

Leveraging big data using a novel clinical database and analytic platform based on 323,145 individuals with and without of Diabetes

Hiranjith G.H¹, Anjana Ranjit Mohan², Praveen Raj³, Jebarani Saravanan², Srinivasan Vedantham³, Radha Venkatesan², Muthu Narayanan², Pradeepa Rajendra², Ranjit Unnikrishnan², Somasekhar Jayaram³, Rohit Gupta³, Paul George³, Brijendra Kumar Srivastava², Uthra Subash Chandra Bose², Lovelena Munawar², Sam Santhosh³, Mohan Viswanathan²

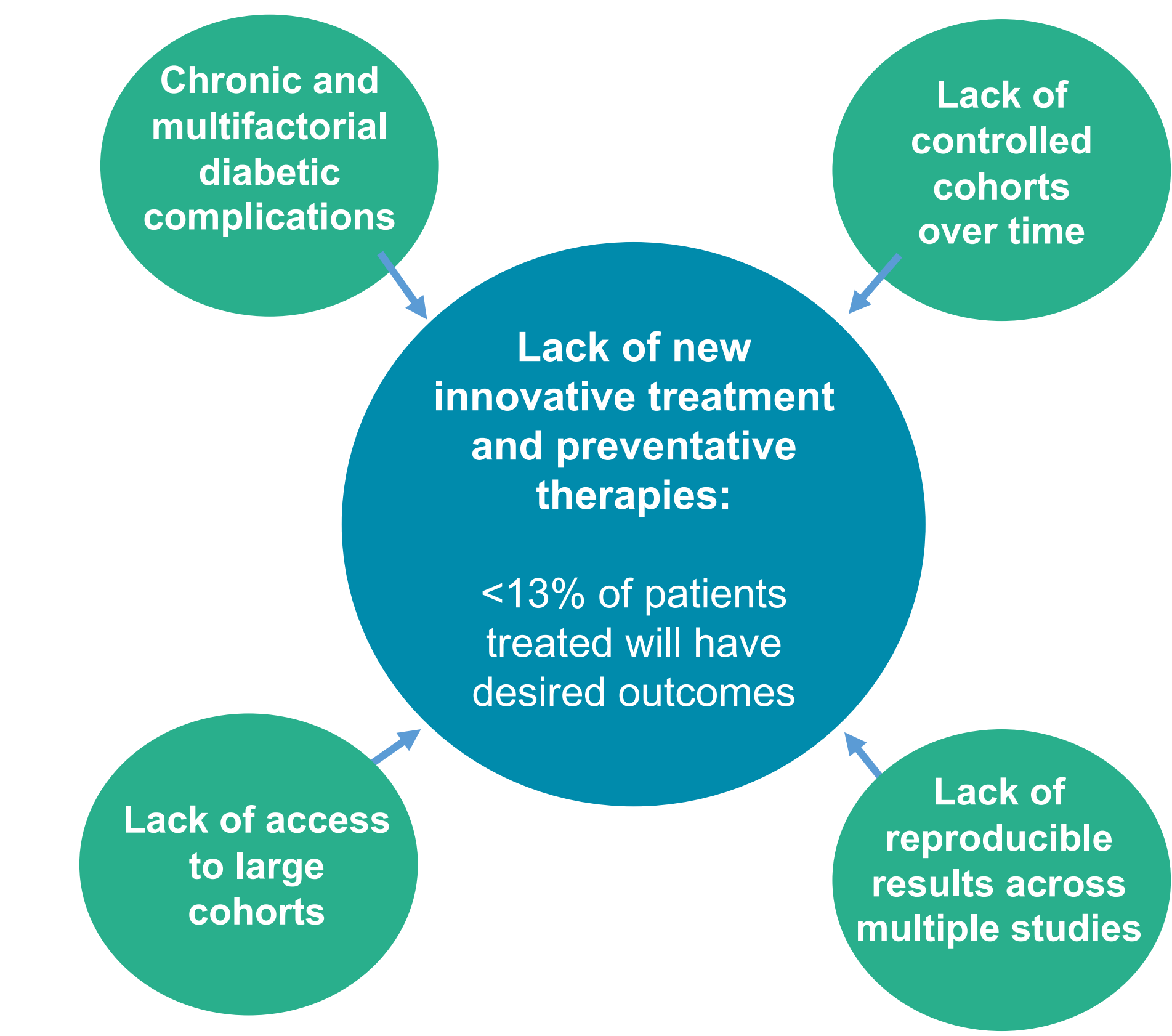
¹MedGenome Inc, Foster City, CA, USA, ²Dr. Mohan's Diabetes Specialities Centre, Chennai, India
³MedGenome Labs Pvt. Ltd., Bangalore, India

