences in response with SGLT2i have also been observed by baseline renal function. Whilst the reduced efficacy of SGLT2i at eGFR's less than 60 is well-established,(29) pooled trial analysis has demonstrated this likely extends across the normal range, meaning that people with baseline eGFR >90 have a greater response compared to those with eGFR 60-

90.(30) In contrast, with DPP4i response is likely maintained in people at lower eGFRs.(31) Early work by our group suggests that these differential treatment effects for SGLT2i and DPP4i are replicated in UK primary care data (**Figure 5**).

Factors altering treatment response may relate to the underlying mechanism of action of different drug classes

The identified clinical features associated with HTE in many cases relate to the known mechanisms of action of the different drug classes. Such 'plausibility of effect modification' greatly strengthens the credibility of HTE analysis.(13) For TZD, as well as the increased insulin resistance with higher BMI, variation in response by sex and obesity is likely to reflect associated differences in adipocyte distribution and function as these drugs primarily act on adipose tissue.(32, 33) For SU and DPP4i, which stimulate insulin secretion by the β-call, the association between reduced insulin sensitivity and higher BMI possibly explains greater response in non-obese people. However, this does not explain the lack of association between insulin resistance and glucose lowering for the other incretin based drug class, GLP1-RA; it is possible this difference could relate to the added weight loss effects of this medication class, or that GLP-1RA response was studied in an almost entirely obese (and therefore insulin resistant) population.(23) The lack of GLP-1RA glycaemic benefit in insulin treated participants with very severe endogenous insulin deficiency is also consistent with the known role of potentiation of endogenous insulin secretion in their action. Effects on urinary glucose excretion provides a likely explanation for the variation in glucose-lowering efficacy of SGLT2i with baseline HbA1c and eGFR.(30, 31)

How can differences in treatment response inform selection of optimal treatment?

Whilst evidence of robust differences in type 2 diabetes treatment response is growing, there is current debate and considerable uncertainty about how to translate this to inform decision making in clinical practice. Recent literature has focused on two approaches (**Figure 6**):

1) A 'subtypes' approach, in which people with type 2 diabetes are subclassified based on their underlying pathophysiology (whether clinical, genetic, phenotypic or biomarker traits), on the assumption that once subtypes are defined they will have utility to stratify therapeutic decisions and other outcomes such as progression to complications. This was recently and notably proposed by Ahlqvist et al. in a sex-stratified data-driven cluster analysis of people close to diabetes diagnosis that grouped individuals with similar underlying pathophysiology using 5 clinical features (age at diagnosis, BMI, HbA1c, and HOMA-measured insulin resistance and insulin sensitivity) in Scandinavian registry data.(34) Importantly, similar looking subgroups were identified when the analysis was repeated in multiple international population-based cohorts.(35, 36) Subgroups showed differences in outcomes in observational follow-up, although differential treatment response was not assessed. Several other data-driven classifications have recent been proposed with substantial variation in the features used for classification and the numbers of subgroups identified,(37-39) including genetically defined clusters(40, 41), but their utility to stratify treatment response has similarly not been assessed.

2) To use a person's specific clinical information in a probabilistic 'individualised prediction' approach. In this approach markers reflecting underlying pathophysiology are used as continuous traits to directly predict an individual's treatment response for each drug. An individuals' specific information can then be used to predict their likely best drug in terms of glucose-lowering response (or alternatively to identify the absence of clinically relevant differences in response across treatments), and these predictions can guide selection of optimal treatment. The model developed is specific to the outcome of treatment response, and

can be deployed based on a person's current information at the point a decision to escalate treatment is made. Although subtypes could then in theory be specified based on the prediction of differential response or optimal therapy, this would make little sense as the subtypes would be based on clinical parameters that vary over time and are affected by treatment, meaning for an individual subtype assignment is unlikely to be stable. This proposed approach is consistent with the ideas underlying the recently proposed 'palette model' of diabetes,(7) which, at a specific point in time, positions an individual with diabetes on a spectrum of phenotypic variation and using this position to predict likely outcome.

Whilst the advantages and disadvantages of each approach in the context of selecting optimal treatment are shown in **Figure 6**, the fundamental difference between the two approaches is that the subtypes approach assumes homogeneity of differential treatment response for all individuals within a subtype, whereas the individualised prediction approach allows for estimation of differential treatment effects at the individual level. The use of individual level data means that the individualised prediction approach will almost certainly provide more precise estimates of treatment response, and thus more accurately guide optimal treatment selection, than approaches that lose information by classifying individual into subgroups.(42) The same principles will apply to prediction of any other outcome, for example predicting disease progression or development of microvascular and macrovascular complications.

Evaluating performance of strategies for selecting optimal treatment

Our group has recently applied a novel framework to evaluate treatment selection models in type 2 diabetes. Novel approaches are required as conventional measures of prediction model performance are of limited utility when evaluating treatment selection models,(13) as the focus is not the overall ability of a model to predict response, but rather accurate

identification of differences in response between treatments. At the individual level these differences are unobservable,(13) as at one point in time the response of a person to multiple different therapies cannot simultaneously be evaluated.

Our framework was applied to test head-to-head the Ahlqvist clusters strategy against an individualised prediction strategy for selecting optimal treatment, in post-hoc analysis of individual level data from two large clinical trials (ADOPT & RECORD, n=8,798).(43-45) This was important as a key discussion point raised in the Ahlqvist et al. study was that the clusters identified could be used to 'guide therapy.'(34) In both trials, participants were randomised to either SU, TZD, or metformin treatment. The same 5 subtypes proposed using the Scandinavian data were reproduced in ADOPT using the same data-driven cluster analysis approach.(34, 46) Then, within each subtype, average glycaemic response for each of the three treatments was estimated, and the treatment associated with the greatest average glycaemic response was allocated as the optimal treatment for all people assigned to that subtype. The utility of the subtypes was compared to an individualised prediction strategy that assigned optimal treatment on an individual rather than subtype level, using a model that estimated response for each drug for each participant based on their specific features. Notable, only the simple routine clinical features (sex, and BMI, HbA1c and age at diagnosis as continuous markers) were used for the individualised prediction model; two features used to inform the cluster analysis, HOMA-IR and HOMA-B (respectively, measures of insulin resistance and insulin secretion), were not included as they are not routinely available in clinical practice.

Despite including only simple markers, the individualised prediction strategy markedly outperformed the subtypes strategy in the external validation trial dataset (RECORD trial, n=4,057) (**Figure 7**).(43) For each strategy, the approach used was to define two subgroups of participants: 1) a concordant subgroup whose randomised treatment was the same as their

predicted optimal treatment; 2) a discordant subgroup whose randomised treatment differed to their predicted optimal treatment.(47) The difference between the concordant versus discordant subgroups was then contrasted for each strategy, with a bigger difference indicating a more useful treatment selection strategy. Where external test datasets are available, this evaluative framework represents a novel and cost-effective means of evaluating the utility of treatment selection models, whether on their own or in head-to-head comparison, and can be applied for other outcomes as well as treatment response.

Future directions: 'omics' and beyond HbA1c

Whilst this perspective has focused only on glycaemic response to diabetes treatment, the approaches outlined can easily be extended to non-glycaemic endpoints including microvascular and macrovascular complications. The ideal precision medicine approach in type 2 diabetes will maximise therapeutic benefit while limiting risks,(48) which will also require evaluation of HTE for side-effects, glycaemic progression, and risk of microvascular or macrovascular complications. Particular subgroups at higher risk of common treatment-specific side effects are already established for several drug classes, for example the risk of fracture with TZD is limited mainly to females.(49) Methods to overcome unmeasured confounding, such as the Prior Event Rate Ratio, may have particular utility for evaluating side-effect risk in observational routine care data where allocation to therapy is not randomised.(50, 51) A related but overlooked question for precision medicine, with great clinical relevance, is whether the benefits and risks of a treatment are positively associated. This is likely the case for TZD; the risk of oedema and likelihood of weight gain increase with greater glucose-lowering response,(21, 52) and this should be an important consideration when choosing treatment. A further extension of the current work would be

evaluation of effects of higher-order drug combinations. This will be possible in large routine clinical datasets where substantial numbers of patients are on specific combination therapies, although robust validation approaches will be required.

A key question is where genetics can add. Proposed genetically defined type 2 diabetes subtypes reflect and help to understand underlying pathophysiology.(40, 41) The clear advantage of using genetics is stability, as subtypes defined solely by genetics will be constant throughout life. At the moment it is unknown whether the continuous polygenic scores underlying genetic subtypes add over routine clinical features and biomarkers when predicting outcome and optimising treatment. For treatment response, individual genetic markers have shown differences for specific treatments, and may be of clinical utility when genetic information is routinely available in the medical records.(53, 54) If clinically relevant benefit can be demonstrated for polygenic scores and implementation is cost-effective, such scores can similarly be integrated into models based on routine clinical features.

