

Clinical utility of data-driven clusters of type 2 diabetes

We thank Emma Ahlqvist and colleagues for their comments on our study,¹ itself a follow-up to their original paper which proposed 5 novel subgroups of adult-onset diabetes based on data-driven cluster analysis.²

We agree that our analysis was limited to assessing the clinical utility of the clusters and did not assess any broader insights into the pathophysiology of Type 2 diabetes or the aetiology of complications arising from the clustering approach. The potential insights into pathophysiology of complications in Type 2 diabetes are exciting and we look forward to hearing more about this in future publications. We accept the models we propose are constructed to accurately predict clinical outcomes and do not readily lead to pathophysiological insights.

We think the individual data from clinical trials used in our study offer considerable advantages over population-based cohorts for the assessment of optimal treatment and clinical utility. Treatment is randomly allocated, and follow-up is at set times and protocol driven so provides robust information on treatment response and clinical outcomes. We agree that clinical trials will represent a subgroup of the total population with diabetes. The selection of specific patients will have a disproportionate impact on some of the clusters however it is interesting that the proportion of individuals allocated to each cluster in the trial data were similar to those described in population cohorts.

In our view, precision medicine approaches in type 2 diabetes are likely to have greatest impact on clinical practice if based on simple and reproducible clinical measures available in any diabetes clinic.³ The proposed 5 clusters approach is limited as it requires HOMA-measures at diagnosis, based on fasting glucose and either fasting insulin or C-peptide, which are rarely measured in the clinic. Variability of the precision of fasting insulin or C-peptide assays provide an additional barrier to the use of HOMA-measures in the clinic.^{4,5} In contrast, our approach used only routine clinical measures (BMI, age at diagnosis, HbA1c and renal function), and showed that, when modelled continuously, these simple measures outperform the more complex 5 clusters to select treatment and predict disease progression.

In conclusion we think the approach used should depend on the outcome you want. Exciting new aetiological insights into diabetes complications may arise from data-driven approaches to classification such as the 5 clusters proposed by Ahlqvist and colleagues. However, for practical approaches to personalising type 2 diabetes care, models using an individual's precise clinical measures are likely to have more utility than classification approaches that use clinical measures to assign individuals into subgroups. This means there is a need for both approaches as they have different roles. The final arbitrators of the most useful approach will be the clinicians who need to select treatment or predict likely outcomes.

References

1. 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.
2. Ahlqvist E, Storm P, Käräjämäki 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.
3. Hattersley AT, Patel KA. Precision diabetes: learning from monogenic diabetes. *Diabetologia* 2017; **60**(5): 769-77.
4. Wallace TM, Levy JC, Matthews DR. Use and abuse of HOMA modeling. *Diabetes Care* 2004; **27**(6): 1487-95.
5. Jones AG, Hattersley AT. The clinical utility of C-peptide measurement in the care of patients with diabetes. *Diabetic medicine : a journal of the British Diabetic Association* 2013; **30**(7): 803-17.