precisionlife
driving precision medicine

Disease Study

Identification of Genetic Drivers of Complications in Type 2 Diabetes

precisionlife.com

Identification of Genetic Drivers of Complications in Type 2 Diabetes

Executive Summary

Type 2 diabetes-related complications create some of the most significant health and economic burdens in developed and developing countries. Complications associated with type 2 diabetes lead to high levels of expensive additional hospitalizations and more radical treatment interventions (including dialysis and amputations) as the disease progresses. They are directly responsible for poor quality of life for patients, with higher long-term social care costs and mortality rates. They are also predisposing features that can be associated with debilitating conditions such as cardiovascular disease and various forms of dementia.

PrecisionLife analyzed a dataset from the UK Biobank, comparing cases with a variety of type 2 diabetes-associated complications (such as ketoacidosis, nephropathy, and retinopathy) against gender-matched controls who had also been diagnosed with diabetes and had similar BMI measurements, but had not (yet) developed any complications.

We found significant genetic differences between the two populations, indicating that there is a subset of diabetic patients with additional genetic features that predispose them to severe diabetes and diabetes-related complications, independent of lifestyle and environmental factors.

Our analysis revealed several single nucleotide polymorphisms (SNPs) and genes that were highly associated with the development of specific diabetes-related complications. These associations would not have been found using standard Genome-Wide Association Studies (GWAS) analysis techniques.

The diabetes complication subpopulations for which genetic signatures were identified included:

- general (five genes associated with an increase in all complications)
- acute (five genes—three ketoacidosis, two coma)
- neurological (six genes)
- peripheral (eight genes)
- renal (four genes)
- ophthalmic (seven genes)
- type 1 diabetes (four genes).

Having stratified this case population into distinct complication-specific subtypes, we identified individual genes and biological mechanisms associated with each. These high-resolution disease insights enable novel drug discovery, the development of combinatorial risk scores, and combinatorial biomarkers to distinguish the patient subgroups and evaluate individual risk.

These can be used to identify patients most at risk and inform development of targeted lifestyle and therapeutic interventions, in order to prevent the development of these complications. This could result in better patient outcomes with respect to disease progression, and reduction of the health and economic burden caused by diabetes complications.

Introduction

Type 2 diabetes is among the most common chronic diseases globally, affecting around 5–10% of populations in both developed and developing nations worldwide. It leads to long-term reduction in quality of life and significantly decreases life expectancy by up to 10 years.[1]

Due to their severity and long-term nature, diabetes-related complications impose one of the greatest economic and health burdens associated with any disease. In the UK, the cost of diabetes treatment corresponds to around £14 billion—10% of the National Health Service (NHS) budget. Around 80% of this is spent on the management of disease complications, corresponding to about 0.65% of UK GDP. These costs are expected to grow as a result of an increasingly overweight and ageing population.[2, 3]

Diabetes-related complications can broadly be split into several categories, including micro- and macrovascular complications, as well as peripheral neuropathies and renal, ophthalmic, and acute conditions such as ketoacidosis and coma. Such complications are one of the main causes of reduced quality of life, long-term care needs, and mortality in diabetic patients. Both the primary disease itself and, particularly, its complications are preventable with the appropriate management of blood glucose levels, lifestyle, and educational and therapeutic interventions.

In a previous study, we analyzed a cohort of White British type 2 diabetes cases found in the UK Biobank and compared them against BMI-matched healthy controls. We found that this White British type 2 diabetes population stratified into five distinct clusters based on their genotypes only. Adding phenotypic and clinical information revealed that two of the clusters were particularly associated with type 2 diabetes-related complications (see Figure 1). This indicates that subsets of

type 2 diabetes patients are at a higher risk of developing complications as a result of their genotype. In a separate study we have developed these insights into a highly predictive model that stratifies these patients into five classifications with respect to their degree and type of complication risk.

Figure 1 Heatmap showing the frequency of ICD-10 codes for diabetes-related complications for five type 2 diabetes case clusters identified from a White British population's medical histories, as reported in UK Biobank. The relative incidence of a disease in the case clusters are colored from green (high) through to red (low). The case clusters were identified by clustering combinatorial disease signatures vs. case samples using a k-modes clustering method.

In this study, we analyzed a different dataset from the UK Biobank, comparing all cases (regardless of ethnicity) with a variety of type 2 diabetes complications against controls who had also been diagnosed with diabetes and had similar BMI measurements, but had not (yet) developed any complications. We aim to increase the understanding of the genetic risk factors and biological mechanisms underlying the development of diabetes-related complications, in order to identify those diabetic patients who are at higher genetic risk of developing them, and to inform targeted prevention/treatment strategies.