A further exciting opportunity is the application of causal inference, data-driven machine learning and AI-based approaches to improve HTE prediction accuracy and generalisability of findings from large data sources such as electronic health records. Data-driven approaches may be of particular utility when databases start to incorporate high-dimensional genetic information.(55) One possibility is that individualised prediction models developed with standard statistical methods based on classical risk factors could be augmented with data-driven classification approaches, if data-driven approaches are able to improve prediction by capturing higher order complex traits missed by the standard methods.

Although existing data can be used to develop and test candidate type 2 diabetes precision medicine approaches, ultimately prospective trials, as done in cancer and monogenic diabetes,(4, 56, 57) will likely be needed to demonstrate clinical utility. TRIMASTER, an

ongoing 3-way crossover randomised trial due to report in May 2021, is one such study in type 2 diabetes (NCT02653209). TRIMASTER will directly test the hypotheses that simple subgroups defined by baseline BMI and eGFR alter response with DPP4i, SGLT2i and TZD treatment.(58) Not only will this provide the first prospective randomised evaluation of a precision medicine approach for glycaemic response, the 3-way crossover design will allow an "n of 1" analysis of patient preferences regarding the three treatments when they are tried in randomised order in blinded conditions. However, running prospective trials to test potential candidate factors one-at-a-time for personalisation is not a feasible, cost-effective, or efficient strategy. Future trials could instead test specific precision medicine algorithms based on multiple factors (potentially both clinical and genetic features), to test whether use of an algorithm results in improved outcomes for patients. One simple trial design for this would be to cluster randomise health centres (e.g. GP practices in the UK) to either receive or not receive an algorithm – comparing centres with and without the algorithm would enable evaluation of its effectiveness and efficacy. If two competing algorithms or strategies need to be tested this could be done using 3-way cluster randomisation.

A final key challenge is implementation of algorithms, which to ensure patient benefit should be not only effective but transparent, reproducible and ethically sound,(59) and which should be equally and freely accessible to all health professionals and patients. A type 2 diabetes treatment selection model would likely be most appropriately positioned within clinical practice software systems, so that it can be automatically populated with relevant clinical information from the electronic health record and function as a decision aid at the point of care. Development of software infrastructure that can utilise routinely collected health records to support delivery of such probabilistic algorithms will be required before precision medicine can truly become a reality for common diseases such as type 2 diabetes.

Conclusions

Recent demonstration of robust, clinically relevant differences in glycaemic response suggest a precision medicine approach to selecting optimal type 2 diabetes treatment will soon be possible. The most practical way to implement this in the near future will be to focus on routine clinical markers, and the most accurate approach will be integration of continuous features into individualised, probabilistic, prediction models that can be deployed at the point a decision to escalate treatment is made, rather than subtyping. Estimates of differences in treatment response can augment the limited existing stratification of people with type 2 diabetes based on cardiovascular and renal comorbidity, and will be applicable to everyone requiring glucose-lowering treatment. For people for whom differences in response between treatments are modest, this information is still important as it can facilitate selection of treatment based on other criteria. A framework of discovery in routine data, followed by replication and testing in existing clinical trial datasets, offers a low-cost and principled way to evaluate the potential of precision medicine, applicable to other chronic diseases as well as type 2 diabetes.

Acknowledgments: The author is grateful to colleagues Dr Beverley Shields, Dr Angus Jones, Professor Andrew Hattersley, Dr Andrew McGovern and Dr Nick Thomas for insightful discussions and recommendations when preparing this article. The author also thanks the MASTERMIND consortium who have contributed to many of the studies discussed in this article. Data for the ADOPT trial were accessed through the Clinical Trial Data Transparency Portal under approval from GSK (Proposal 930). This article is based in part on data from the Clinical Practice Research Datalink obtained under licence from the UK Medicines and Healthcare products Regulatory Agency. CPRD data is provided by patients and collected by the NHS as part of their care and support. The author is supported by the NIHR Exeter Clinical Research Facility. The interpretation and conclusions contained in this study are those of the author alone.

Author contributions: J.D. is the sole author and guarantor of the manuscript.

Conflict of interest: No potential conflicts of interest relevant to this article were reported.

Funding: J.D. is supported by an Independent Fellowship funded by Research England's Expanding Excellence in England (E3) fund. Research discussed is this article was supported by the Medical Research Council (UK) (MR/N00633X/1).

References

1. Davies MJ, D'Alessio DA, Fradkin J, Kernan WN, Mathieu C, Mingrone G, et al. Management of Hyperglycemia in Type 2 Diabetes, 2018. A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Care. 2018;41(12):2669-701.
2. McGovern A, Feher M, Munro N, de Lusignan S. Sodium-Glucose Co-transporter 2 (SGLT2) Inhibitor: Comparing Trial Data and Real-World Use. Diabetes Ther. 2017;8(2):365-76.
3. Hinton W, Feher M, Munro N, Walker M, de Lusignan S. Real-world prevalence of the inclusion criteria for the LEADER trial: Data from a national general practice network. Diabetes, Obesity and Metabolism. 2019;21(7):1661-7.
4. Pearson ER, Flechtner I, Njolstad PR, Malecki MT, Flanagan SE, Larkin B, et al. Switching from insulin to oral sulfonylureas in patients with diabetes due to Kir6.2 mutations. N Engl J Med. 2006;355(5):467-77.
5. Moscow JA, Fojo T, Schilsky RL. The evidence framework for precision cancer medicine. Nat Rev Clin Oncol. 2018;15(3):183-92.
6. Hattersley AT, Patel KA. Precision diabetes: learning from monogenic diabetes. Diabetologia. 2017;60(5):769-77.
7. McCarthy MI. Painting a new picture of personalised medicine for diabetes. Diabetologia. 2017;60(5):793-9.
8. Pearson ER. Diabetes: Is There a Future for Pharmacogenomics Guided Treatment? Clinical Pharmacology & Therapeutics. 2019;106(2):329-37.
9. Gloyn AL, Drucker DJ. Precision medicine in the management of type 2 diabetes. Lancet Diabetes Endocrinol. 2018;6(11):891-900.
10. McGovern AP, Dennis JM, Shields BM, Hattersley AT, Pearson ER, Jones AG, et al. What to do with diabetes therapies when HbA1c lowering is inadequate: add, switch, or continue? A MASTERMIND study. BMC Medicine. 2019;17(1):79.
11. Senn S. Mastering variation: variance components and personalised medicine. Statistics in medicine. 2016;35(7):966-77.
12. Senn S. Statistical pitfalls of personalized medicine. Nature. 2018;563(7733):619-21.
13. Kent DM, Paulus JK, van Klaveren D, D'Agostino R, Goodman S, Hayward R, et al. The Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement. Ann Intern Med. 2020;172(1):35-45.
14. Wallach JD, Sullivan PG, Trepanowski JF, Sainani KL, Steyerberg EW, Ioannidis JP. Evaluation of Evidence of Statistical Support and Corroboration of Subgroup Claims in Randomized Clinical Trials. JAMA Intern Med. 2017;177(4):554-60.
15. Kent DM, Steyerberg E, van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. BMJ. 2018;363:k4245.
16. The YODA Project [Internet] [cited 14 Feb 2017]. Available from: http://yoda.yale.edu/.
17. Clinical Study Data Request [Internet] [cited 14 July 2017]. Available from: https://clinicalstudydatarequest.com/.
18. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol. 2015;44(3):827-36.
19. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med. 2015;34(28):3661-79.
20. Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183(8):758-64.
21. Dennis JM, Henley WE, Weedon MN, Lonergan M, Rodgers LR, Jones AG, et al. Sex and BMI Alter the Benefits and Risks of Sulfonylureas and Thiazolidinediones in Type 2 Diabetes: A Framework for Evaluating Stratification Using Routine Clinical and Individual Trial Data. Diabetes Care. 2018;41(9):1844-53.