Introduction

Diabetes is a chronic disorder of glucose metabolism and is a major cause of heart disease and end-stage renal disease in world populations. It is also the single biggest cause of preventable blindness, the leading cause of non-traumatic lower extremity amputation and major cause of premature mortality. **415 million people have diabetes globally and is expected to reach 642 million by 2040¹**. Large volumes of diabetes biomedical data are being produced every day, but it has not been used effectively. Leveraging such voluminous amount of patient data using data science approaches help to uncover hidden patterns, unknown correlations, and other insights of the disease. Integration of diverse genomic data with comprehensive electronic health records (EHRs) exhibit challenges, but essentially, they provide a feasible opportunity to better understand the underlying diseases, treatment patterns and develop an efficient and effective approach to identify biomarkers for diagnosis and improve therapy.

Here, we describe a big data solution for diabetes - **Diabetome**, providing an efficient and responsive scientific discovery platform for researchers.

Challenges in diabetes R&D today



Diabetome

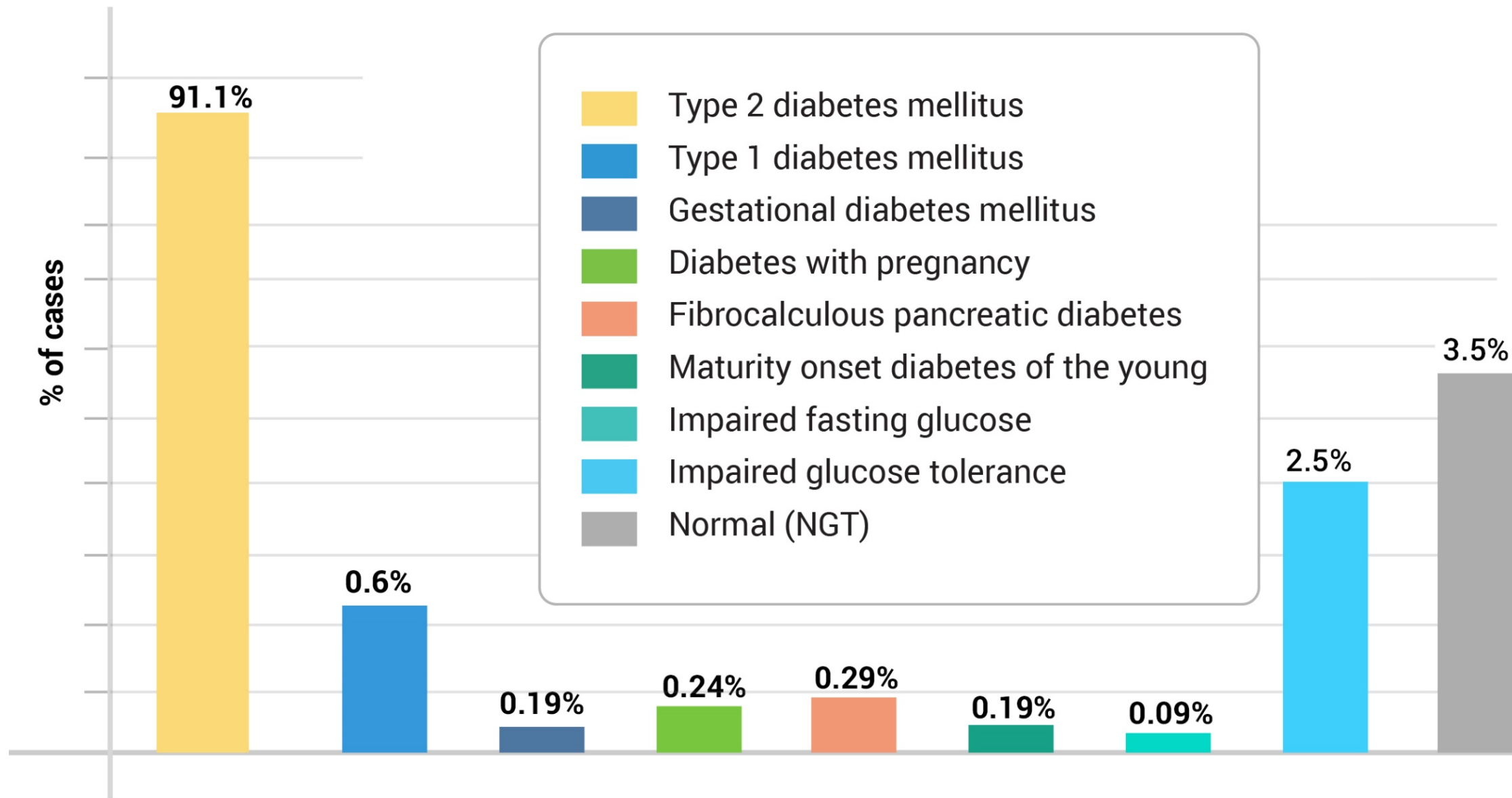
A unique clinical database for diabetes R&D to enable new insights into diabetes, and its risk factors and complications for research, development and clinical diagnosis.