Study Design and Methods

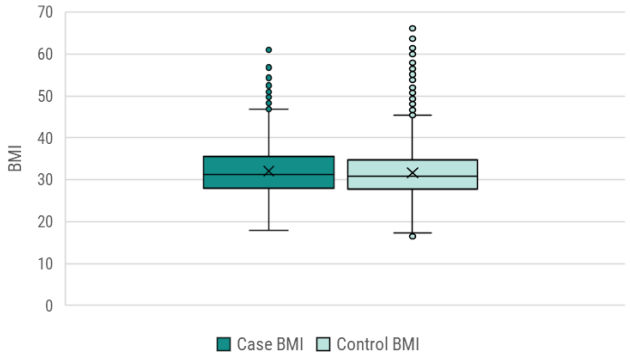
We obtained genotype data from patients (cases) in the UK Biobank[4] who have been diagnosed with type 2 diabetes (ICD-10 code E11), and who have gone on to develop at least one of the main complications associated with diabetes, including ketoacidosis, neurological complications, and kidney disease (see Appendix for more detail). After quality control and removal of samples with missing data, we identified 2,900 cases (1,924 male, 976 female) who had been diagnosed with type 2 diabetes-related complications.

We performed a case-control study, comparing these cases against 5,800 gender-and BMI-matched controls who had been diagnosed with type 2 diabetes, but had not (yet) developed any of the associated complications. We selected the oldest possible controls who met these criteria in order to have a "supercontrol" population. This meant that the control age distribution was skewed more heavily to older age groups than the case population. This would not be appropriate for some studies, but for the purposes of this study it avoids, as far as possible, the potential for undiagnosed "controls" who might go on to develop a complication later.

Cohort Analysis

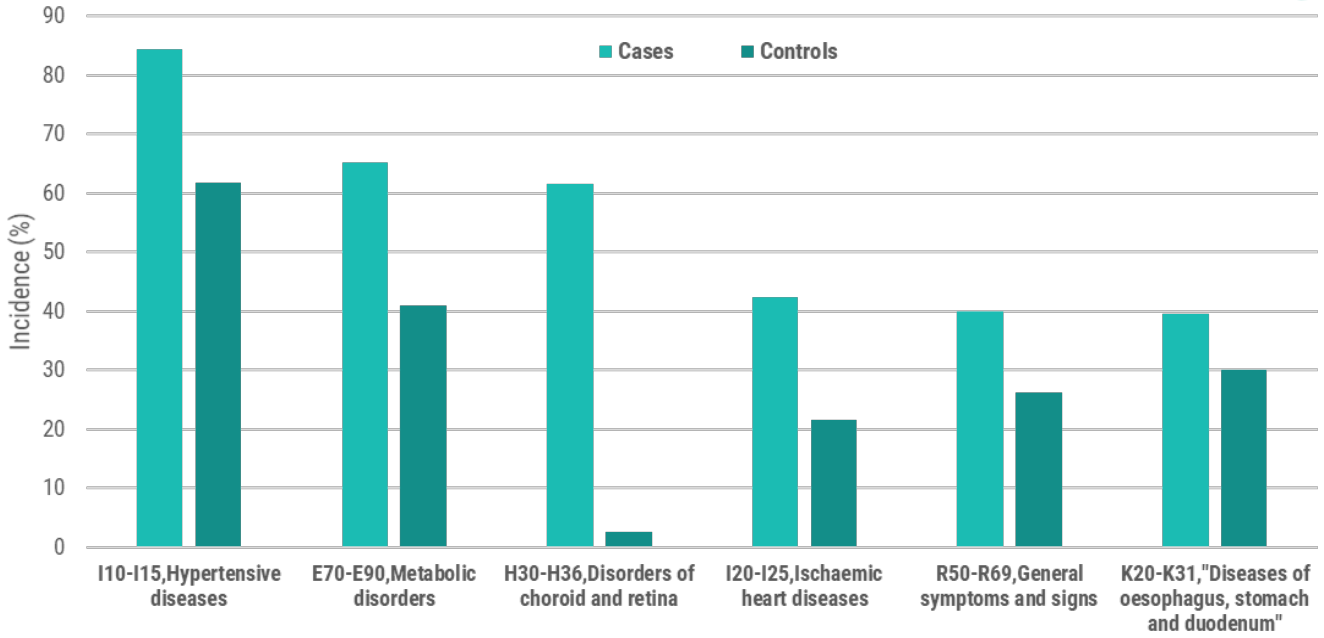
As obesity is one of the key drivers of insulin resistance and type 2 diabetes progression, we first compared the BMI measures from the diabetic cases and control populations. Although the mean BMI of cases was very slightly higher than controls, there was no significant difference between the two cohorts (see Figure 2). This indicates that any signals observed in the cases are unlikely simply to be products of greater adiposity and poor dietary lifestyle-related factors driving progression of diabetes complications.

Figure 2 Boxplot showing the mean and range of BMI measures found in diabetic complications cases and controls



The most common comorbidities seen in cases and controls were also very similar, with hypertensive disease, metabolic disorders, ischemic heart disease, and digestive health problems most frequently recorded for both (see Figure 3). However, the proportion of cases with these comorbidities was slightly higher than in the control non-complication diabetic population.