22. Jones AG, Lonergan M, Henley WE, Pearson ER, Hattersley AT, Shields BM. Should Studies of Diabetes Treatment Stratification Correct for Baseline HbA1c? PLoS One. 2016;11(4):e0152428.
23. Dennis JM, Shields BM, Hill AV, Knight BA, McDonald TJ, Rodgers LR, et al. Precision Medicine in Type 2 Diabetes: Clinical Markers of Insulin Resistance Are Associated With Altered Short- and Long-term Glycemic Response to DPP-4 Inhibitor Therapy. Diabetes Care. 2018;41(4):705-12.
24. Jones AG, McDonald TJ, Shields BM, Hill AV, Hyde CJ, Knight BA, et al. Markers of β-Cell Failure Predict Poor Glycemic Response to GLP-1 Receptor Agonist Therapy in Type 2 Diabetes. Diabetes Care. 2016;39(2):250-7.
25. Bihan H, Ng WL, Magliano DJ, Shaw JE. Predictors of efficacy of GLP-1 agonists and DPP-4 inhibitors: A systematic review. Diabetes Res Clin Pract. 2016;121:27-34.
26. Davis TME, Mulder H, Lokhnygina Y, Aschner P, Chuang L-M, Raffo Grado CA, et al. Effect of race on the glycaemic response to sitagliptin: Insights from the Trial Evaluating Cardiovascular Outcomes with Sitagliptin (TECOS). Diabetes, Obesity and Metabolism. 2018;20(6):1427-34.
27. DeFronzo RA, Ferrannini E, Schernthaner G, Hantel S, Elsasser U, Lee C, et al. Slope of change in HbA1c from baseline with empagliflozin compared with sitagliptin or glimepiride in patients with type 2 diabetes. Endocrinology, Diabetes & Metabolism. 2018;1(2):e00016.
28. Rosenstock J, Hansen L, Zee P, Li Y, Cook W, Hirshberg B, et al. Dual add-on therapy in type 2 diabetes poorly controlled with metformin monotherapy: a randomized double-blind trial of saxagliptin plus dapagliflozin addition versus single addition of saxagliptin or dapagliflozin to metformin. Diabetes Care. 2015;38(3):376-83.
29. Macha S, Mattheus M, Halabi A, Pinnetti S, Woerle HJ, Broedl UC. Pharmacokinetics, pharmacodynamics and safety of empagliflozin, a sodium glucose cotransporter 2 (SGLT2) inhibitor, in subjects with renal impairment. Diabetes Obes Metab. 2014;16(3):215-22.
30. Cherney DZI, Cooper ME, Tikkanen I, Pfarr E, Johansen OE, Woerle HJ, et al. Pooled analysis of Phase III trials indicate contrasting influences of renal function on blood pressure, body weight, and HbA1c reductions with empagliflozin. Kidney Int. 2018;93(1):231-44.
31. Goldenberg RM. Choosing Dipeptidyl Peptidase-4 Inhibitors, Sodium-glucose Cotransporter-2 Inhibitors, or Both, as Add-ons to Metformin: Patient Baseline Characteristics Are Crucial. Clinical Therapeutics. 2017;39(12):2438-47.
32. Nesto RW, Bell D, Bonow RO, Fonseca V, Grundy SM, Horton ES, et al. Thiazolidinedione use, fluid retention, and congestive heart failure: a consensus statement from the American Heart Association and American Diabetes Association. Diabetes Care. 2004;27(1):256-63.
33. Fonseca V. Effect of thiazolidinediones on body weight in patients with diabetes mellitus. Am J Med. 2003;115 Suppl 8A:42S-8S.
34. 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. The Lancet Diabetes & Endocrinology. 2018;6(5):361-9.
35. 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. The Lancet Diabetes & Endocrinology. 2019;7(9):684-94.
36. Zou X, Zhou X, Zhu Z, Ji L. Novel subgroups of patients with adult-onset diabetes in Chinese and US populations. The Lancet Diabetes & Endocrinology. 2019;7(1):9-11.
37. CHRISTIAENS A, HERMANS MP, BOLAND B, HENRARD S. Distinction of cardiometabolic profiles among people ≥75 years with type 2 diabetes: a latent profile analysis. BMC Endocrine Disorders. 2019;19(1):85.
38. Bancks MP, Casanova R, Gregg EW, Bertoni AG. Epidemiology of diabetes phenotypes and prevalent cardiovascular risk factors and diabetes complications in the National Health and Nutrition Examination Survey 2003–2014. Diabetes Research and Clinical Practice. 2019;158:107915.
39. Safai N, Ali A, Rossing P, Ridderstråle M. Stratification of type 2 diabetes based on routine clinical markers. Diabetes Research and Clinical Practice. 2018;141:275-83.

40. Udler MS, Kim J, von Grotthuss M, Bonàs-Guarch S, Cole JB, Chiou J, et al. Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: A soft clustering analysis. PLoS medicine. 2018;15(9):e1002654-e.
41. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nature Genetics. 2018;50(11):1505-13.
42. van Smeden M, Harrell FE, Jr., Dahly DL. Novel diabetes subgroups. The Lancet Diabetes & Endocrinology. 2018;6(6):439-40.
43. 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. The Lancet Diabetes & Endocrinology. 2019;7(6):442-51.
44. Kahn SE, Haffner SM, Heise MA, Herman WH, Holman RR, Jones NP, et al. Glycemic durability of rosiglitazone, metformin, or glyburide monotherapy. N Engl J Med. 2006;355(23):2427-43.
45. Home PD, Pocock SJ, Beck-Nielsen H, Curtis PS, Gomis R, Hanefeld M, et al. Rosiglitazone evaluated for cardiovascular outcomes in oral agent combination therapy for type 2 diabetes (RECORD): a multicentre, randomised, open-label trial. Lancet. 2009;373(9681):2125-35.
46. Lynam A, McDonald T, Hill A, Dennis J, Oram R, Pearson E, et al. Development and validation of multivariable clinical diagnostic models to identify type 1 diabetes requiring rapid insulin therapy in adults aged 18–50 years. BMJ Open. 2019;9(9):e031586.
47. Janes H, Pepe MS, Bossuyt PM, Barlow WE. Measuring the performance of markers for guiding treatment decisions. Ann Intern Med. 2011;154(4):253-9.
48. Riddle MC, Gerstein HC, Holman RR, Inzucchi SE, Zinman B, Zoungas S, et al. A1C Targets Should Be Personalized to Maximize Benefits While Limiting Risks. Diabetes Care. 2018;41(6):1121-4.
49. Zhu ZN, Jiang YF, Ding T. Risk of fracture with thiazolidinediones: an updated meta-analysis of randomized clinical trials. Bone. 2014;68:115-23.
50. Rodgers LR, Dennis JM, Shields BM, Mounce L, Fisher I, Hattersley AT, et al. Prior event rate ratio adjustment produced estimates consistent with randomized trial: a diabetes case study. J Clin Epidemiol. 2020;122:78-86.
51. Lin NX, Henley WE. Prior event rate ratio adjustment for hidden confounding in observational studies of treatment effectiveness: a pairwise Cox likelihood approach. Stat Med. 2016;35(28):5149-69.
52. Dennis JM, Shields BM, Jones AG, Pearson ER, Hattersley AT, Henley WE. Evaluating associations between the benefits and risks of drug therapy in type 2 diabetes: a joint modeling approach. Clin Epidemiol. 2018;10:1869-77.
53. Klonoff DC, Florez JC, German M, Fleming A. The Need for Precision Medicine to be Applied to Diabetes. J Diabetes Sci Technol. 2020:1932296819894295.
54. Mannino GC, Andreozzi F, Sesti G. Pharmacogenetics of type 2 diabetes mellitus, the route toward tailored medicine. Diabetes Metab Res Rev. 2019;35(3):e3109.
55. Kalina J. Classification methods for high-dimensional genetic data. Biocybernetics and Biomedical Engineering. 2014;34(1):10-8.
56. Pearson ER, Starkey BJ, Powell RJ, Gribble FM, Clark PM, Hattersley AT. Genetic cause of hyperglycaemia and response to treatment in diabetes. Lancet. 2003;362(9392):1275-81.
57. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, et al. Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non–Small-Cell Lung Cancer to Gefitinib. New England Journal of Medicine. 2004;350(21):2129-39.
58. TriMaster: Study of a DPP4 Inhibitor, SGLT2 Inhibitor and Thiazolidinedione as Third Line Therapy in Patients With Type 2 Diabetes. (TriMaster). (2016). Retrieved from http://clinicaltrials.gov/ct2 (Identification No. NCT02653209).

59. Vollmer S, Mateen BA, Bohner G, Király FJ, Ghani R, Jonsson P, et al. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. BMJ. 2020;368:l6927.