Data

Clinical phenotype data was collected from EHR of a tertiary care diabetes centre across 20 locations in India. It encompasses >20 million data points on **323,145** patients registered over 25 years.

Type of patient data

- Demographics
- Clinical characteristics
- Biochemical results
- Family history
- Treatment history
- Drug responses
- Genotypes
- Secondary complications of diabetes



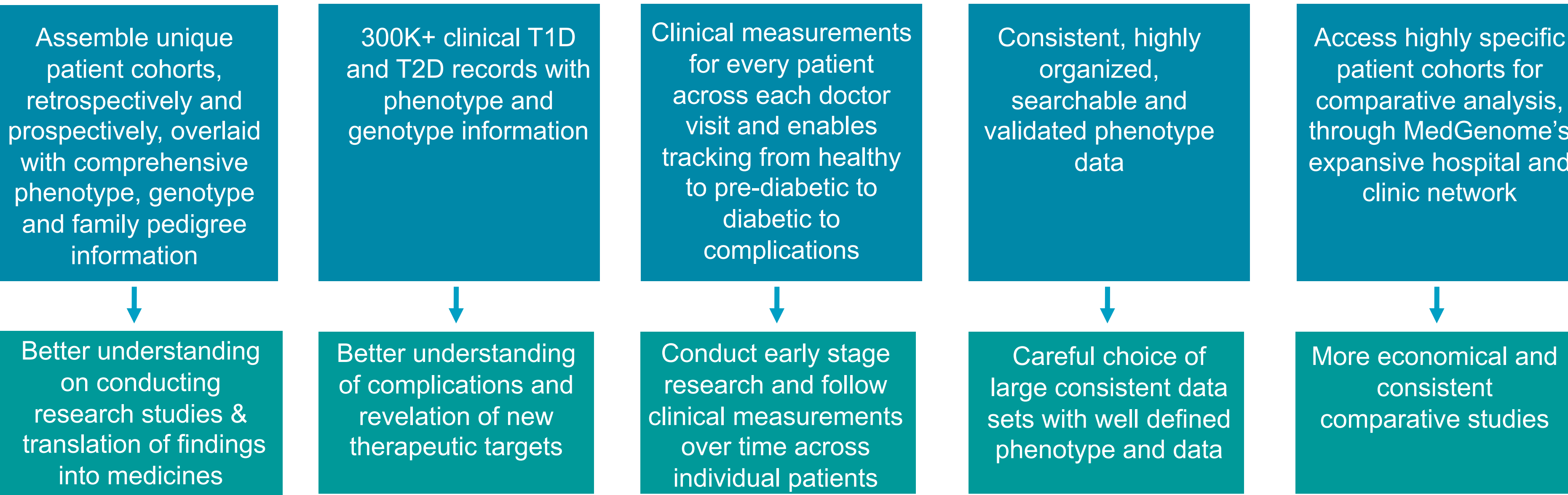
Subjects with secondary complications of diabetes