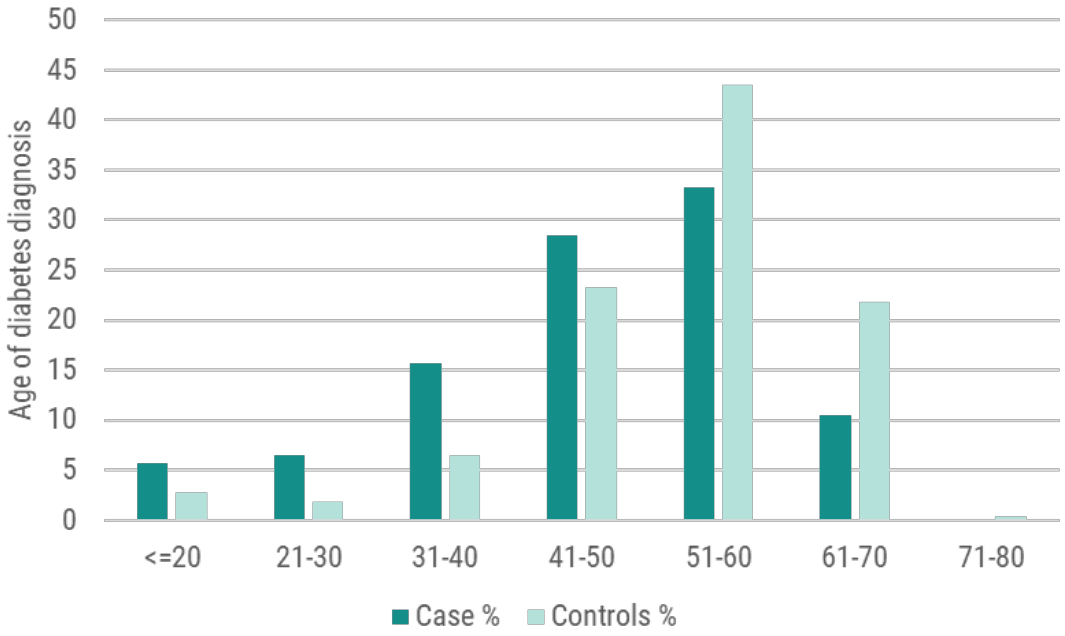
Figure 3 Incidence (percentage of patients) with comorbidities by ICD-10 code from the UK Biobank for diabetic cases (with complications) and controls (with no complications)



Finally, complications cases in the UK Biobank were, on average, diagnosed with diabetes earlier than the non-complication controls (see Figure 4). This may indicate that these cases might be diagnosed earlier because they have more severe forms of diabetes, and so display symptoms earlier than the complication-free controls.

This could indicate an additional genetic driver of disease, causing faster and earlier disease progression, while the disease in the control population may occur as result of a cumulation of lifestyle and environmental factors, and so develops later.

Figure 4 Distribution plot showing age of diabetes diagnosis for cases and controls



We analyzed this case-control dataset using the PrecisionLife® platform to identify risk-associated combinatorial signatures comprising multiple SNPs that, in combination, were strongly associated with complications arising from type 2 diabetes. The PrecisionLife platform identifies high-order, disease-associated combinations of features at whole-genome resolution in large patient cohorts. These can be based on genotype only, phenotype only, or mixed (e.g. SNPs, transcriptomic, epidemiological, and/or clinical) datasets. The improved discovery and stratification performance of PrecisionLife over techniques such as GWAS, Phe-WAS, etc. has been validated across multiple different disease populations.[5, 6, 7]

When applied to genomic data, PrecisionLife finds high-order epistatic interactions (multi-SNP genotype signatures—typically of combinatorial order between 3 and 8) that are significantly more predictive of patients' phenotype than those identified using existing single SNP-based methods such as GWAS. Some of the individual SNPs making up these signatures may, if assessed individually across the whole population, fall below the GWAS significance thresholds. However, we have demonstrated that when evaluated in combination with each other using multiple statistical validation techniques, these SNPs can be highly predictive in delineating disease subpopulations. The combinatorial signatures encode non-linear additive effects from multiple SNPs, providing more predictive power than polygenic methods.

Case-associated SNP genotypes and their high-order combinations were validated in the PrecisionLife methodology using 1,000 cycles of fully random permutation, with a False Discovery Rate filter of 1% in conjunction with a Benjamini-Hochberg correction. The validated SNPs were then scored using a Random Forest algorithm based on a five-fold cross-validation method to evaluate the accuracy with which the combinations predict the observed case-control split.

The SNP combinations were then mapped to the human reference genome[8] to identify disease-associated and clinically relevant target genes. A semantic knowledge graph derived from multiple public and private data sources was used to annotate the SNP and gene targets, including relevant tissue expression, chemical activity/tractability for gene targets, functional assignment, and disease-associated literature.