Figure legends

Figure 1: The marked individual-level variation in change in HbA1c at 6 months (6 month HbA1c minus baseline HbA1c) by drug in ADOPT trial for 3,707 participants with a valid measure of HbA1c at 6 months. A negative change is represented an improvement in HbA1c. Mean (standard deviation) improvement in HbA1c was greatest at 6 months for sulfonylureas -9.4 (8.6) mmol/mol [0.9%], compared to metformin -7.5 (8.1) mmol/mol [0.7%] and thiazolidinedione treatment -6.4 (8.6) mmol/mol [0.6%].

Figure 2: **A 'triangulation' approach using routine clinical and trial data to evaluate differences in drug response, and the strengths and weaknesses of each data source.**

Figure 3: Five-year glycaemic response (change from baseline in HbA1c) with thiazolidinedione (TZD) and sulfonylurea (SU) treatment in males without obesity (BMI<30) and females with obesity (BMI≥30) subgroups in 1,232 participants in the ADOPT clinical trial.(21) Data are presented as means ± standard errors at each study visit from mixed-effects models. A reduction (improvement) in HbA1c is represented as a negative value. For area-under-the-curve difference estimates (AUC), positive values favour SU, and negative values favour TZD. Adapted from Dennis et al.(21).

A) Males without obesity subgroup

B) Females with obesity subgroup

Figure 4: Associations between markers of insulin resistance and HbA1c response with DPP4 inhibitor and GLP-1 receptor agonist treatment in the PRIBA study (n=593), in participants not on insulin co-treatment. Estimates denote the mean HbA1c change (mmol/mol) at 6 months (Baseline HbA1c – 6 month HbA1c) per 1-SD higher baseline value of each marker. Associations were tested in a series of independent linear regression models adjusted for baseline HbA1c and co-treatment change. Error bars denote 95% confidence intervals. Adapted from Dennis et al.(23)

Figure 5: Associations between baseline HbA1c and baseline eGFR (CKD-EPI formula) and HbA1c response at 6 months (Baseline HbA1c – 6 month HbA1c) with SGLT2 inhibitor and DPP4 inhibitor treatment in UK primary care data (Clinical Practice Research Datalink) [n=20,965]. Results are predicted values from a linear regression model including baseline HbA1c-by-drug and eGFR-by-drug interaction terms (with each modelled as a restricted cubic spline with 3 knots), with additional adjustment for number of diabetes treatments ever initiated, number of current diabetes treatments, age at treatment, duration of diabetes, sex and BMI. Grey shading represents 95% confidence intervals.

A) Baseline HbA1c **B) Baseline eGFR**

Figure 6: Individualised prediction compared to classification into subtypes – advantages and disadvantages of two strategies to apply a precision medicine approach in type 2 diabetes

A) Classification into subtypes

B) Individualised prediction

Figure 7: Three-year glycaemic response (change from baseline in HbA1c) with concordant and discordant subgroups using the subtypes strategy and the individualised prediction strategy in the RECORD trial independent validation set (n=4,057). Each strategy was developed in the ADOPT trial (n=3,785). Data are presented as means ± 95% confidence intervals at each study visit from mixed-effects models. A reduction (improvement) in HbA1c is represented as a negative value. For AUC estimates, a more negative value represents a greater response. Adapted from Dennis et al. (43)

A) Subtypes strategy (left panel)

B) Individualised prediction strategy (right panel)

Figure 1

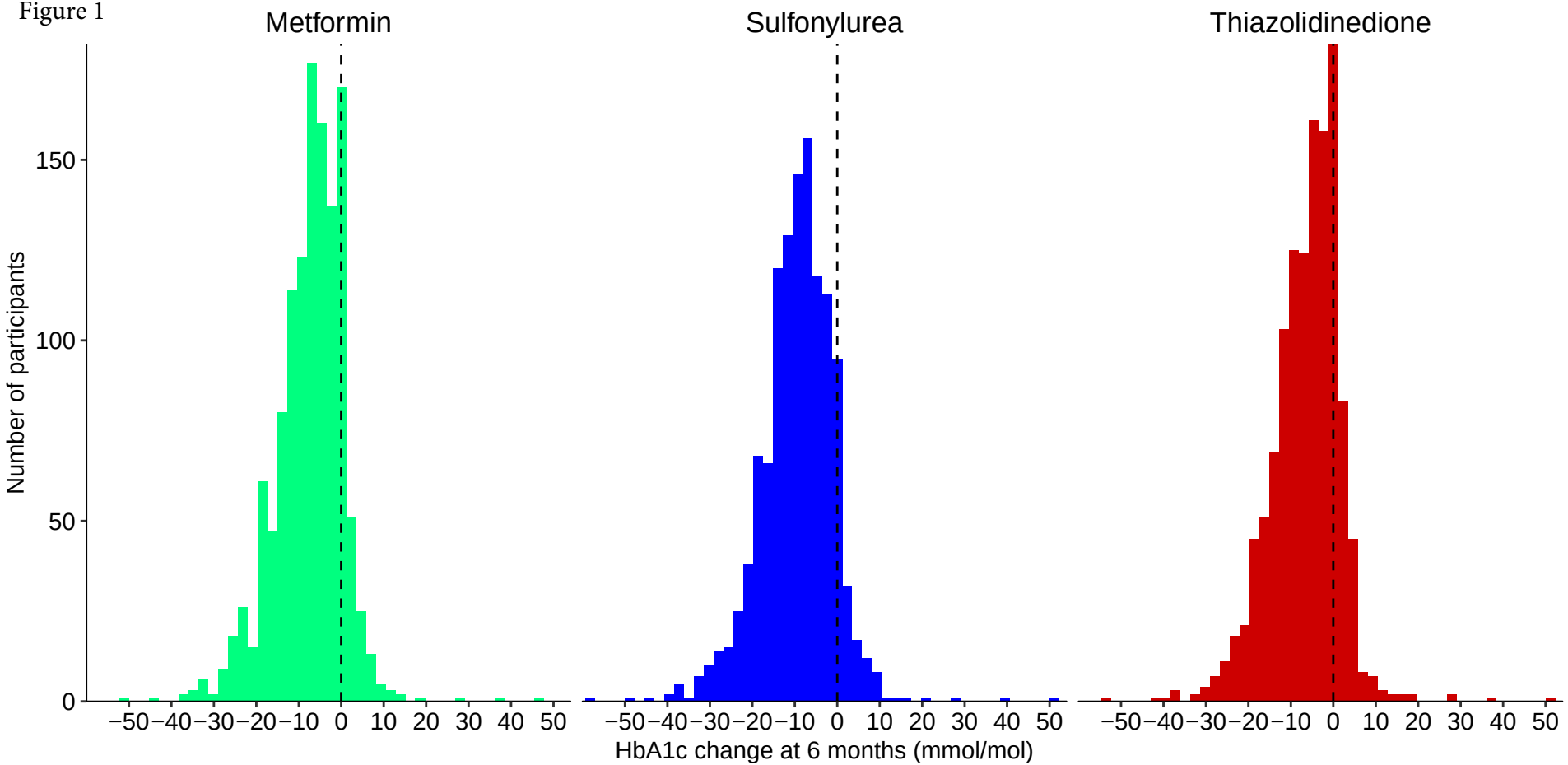
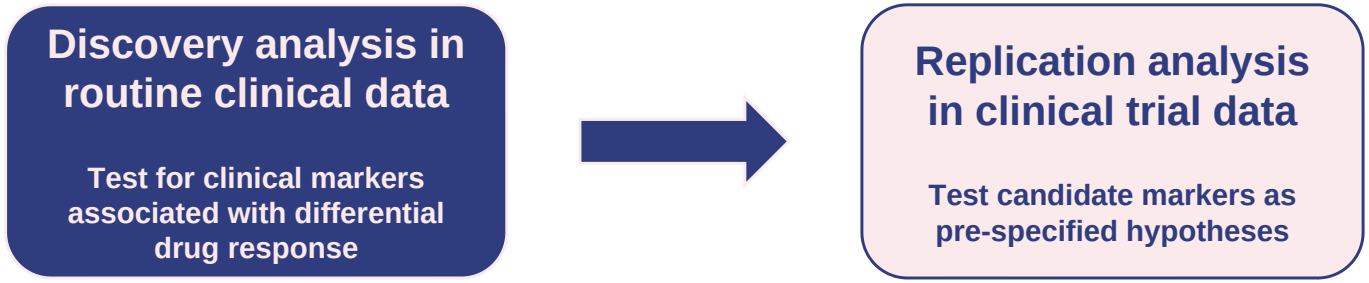
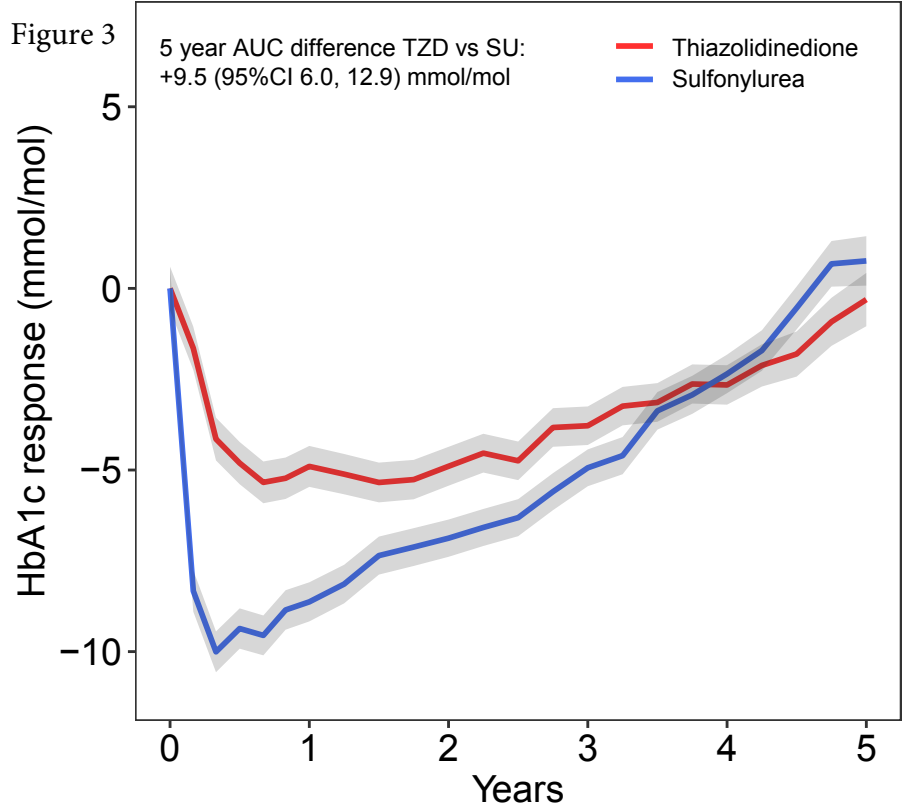