Diabetic retinopathy (75,449)	Cardio vascular disease (8,508)
Diabetic nephropathy (66,395)	Peripheral vascular disease (6,500)
Diabetic neuropathy (74,833)	Non-alcoholic fatty liver disease (13,157)

Tools

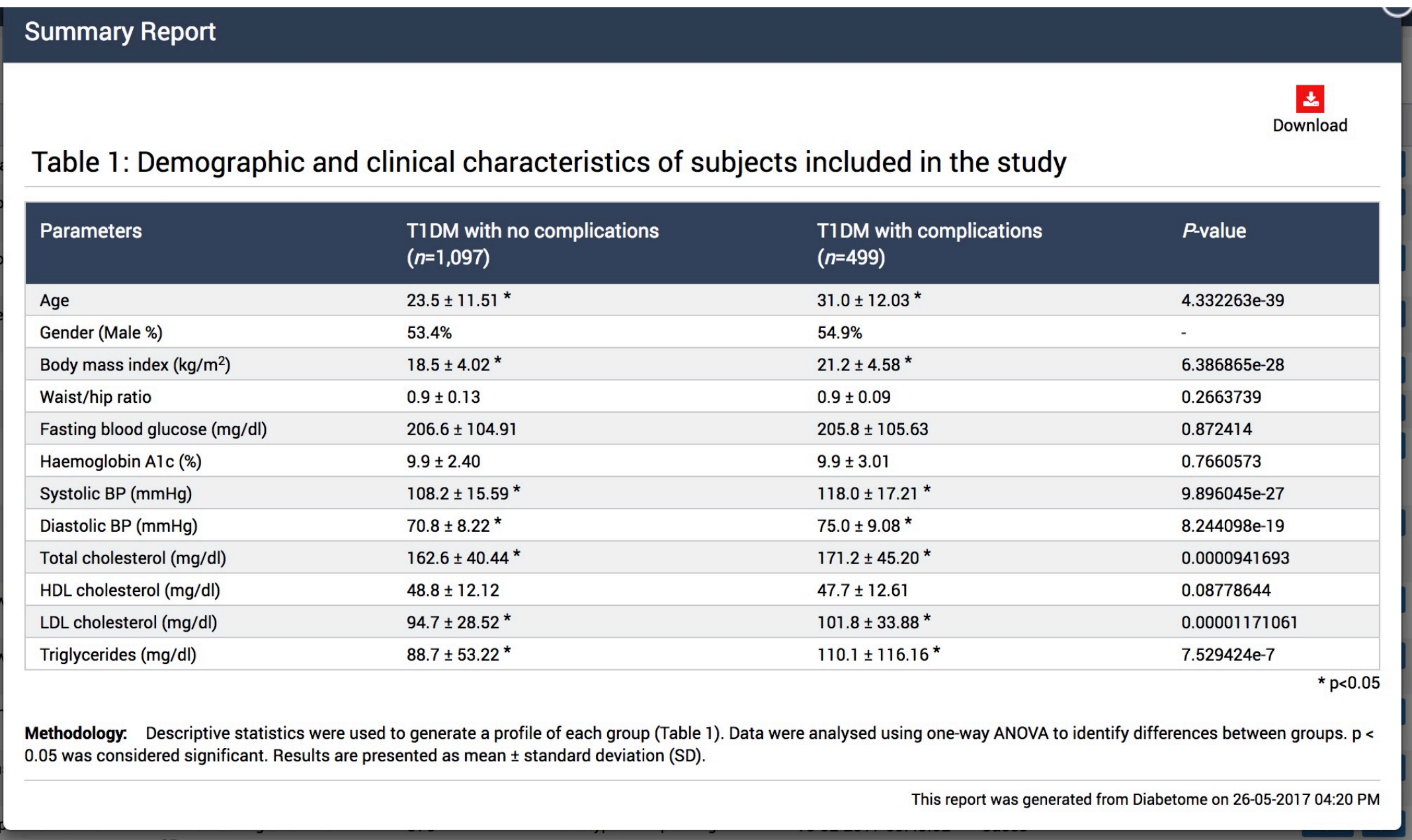
Analytical tools to compute descriptive statistics and compare cohorts and **visualization** techniques to understand the disease progression and uncover patterns of high glucose levels, or drug treatments and their association with progression of disease over time.