We then traced back all the significant disease signatures to the diabetic cases in which they were found, and associated them with ICD-10 codes relating to specific diabetes-associated complications. This generated high-resolution patient stratification insights, allowing us to identify specific genetic signatures associated with each complication.

Results

GWAS analysis of this dataset using PLINK 1.9[9] yielded no significant results (below 5×10^{-8} *p*-value; see Figure 5).

The PrecisionLife platform identified 135 combinations of SNP genotypes (N-states) that were highly associated with the development of type 2 diabetes-specific complications (Table 1). All of the SNPs were found in combinations of three SNPs or more, and so could not have been discovered using standard GWAS analysis methods.

Figure 5 Manhattan plot showing results of GWAS analysis of diabetes complications cases vs. controls

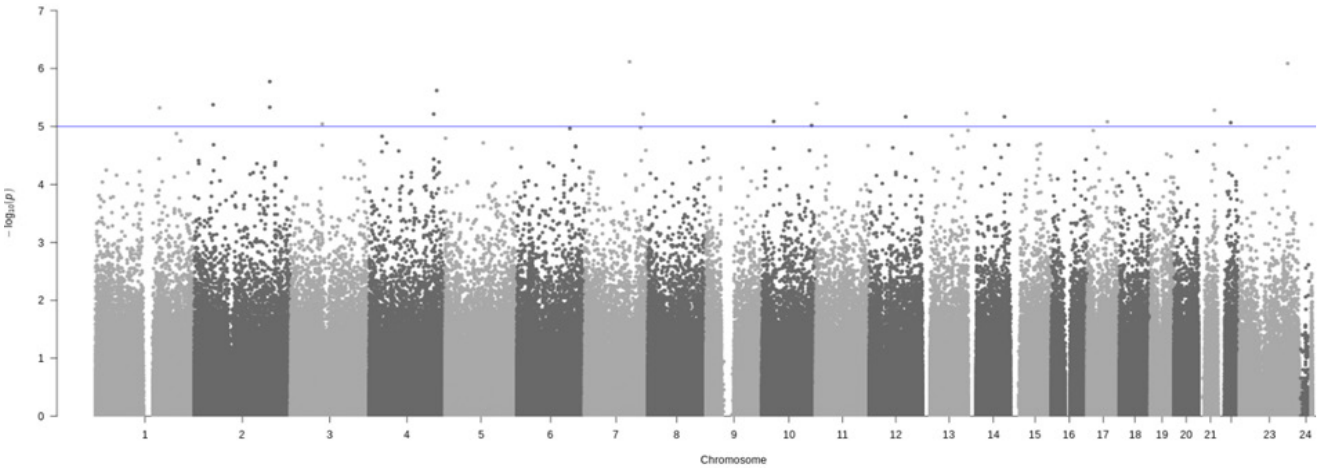


Table 1 Summary of PrecisionLife diabetes disease study comparing complications cases vs. no complications controls

	Diabetes Complications Study Results
Total N-states	135
Unique SNPs	35
Genes	20
Penetrance (cases represented by all signatures)	64%