Metformin
Sulfonylurea
Thiazolidinedione
Number of participants
0
50
100
150
−50 −40 −30 −20 −10 0 10 20 30 40 50
HbA1c change at 6 months (mmol/mol)

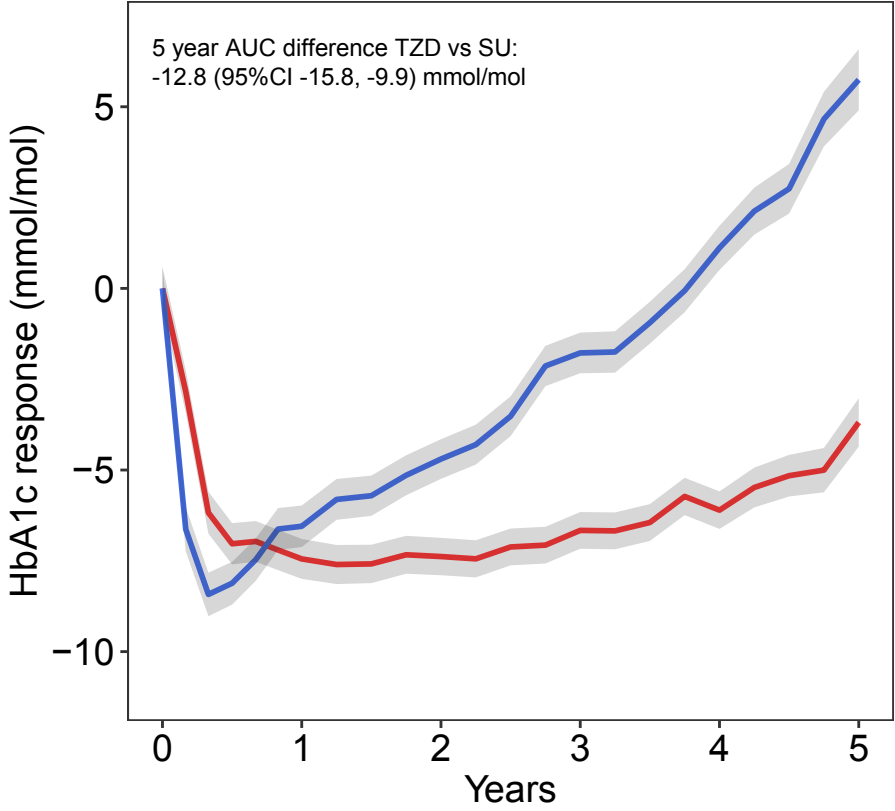
Figure 2



Routine clinical data	Clinical trial data
Strengths	**Strengths**
Large sample size meaning data on outcomes for thousands of people	Participants randomised and blinded to treatment received
Data on people initiating all available diabetes treatments for head-to-head comparison	Systematic baseline assessment of all participants
All people with type 2 diabetes can be included	Systematic follow-up and assessment of outcome
Only factors routinely available in clinical care are evaluated, ensuring any outputs can be integrated into practice at low cost	Non-routine markers available, which may have particular utility for study of underlying mechanisms
Weaknesses	**Weaknesses**
Treatment choice is not random and the reason underlying the treatment choice not available – high risk of bias	Smaller sample size typically means underpowered for subgroup analysis or evaluation of HTE (although individual data meta-analysis can be used to combine trial datasets[52])
Missing baseline information for many people, and the missingness may be informative and bias findings	Trial inclusion and exclusion criteria mean participants not representative of broader type 2 diabetes population
Outcome information only available if person returns for a primary care consultation	Typically placebo controlled or one comparator treatment only