Applications of Diabetome

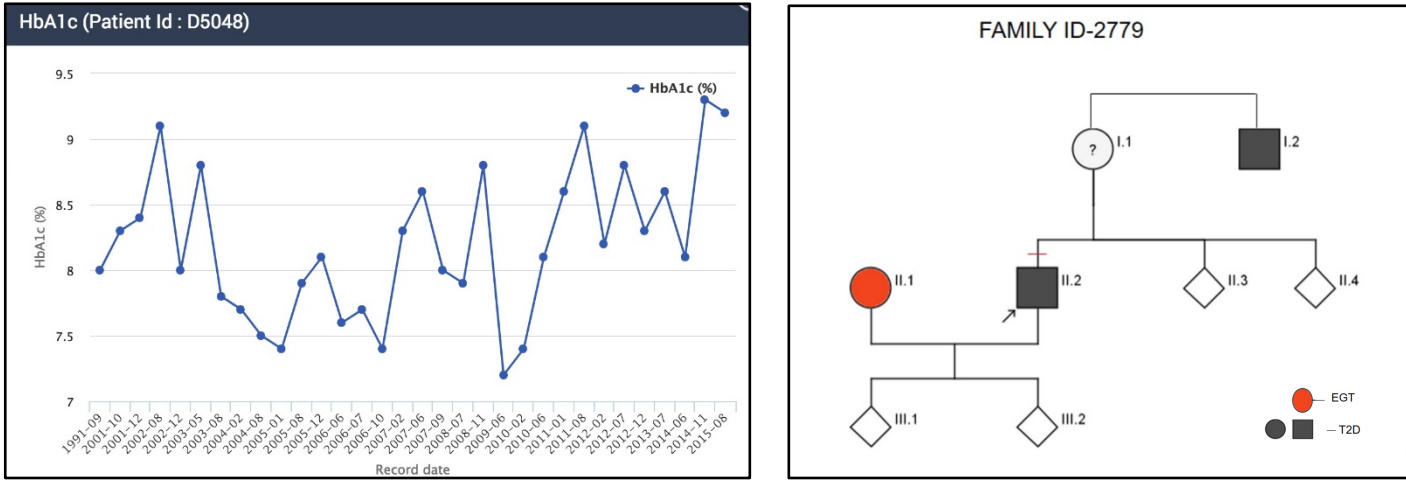


Sample reports and visualizations in Diabetome

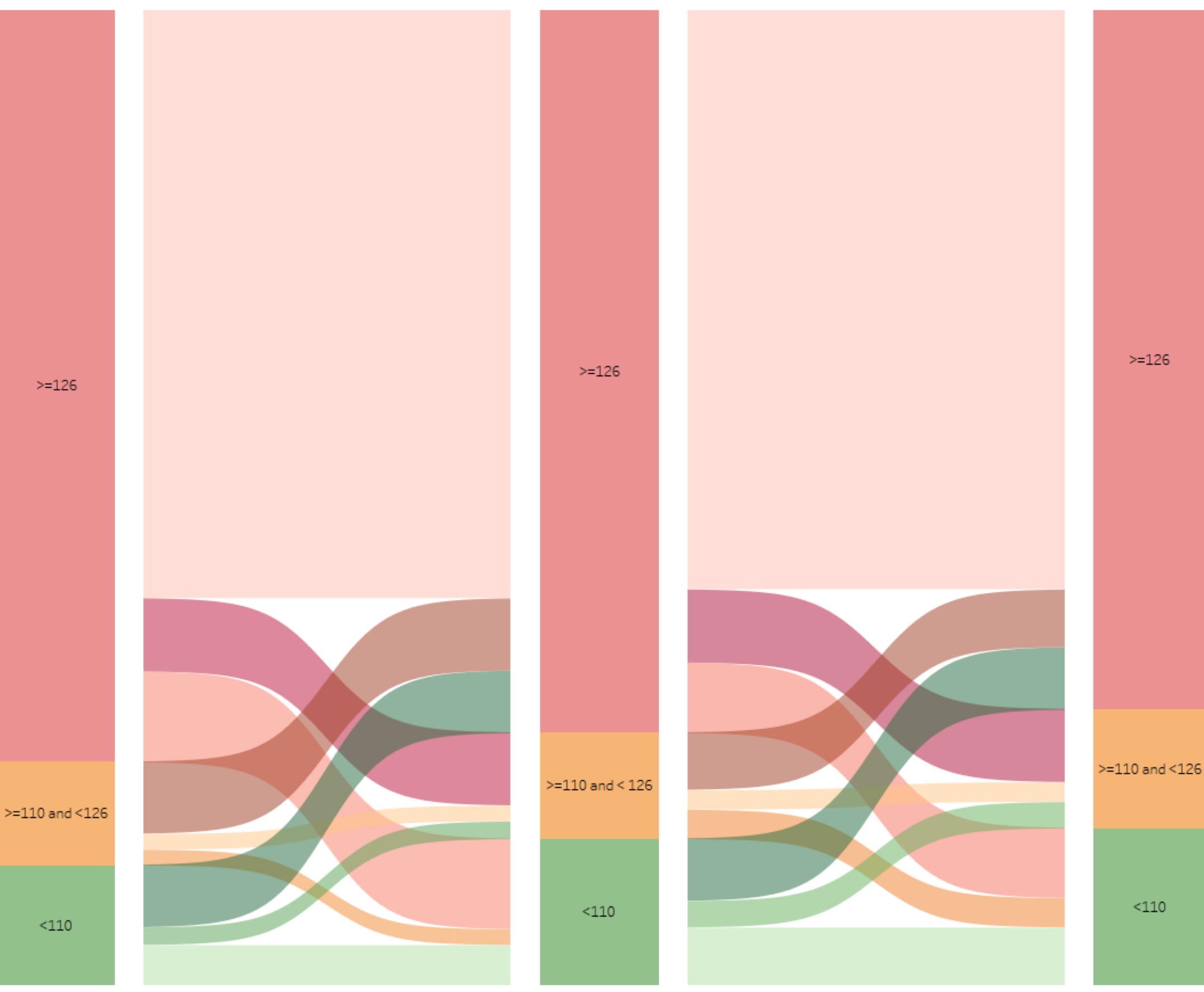
(1) A summary report showing the demographics and clinical characteristics between two cohorts – T1DM with complications and those without complications.



(3) Longitudinal view of HbA1c and pedigree chart of family



(2) Visualization of progression of diabetes in normal glucose tolerant, pre-diabetes and Type 2 diabetes subjects based on fasting blood glucose



Conclusion

- Diabetome is a large, unique repository of clinical phenotype data of diabetes, and a scientific discovery tool for researchers to investigate diabetes and related complications at clinical and molecular levels.
- It caters to early stage pharma drug discovery and development endeavours by leveraging the abundant clinical data.

References

1. International Diabetes Federation. IDF Diabetes Atlas. 7th edn. Brussels, Belgium: International Diabetes Federation. 2015.