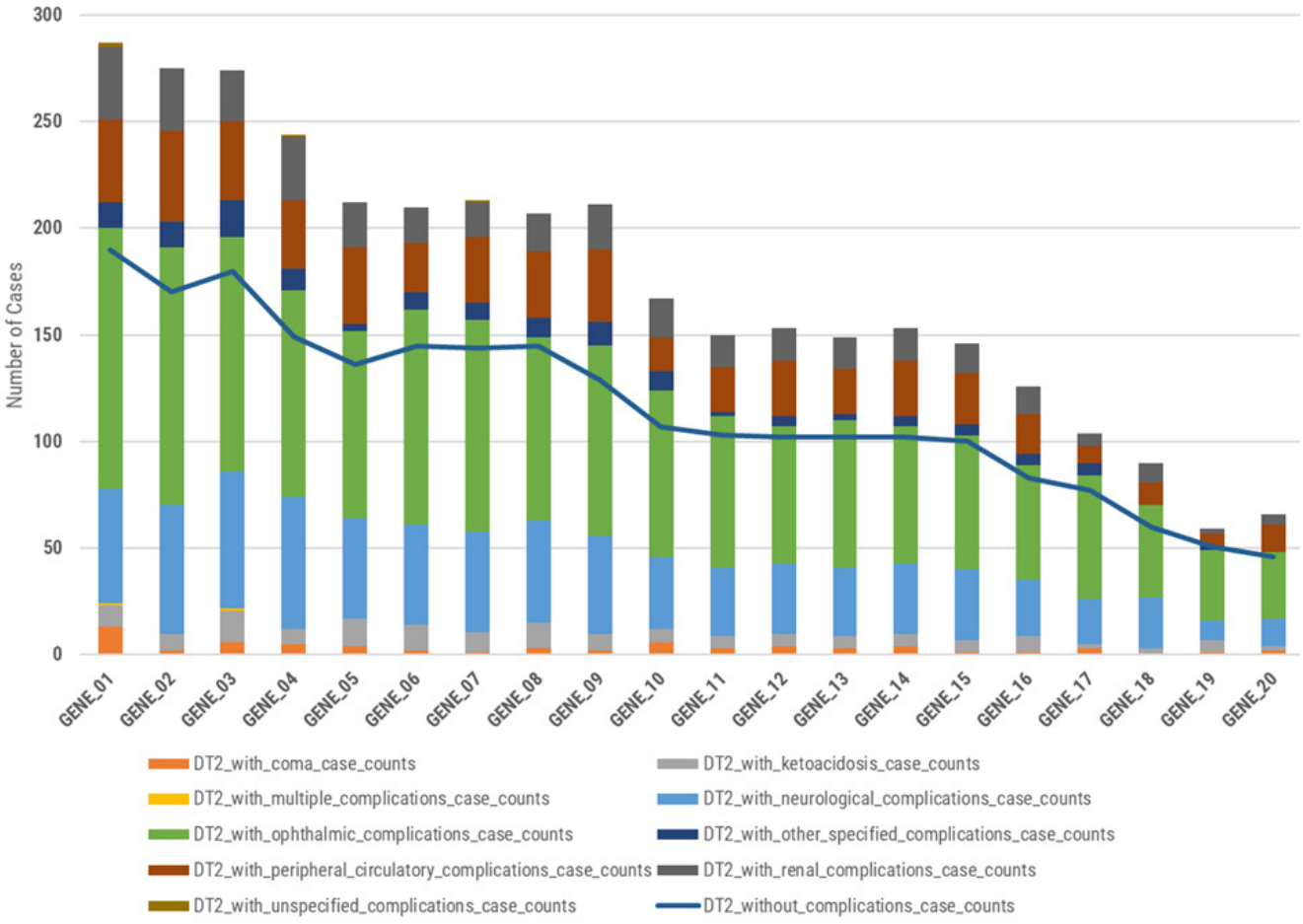
Mapping the SNPs to genes revealed 20 protein-coding genes that are highly associated with the risk of developing type 2 diabetes-related complications. Several of these genes are implicated in diabetes-related pathological mechanisms such as insulin resistance and angiogenesis, while others have already been associated with diabetes-specific complications such as diabetic retinopathy and end-stage renal disease.

This provides validation for the combinatorial analytics approach, and indicates that there are significant genetic differences between individuals who develop diabetes-related complications and those who do not.

We can trace back the genetic variants that are particularly associated with patients developing specific diabetes-related complications, enabling higher-resolution patient stratification insights (see Figure 6). We can use these insights to gain further understanding of the biological mechanisms underpinning each of these complications, and the genetic factors increasing the risks of developing them.

In separate studies, screening these gene targets allows us to find highly tractable targets for novel drug development. We have also used this analysis to identify combinatorial risk factors associated with five subpopulations of White British diabetics, and a sixth subgroup of mainly South Asian patients. The risk factors between these populations are quite different, with causative factors in White British being more phenotypic, and South Asian being more genetic.

Figure 6 Stacked bar plots showing the number of cases with different complications who are affected by the risk-associated genes identified by the PrecisionLife platform. The line plot shows the total unique number of cases who are affected for each gene.



Discussion

The resolution of our insights into complication-specific patient stratification is slightly limited by the summary diagnostic ICD-10 codes that are assigned to patients in the UK Biobank.

Although several of the macrovascular complications such as coronary artery disease and stroke are listed in the UK Biobank, there is no indication of whether a patient developed them as a direct result of hyperglycemia/insulin resistance, or if the cardiovascular outcome occurred before the type 2 diabetes developed. For this reason, we are unable to analyze the genetic factors underlying major cardiovascular events that are a direct consequence of diabetes pathology.

Furthermore, the subcategories relating to diabetes-specific complications (E11.x) are only split into general categories such as renal, ophthalmic, and peripheral complications. This means that we cannot refine the patient groups into more specific diseases, like diabetic ulcer or macular edema.

General Complication Risk Genes

Many of the biological processes underpinning these ICD-10 subcategories—such as renal, ophthalmic, and peripheral complications—are interconnected, and so it is unsurprising that a high proportion of patients have been diagnosed with multiple complications.

When performing a gene-wise stratification analysis, we identified five gene variants that were not particularly enriched for any specific complication, but were common among many patients who develop diabetes-related complications. Among these genes, we identified several that were involved in regulation of angiogenesis. Dysregulated angiogenesis is a major driver of several different diabetes-associated complications, including nephropathy, retinopathy, and vascular dementia.[10]

We also found several genes that are associated with insulin resistance, glucose oxidation, and hypertension risk. These genes may be predisposing patients to general diabetes disease severity, rather than the development of a particular complication, which is why they were not particularly enriched in specific patient subtypes, but common across the whole case population.

Acute Complications

Ketoacidosis and diabetic coma are two subtypes of complications associated with type 2 diabetes that are acute systemic conditions, rather than a chronic disease.