Figure 3





TZD (n)	359	300	277	242	219	75
SU (n)	391	328	297	265	219	91

399	327	304	266	240	91
375	287	244	181	138	48

Figure 4

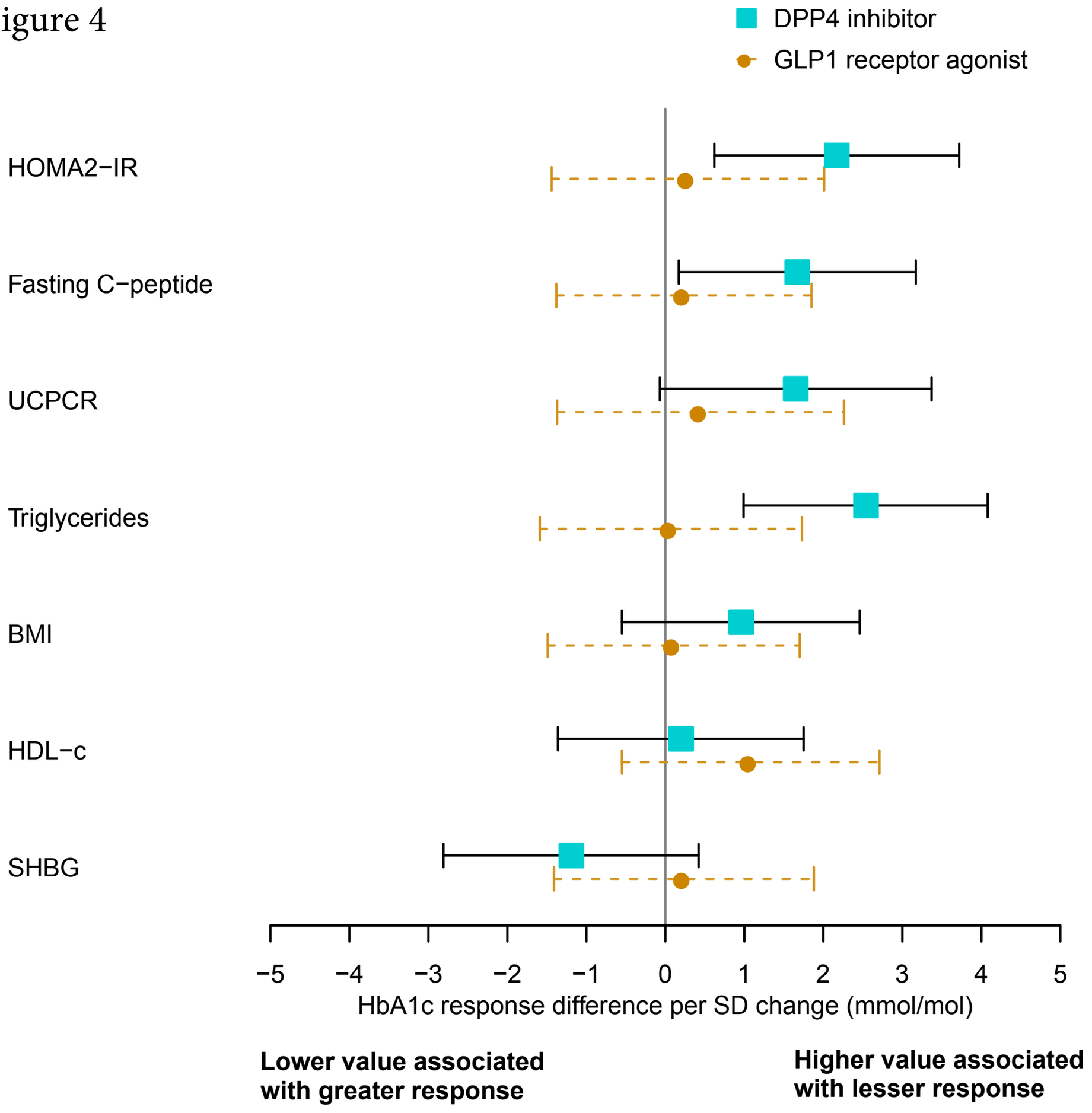

DPP4 inhibitor
GLP1 receptor agonist
HOMA2−IR
Fasting C−peptide
UCPCR
Triglycerides
BMI
HDL−c
SHBG
−5
−4
−3
−2
−1
0
1
2
3
4
5
HbA1c response difference per SD change (mmol/mol)
Lower value associated with greater response
Higher value associated with lesser response

Figure 5A

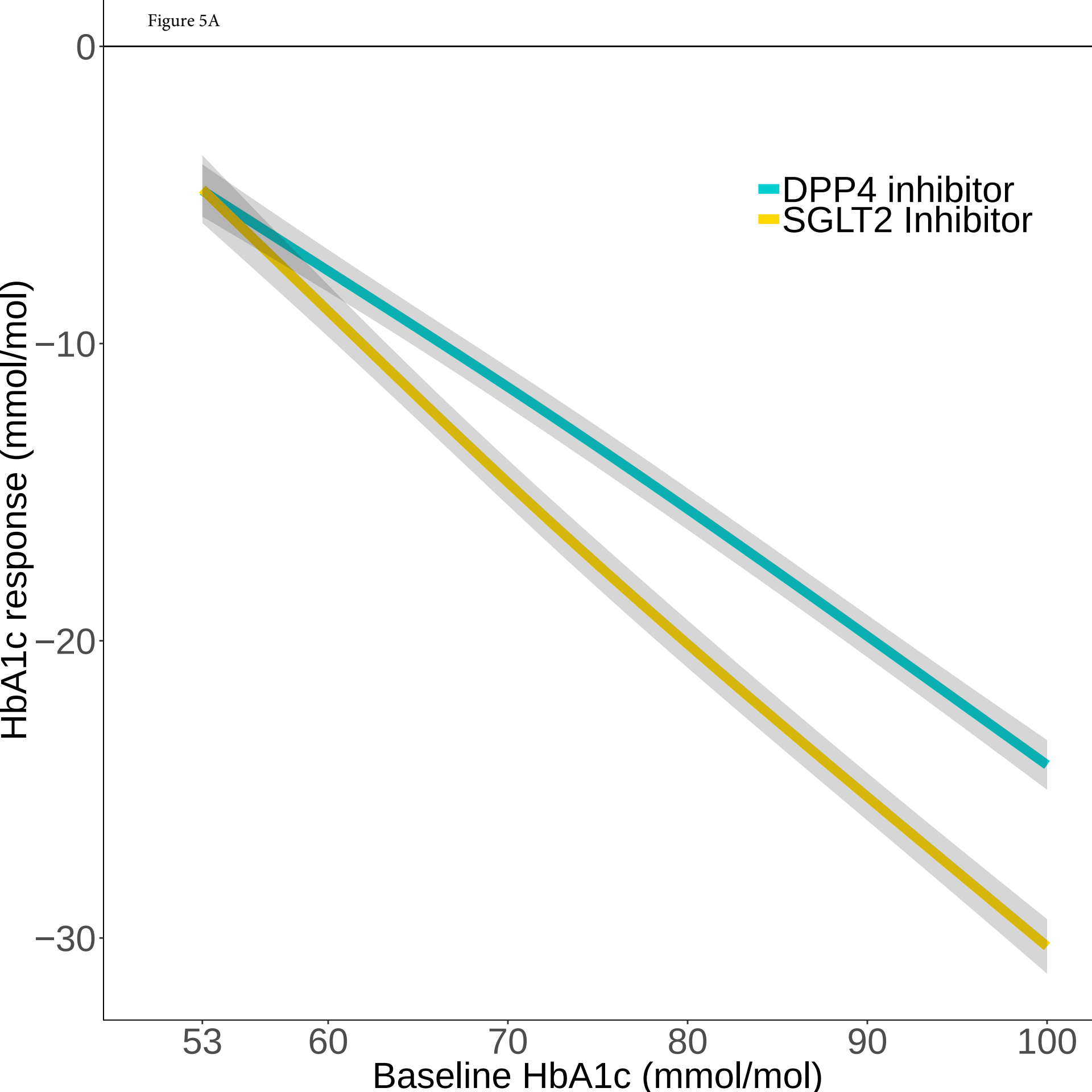

0
−10
−20
−30
HbA1c response (mmol/mol)
DPP4 inhibitor
SGLT2 Inhibitor
53
60
70
80
90
100
Baseline HbA1c (mmol/mol)

Figure 5B

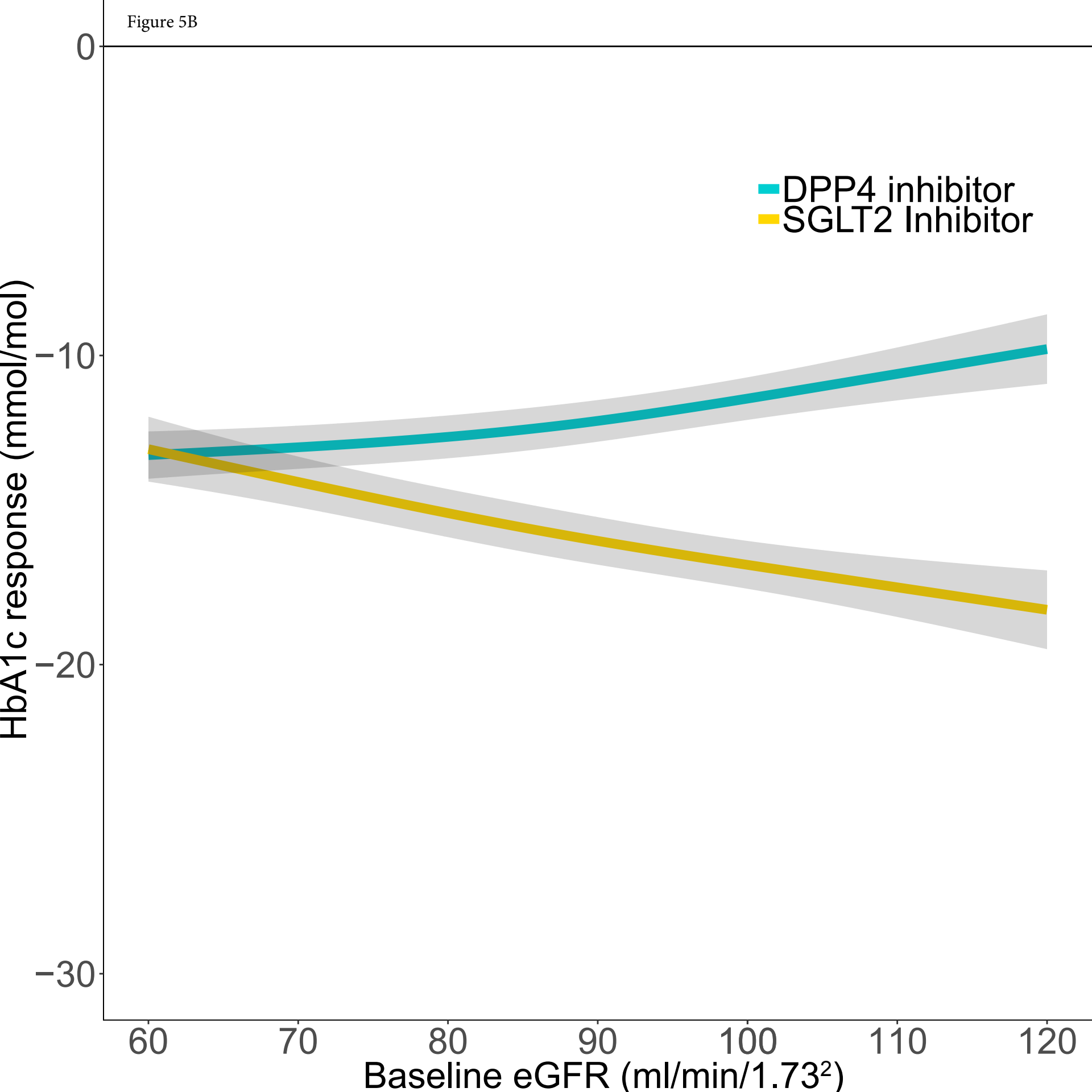
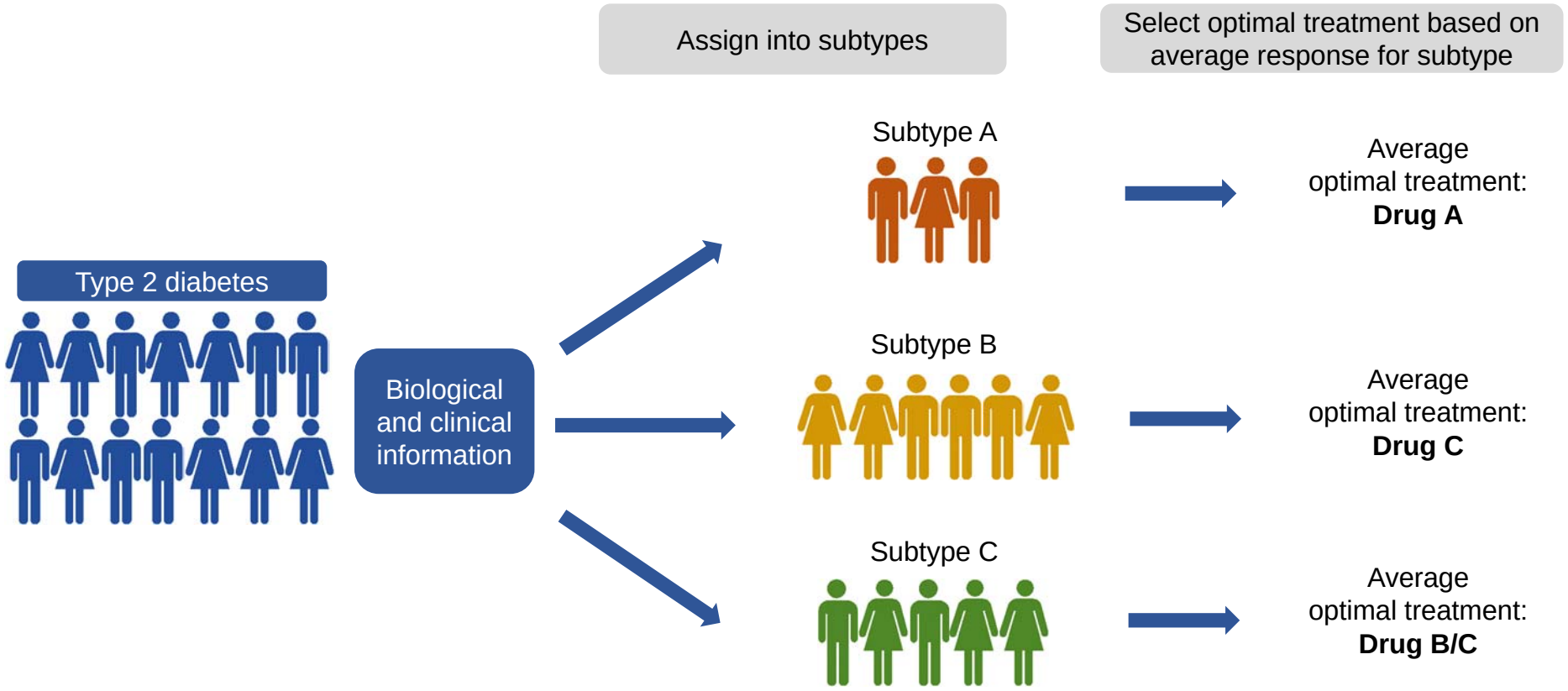

0
−10
−20
−30
HbA1c response (mmol/mol)
DPP4 inhibitor
SGLT2 Inhibitor
60
70
80
90
100
110
120
Baseline eGFR (ml/min/1.73²)

Figure 6A



Advantages

- Simple to communicate.

- Can assess risk of multiple outcomes based on subgroup assignment.
- Classification could take place at one time point only e.g. around diagnosis (important if non-routine testing is required).
- May enhance understanding of the pathophysiological basis for type 2 diabetes

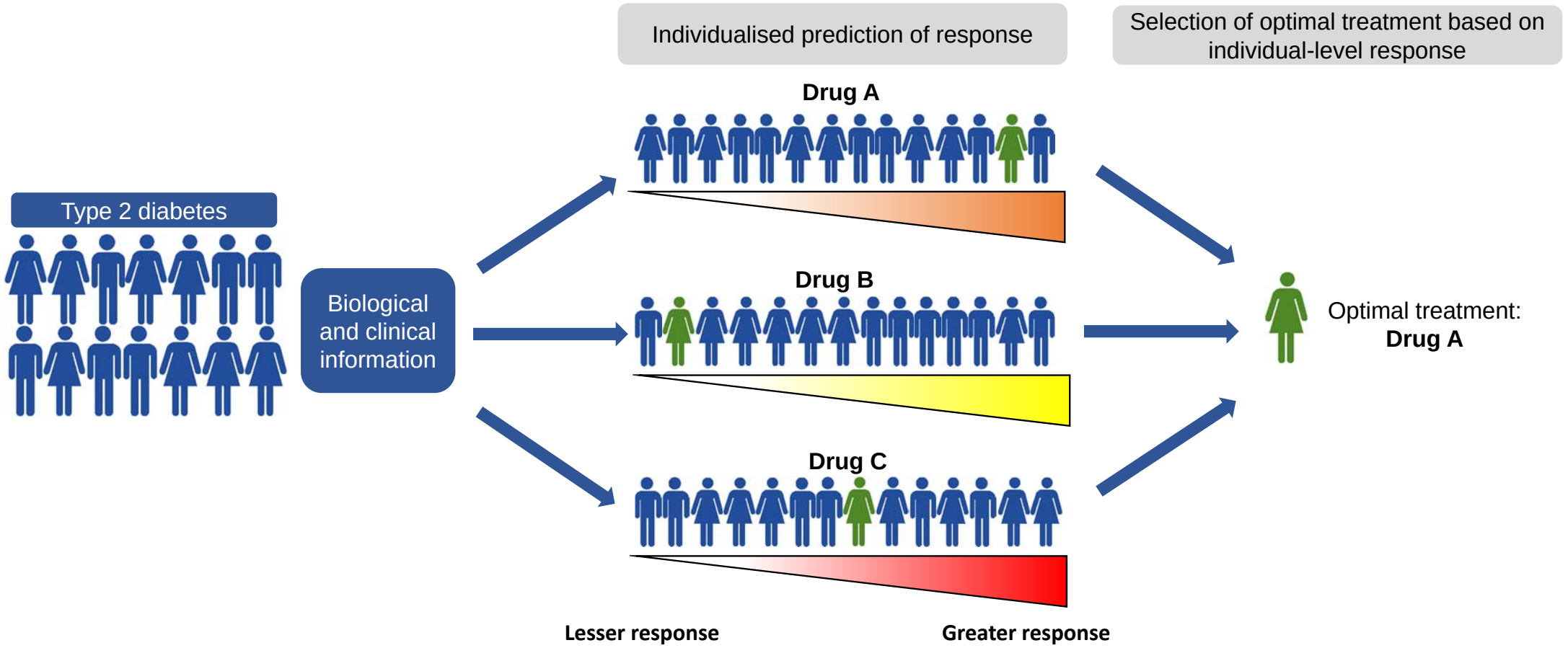
Disadvantages

- People within a subtype may be very different but are assumed to have the same outcome.
- Cannot be assumed to represent true pathophysiological subtypes - highly dependent on features used to classify them.

- Subtypes are not discrete, but overlap in phenotypic characteristics.

- Subtypes not stable – unless defined solely by genetics, a person can shift from one subgroup to another over time.