Ketoacidosis is caused by extreme insulin deficiency, often in combination with increased levels of counterregulatory hormones such as glucagon, resulting in ketonemia, hyperglycemia, and hyperosmolarity.[11] The most common causes of ketoacidosis, such as infection, insulin omission, and new/misdiagnosis, are driven by non-genetic factors, which may explain why we identified few genetic variants that were significantly overexpressed in patients admitted with diabetic coma or ketoacidosis.[12]

However, the majority of the genes that were most enriched in patients with ketoacidosis and coma are implicated in processes such as insulin resistance and glucose homeostasis. This suggests that some patients who develop acute complications such as ketoacidosis may be genetically predisposed toward a more severe form of diabetes and poorer glucose control, rather than specific genetic drivers expressed in the tissues of localized complications.

Neurological Complications

The ICD-10 code for neurological complications (E11.4) includes conditions such as diabetic neuropathy. Diabetic neuropathy is characterized by nerve fiber degeneration, increased oxidative stress and inflammatory cytokines, and axonal loss. These are also hallmarks of neurogenerative diseases, and the link between type 2 diabetes and dementia risk could be a result of shared risk-associated genes driving these shared pathological processes.[13, 14] We found several genes enriched in patients with neurological complications that have been implicated in neurodegenerative diseases such as Alzheimer's disease (AD) and amyotrophic lateral sclerosis (ALS).

One of these genes encodes a GABA-uptake transporter, limiting GABA neurotransmission and regulating GABAergic remodeling and neuro-inflammation in AD. Higher GABA concentrations have been observed in type 2 diabetes patients, with increased levels particularly associated with an increased rate of cognitive decline in patients with diabetic neuropathy.[15]

The same study observed that glutamate levels were simultaneously significantly raised in patients with neuropathy when compared against healthy controls. We also identified a genetic variant in a glutamate receptor subtype that was enriched in patients with neurological complications. This gene has been associated with both ALS and AD, and is hypothesized to drive disease both through excitotoxic and neuro-inflammatory processes, resulting in an "advance ageing" phenotype found in diseases such as AD and type 2 diabetes.

Additionally, we found three more genes that have been implicated in either AD or familial ALS, regulating processes such as neuronal function and development, and neuro-inflammation.

Peripheral Complications

The ICD-10 code for peripheral complications includes patients who develop type 2 diabetes-related circulatory complications and peripheral angiopathy.

Endothelial cell dysfunction plays a key role in the development of vascular disease in type 2 diabetes.[16] Dysfunction results in imbalances in vasoconstriction and vasodilation, and pro-inflammatory and anti-inflammatory cytokines, as well as disrupting smooth muscle cell physiology. This contributes to the increased risk of developing atherosclerosis and other vascular complications that are associated with type 2 diabetes. It is also worth noting that we found similar significant signals associated with endothelial cell dysfunction in analysis of patients who had severe COVID-19 infections.[17] This study noted and controlled for type 2 diabetes as a strong predisposing comorbidity for severe COVID-19.

Although endothelial cell dysfunction is predominantly driven by insulin resistance and hyperglycemia, we found several variants in genes expressed in endothelial cells that may result in greater endothelial cell susceptibility to dysfunction.

We identified a transcription factor that is also expressed in microvascular endothelial cells and is required for TGF-β induced fibrosis. Fibrogenesis through this mechanism has been associated with diabetic cardiomyopathy and fibrosis,[18] as well as diabetic nephropathy and retinopathy.[19] The gene we identified was also upregulated in response to glucose, indicating a direct link between this gene and diabetes-induced vascular endothelial dysfunction.

Another gene that was particularly enriched in diabetic patients with peripheral vascular disease encodes a transmembrane receptor that is highly expressed on the surface of vascular endothelial cells, playing a key role in processes such as endothelial cell capillary morphogenesis. Its function is regulated by several proteins that are protective against various pathological mechanisms in diabetes, including angiogenesis, fibrosis, and pro-inflammatory processes, and inhibits the development of vascular dysfunction and atherosclerosis. Therefore, variants in the gene we identified may be preventing the transduction of the protective effects of these signaling molecules, predisposing patients to endothelial dysfunction and development of cardiovascular complications.

Renal Complications

We find evidence of dysregulated mineralocorticoid receptor (MR) signaling in patients who developed diabetic nephropathy and chronic kidney disease. Over-activation of the renin-angiotensin-aldosterone system (RAAS) is a key driver of renal fibrosis in diabetic kidney disease, and using RAAS inhibitors decreases expression of profibrotic markers and renal function in diabetic rats. Furthermore, MR antagonists have been shown to prevent the progression of diabetic nephropathy by improving insulin resistance and lowering blood pressure.