Figure 6B



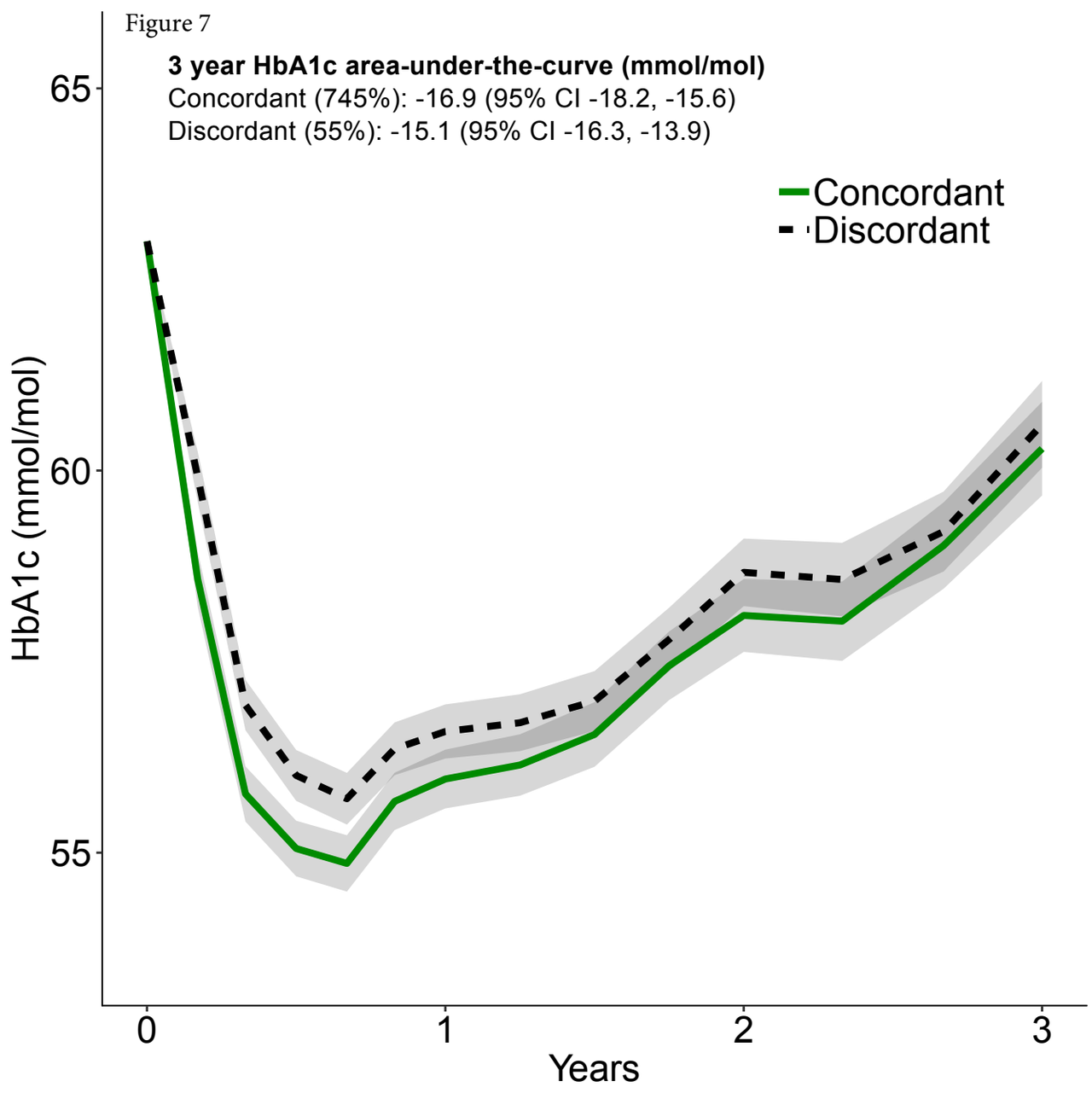
Advantages

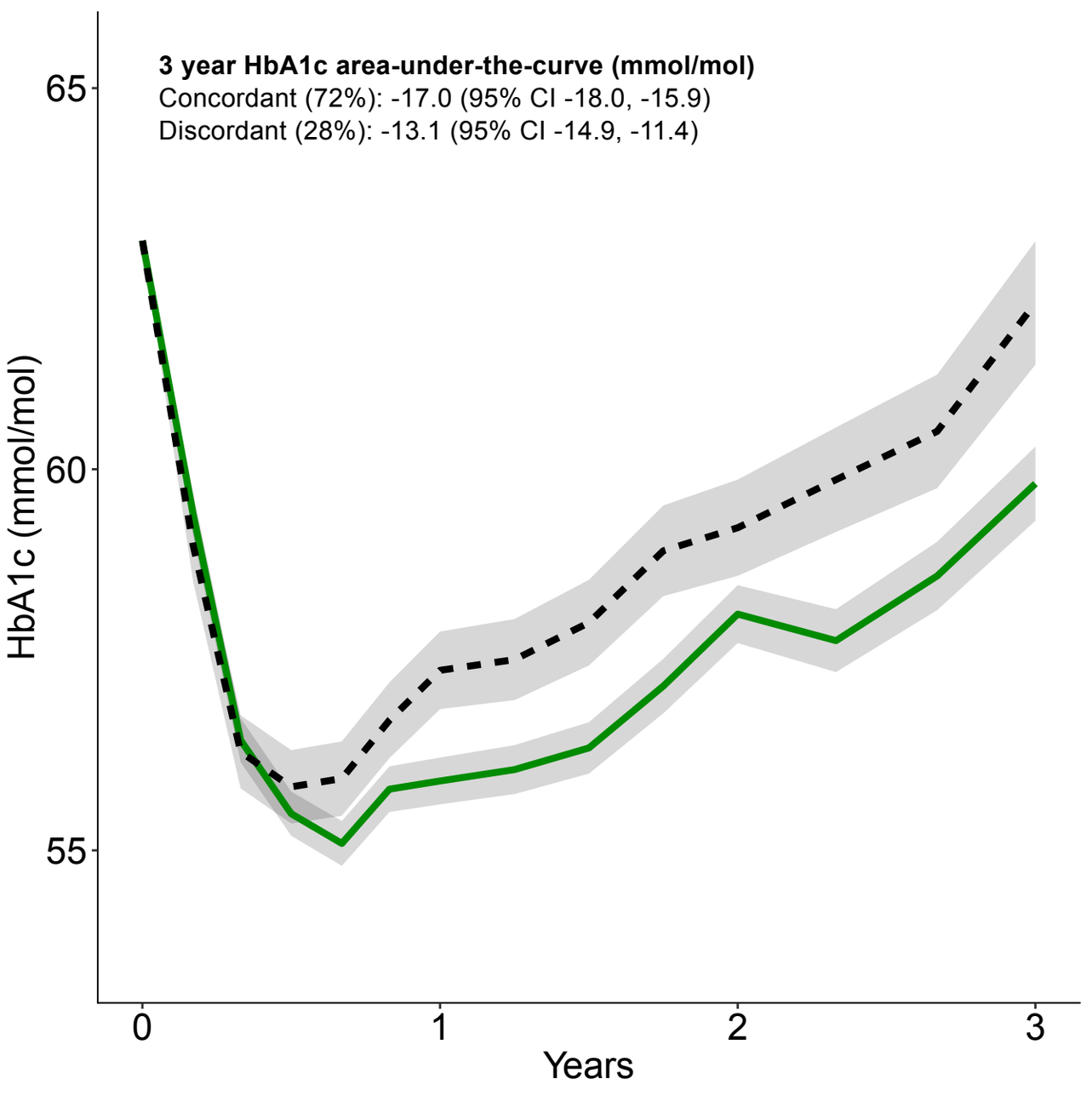
- Optimal prediction of outcome, as predictions based on precise individual level characteristics.[39]
- Predictions specific to a person's characteristics at the point in time an optimal treatment strategy is being considered.

Disadvantages

- Complexity – specific models required for different outcomes e.g. risk of complications.
- Challenging to weigh up models for different outcomes and communicate these.
- Input data for prediction required at different time points.

Figure 7


3 year HbA1c area-under-the-curve (mmol/mol)
Concordant (745%): -16.9 (95% CI -18.2, -15.6)
Discordant (55%): -15.1 (95% CI -16.3, -13.9)
Concordant
Discordant
65
60
55
HbA1c (mmol/mol)
0
1
2
3
Years


3 year HbA1c area-under-the-curve (mmol/mol)
Concordant (72%): -17.0 (95% CI -18.0, -15.9)
Discordant (28%): -13.1 (95% CI -14.9, -11.4)
65
60
55
HbA1c (mmol/mol)
0
1
2
3
Years