Ophthalmic Complications

Our analysis revealed several genes implicated in adherens junction formation, actin polymerization, extracellular matrix signaling, and basement membrane function. This could indicate a genetic predisposition to breakdown of retinal epithelial and endothelial barrier integrity and vascular leakage that has been observed in diabetic retinopathy.[20]

The strongest genetic variant in patients with ophthalmic complications such as diabetic retinopathy and macular edema is in a gene involved in the production of a proteoglycan. This family has been shown to contribute to the assembly of the retinal basement membrane, and knockout models demonstrate lack of proteoglycan expression results in hallmarks of retinopathy, such as hyaloid vessel persistence and disruption of the inner limiting membrane.[21]

Type 1 Diabetes

Finally, we found that 971 patients within our case population had also been diagnosed with type 1 diabetes. This could indicate a subset of type 1 diabetes patients who were misdiagnosed as having type 2 diabetes as a result of developing the disease at an age later than is often associated with type 1 diabetes.[22] Incorrect diagnosis in these patients leads to ineffectual treatment strategies, resulting in poor glucose control and severe complications such as ketoacidosis and diabetic coma.[23] This could be why these patients were included in our complications case cohort.

The genetic variants found in this subcohort were found in as many as 53% of all cases with an additional type 1 diabetes diagnosis, showing that our hypothesis-free approach can effectively stratify patients by their respective phenotypes using only genotype data. Many of the genes specific to this cohort are either established type 1 diabetes genes, regulate beta-cell function, or are located in regions associated with the disease, such as the HLA-block.

Conclusion

Our analysis revealed several genes that were highly associated with the development of diabetes-related complications that would not have been found using standard GWAS analysis techniques.

We have shown that there are significant genetic differences between patients with type 2 diabetes, and these different subpopulations are at greater risk of developing complications, independent of dietary and lifestyle influences.

We have stratified these patients with complications further, identifying complication-specific subgroups with distinct profiles of genetic variants that are more highly enriched in certain complications than others. A deeper analysis of these genes' functions enabled us to form strong hypotheses as to why variants in these genes may predispose patients to developing specific complications.

We can use these high-resolution patient-stratification insights to identify patients most at risk, and form targeted educational and therapeutic interventions for them. This could result in greater prevention of disease progression and better patient outcomes, reducing the health and economic burden caused by diabetes complications.

Acknowledgements

We would like to acknowledge UK Biobank for providing us access to the data under Application ID 44288.

Notes and References

1. Diabetes UK. (2010). *Diabetes in the UK 2010: Key statistics on diabetes*. https://www.diabetes.org.uk/resources-s3/2017-11/diabetes_in_the_uk_2010.pdf

2. Hex, N., Bartlett, C., Wright, D., Taylor, M., & Varley, D. (2012). Estimating the current and future costs of Type 1 and Type 2 diabetes in the UK, including direct health costs and indirect societal and productivity costs. *Diabetic Medicine*, *29*(7), 855–862. https://doi.org/10.1111/j.1464-5491.2012.03698.x

3. Diabetes UK. (2014). *The Cost of Diabetes Report*. https://www.diabetes.org.uk/resources-s3/2017-11/diabetes%20uk%20cost%20of%20diabetes%20report.pdf

4. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P., & Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 203–209. https://doi.org/10.1038/s41586-018-0579-z

5. Taylor, K., Das, S., Pearson, M., Kozubek, J., Strivens, M., & Gardner, S. (2019). Systematic drug repurposing to enable precision medicine: A case study in breast cancer. *Digital Medicine*, *5*(4), 180–186. https://doi.org/10.4103/digm.digm_28_19

6. Gardner, S., Das, S., & Taylor, K. (2020). AI Enabled Precision Medicine: Patient Stratification, Drug Repurposing and Combination Therapies. In *Artificial Intelligence in Oncology Drug Discovery and Development*, Open Access book. https://doi.org/10.5772/intechopen.92594

7. Mellerup, E., Andreassen, O. A., Bennike, B., Dam, H., Djurovic, S., Jorgensen, M. B., Kessing, L. V., Koefoed, P., Melle, I., Mors, O., & Moeller, G. L. (2017). Combinations of genetic variants associated with bipolar disorder. *PloS One*, *12*(12), e0189739. https://doi.org/10.1371/journal.pone.0189739

8. Church, D. M., Schneider, V. A., Graves, T., Auger, K., Cunningham, F., Bouk, N., Chen, H. C., Agarwala, R., McLaren, W. M., Ritchie, G. R., Albracht, D., Kremitzki, M., Rock, S., Kotkiewicz, H., Kremitzki, C., Wollam, A., Trani, L., Fulton, L., Fulton, R., Matthews, L., … Hubbard, T. (2011). Modernizing reference genome assemblies. *PLoS Biology*, *9*(7), e1001091. https://doi.org/10.1371/journal.pbio.1001091

9. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

10. Tahergorabi, Z., & Khazaei, M. (2012). Imbalance of angiogenesis in diabetic complications: the mechanisms. *International Journal of Preventive Medicine*, *3*(12), 827–838. https://doi.org/10.4103/2008-7802.104853

11. Evans K. (2019). Diabetic ketoacidosis: update on management. *Clinical Medicine*, *19*(5), 396–398. https://doi.org/10.7861/clinmed.2019-0284

12. Karslioglu French, E., Donihi, A. C., & Korytkowski, M. T. (2019). Diabetic ketoacidosis and hyperosmolar hyperglycemic syndrome: review of acute decompensated diabetes in adult patients. *BMJ (Clinical research ed.)*, *365*, l1114. https://doi.org/10.1136/bmj.l1114

13. Hao, K., Di Narzo, A. F., Ho, L., Luo, W., Li, S., Chen, R., Li, T., Dubner, L., & Pasinetti, G. M. (2015). Shared genetic etiology underlying Alzheimer's disease and type 2 diabetes. *Molecular Aspects of Medicine*, *43–44*, 66–76. https://doi.org/10.1016/j.mam.2015.06.006

14. Hu, Z., Jiao, R., Wang, P., Zhu, Y., Zhao, J., De Jager, P., Bennett, D. A., Jin, L., & Xiong, M. (2020). Shared Causal Paths underlying Alzheimer's dementia and Type 2 Diabetes. *Scientific Reports*, *10*(1), 4107. https://doi.org/10.1038/s41598-020-60682-3

15. van Bussel, F., Backes, W. H., Hofman, P., Puts, N., Edden, R., van Boxtel, M., Schram, M. T., Stehouwer, C., Wildberger, J. E., & Jansen, J. (2016). Increased GABA concentrations in type 2 diabetes mellitus are related to lower cognitive functioning. *Medicine*, *95*(36), e4803. https://doi.org/10.1097/MD.0000000000004803

16. Hadi, H. A., & Suwaidi, J. A. (2007). Endothelial dysfunction in diabetes mellitus. *Vascular Health and Risk Management*, *3*(6), 853–876.

17. Taylor, K., Das, S., Pearson, M., Kozubek, J., Pawlowski, M., Jensen, C. E., Skowron, Z., Møller, G. L., Strivens, M., & Gardner, S. (2020). Analysis of Genetic Host Response Risk Factors in Severe COVID-19 Patients (preprint). medRxiv. doi: https://doi.org/10.1101/2020.06.17.20134015

18. Yue, Y., Meng, K., Pu, Y., & Zhang, X. (2017). Transforming growth factor beta (TGF-β) mediates cardiac fibrosis and induces diabetic cardiomyopathy. *Diabetes Research and Clinical Practice*, *133*, 124–130. https://doi.org/10.1016/j.diabres.2017.08.018

19. Wheeler, S. E., & Lee, N. Y. (2017). Emerging Roles of Transforming Growth Factor β Signaling in Diabetic Retinopathy. *Journal of Cellular Physiology*, *232*(3), 486–489. https://doi.org/10.1002/jcp.25506

20. Davidson, M. K., Russ, P. K., Glick, G. G., Hoffman, L. H., Chang, M. S., & Haselton, F. R. (2000). Reduced expression of the adherens junction protein cadherin-5 in a diabetic retina. *American Journal of Ophthalmology*, *129*(2), 267–269. https://doi.org/10.1016/s0002-9394(99)00323-2

21. Tao, C., & Zhang, X. (2016). Retinal Proteoglycans Act as Cellular Receptors for Basement Membrane Assembly to Control Astrocyte Migration and Angiogenesis. *Cell Reports*, *17*(7), 1832–1844. https://doi.org/10.1016/j.celrep.2016.10.035

22. Thomas, N. J., Jones, S. E., Weedon, M. N., Shields, B. M., Oram, R. A., & Hattersley, A. T. (2018). Frequency and phenotype of type 1 diabetes in the first six decades of life: a cross-sectional, genetically stratified survival analysis from UK Biobank. *The Lancet. Diabetes & Endocrinology*, *6*(2), 122–129. https://doi.org/10.1016/S2213-8587(17)30362-5

23. Bao, Y. K., Ma, J., Ganesan, V. C., & McGill, J. B. (2019). Mistaken Identity: Missed Diagnosis of Type 1 Diabetes in an Older Adult. *Medical Research Archives*, *7*(8), 1962.

Appendix

Cases with type 2 diabetes were identified using the ICD-10 codes in Table 3.

Table 3 Relevant ICD-10 codes

Code	Definition
E11.0	With coma
E11.1	With ketoacidosis
E11.2	With renal complications
E11.3	With ophthalmic complications
E11.4	With neurological complications
E11.5	With peripheral circulatory complications
E11.6	With other specified complications
E11.7	With multiple complications
E11.8	With unspecified complications

UK
Unit 8b Bankside
Long Hanborough
Oxfordshire
OX29 8LJ

USA
1 Broadway
Cambridge
MA 02142

DENMARK
Agern Allé 3
DK-2970, Hørsholm

POLAND
CIC, Ul. Chmielna 73
00-801, Warszawa

info@precisionlife.com

