Orchid Genetic Risk Score: Type 1 Diabetes

Orchid has developed advanced genetic risk scores (GRS) for a variety of diseases. Here we present our data on our GRS of Type 1 diabetes.

Written by the Orchid Team

1. Type 1 Diabetes

Type 1 diabetes mellitus is a chronic condition where the pancreas produces little to no insulin and excessive amounts of glucose accumulates in the blood [1]. The precise cause of type 1 diabetes is unknown, but the disease is usually triggered by an autoimmune response that causes the immune system to target insulin-producing beta cells [2]. Genetics impacts the likelihood of developing the disease; an analysis conducted on 2 million children drawn from the Taiwan National Health Insurance Database found the heritability of type 1 diabetes to be approximately 66% [3].

2. Clinical Impact and Prevalence

Approximately 1.6 million Americans are living with type 1 diabetes [4]. Type 1 diabetes is typically diagnosed between the ages of 4 and 14, though adulthood diagnoses do also occur [5]. The typical symptoms include increased thirst and urination, fatigue, weakness, irritability and other mood changes, and blurred vision [1]. The typical course of treatment for type 1 diabetes is insulin, administered regularly and prescribed by an endocrinologist [6].

3. Genetic risk score (GRS)

A genetic risk score quantifies the degree to which an individual's genetics increases their likelihood of developing a specific disease. The GRS for type 1 diabetes includes **64** variants and was developed based on the variants identified in a study that analyzed genomes of about **6481** individuals of European ancestry affected by type 1 diabetes. The study included **6481** cases (individuals with type 1 diabetes) and 9247 healthy controls.

Our type 1 diabetes GRS has some special characteristics relative to our usual GRS. For these diseases, one specific loci, HLA DQ, confers a disproportionate share of genetic risk [7]. That is, they follow an oligogenic model, not a classic polygenic model. For that reason, the resulting GRS is not normally distributed.

Number of variants in genetic risk score	**64**
Discovery GWAS(n=15,728)	Cases: 6481 Controls: 9247

Table 1: Discovery cohort statistics. Variants in GRS and sample number used in the type 1 diabetes GWAS.

4. Performant type 1 diabetes risk stratification

4.1. Comparing Baseline and elevated risk for type 1 diabetes

Individuals in the 99th percentile of genetic risk have a 2.79% prevalence of type 1 diabetes, compared to the baseline rate. Baseline rate is the prevalence of the disease in the entire reference population.

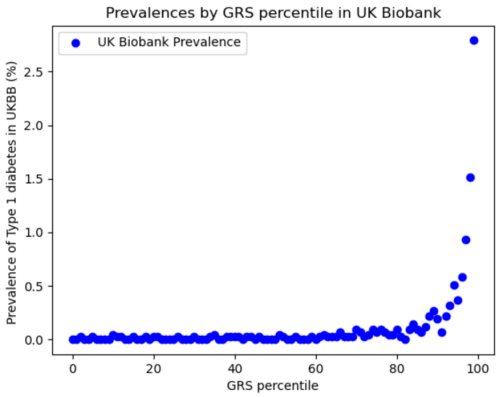

Figure 1: Risk gradient for type 1 diabetes. Each blue dot represents a percentile of Genetic Risk Score, with its percent prevalence in UK Biobank self-reported White British in the y-axis.

Validation in UK Biobank. In the UK Biobank, cases were identified using self-reported type 1 diabetes and relevant ICD-9/ICD-10 diagnosis. See the supplementary table for more details. In the validation, prevalence of type 1 diabetes increased with GRS. We restricted our analysis to self-reported British individuals whose genetic ancestry matched their self-identification. With our phenotype definition there were 421 cases of type 1 diabetes and 408,099 controls.

4.2. Comparing Baseline and elevated risk for type 1 diabetes

Individuals in the 99th percentile of genetic risk develop type 1 diabetes at 27 times the baseline rate. Baseline rate is the prevalence of the disease in the entire reference population.

Elevated Genetic Risk Definition	Prevalence
Baseline Prevalence	0.10%
Top 5% of distribution	1.24%
Top 3% of distribution	1.75%
Top 1% of distribution	2.79%
Top 0.5% of distribution	3.72%

Table 2: Disease prevalence in elevated genetic risk subgroups for White British individuals.

5. Comparison to Published Benchmarks

Orchid's model achieves greater stratification performance with an AUC of 0.904 compared to the benchmark of 0.92.

We compared the performance of our model as validated on the UK Biobank with the performance of the best model in Sharp et al. The benchmarks in this table were generated on different datasets, so they are not precisely comparable.

Model	Orchid	Reference[1]
AUC of model with GRS, age, and PCs[2]	0.904	0.92[3]

Table 3: Accuracy metric comparison. Our model compared to reference.

References

[1] Mayo Clinic. Type 1 diabetes, 2022. URL `https://www.mayoclinic.org/diseases-conditions/type-1-diabetes/symptoms-causes/syc-20353011`.

[2] Mikael Knip, Heli Siljander, Jorma Ilonen, Olli Simell, and Riitta Veijola. Role of humoral beta-cell autoimmunity in type 1 diabetes. *Pediatric diabetes*, 17:17–24, 2016.

[3] Chang-Fu Kuo, I-Jun Chou, Matthew J Grainge, Shue-Fen Luo, Lai-Chu See, Kuang-Hui Yu, Weiya Zhang, Michael Doherty, and Ana M Valdes. Familial aggregation and heritability of type 1 diabetes mellitus and coaggregation of chronic diseases in affected families. *Clinical Epidemiology*, pages 1447–1455, 2018.

[4] American Diabetes Association. Statistics about diabetes. [cited 7 Jan 2022].

[5] S Watts. What is type 1 diabetes. *EndocrineWeb [Internet].[cited 7 Jan 2022]. Available: https://www. endocrineweb. com/conditions/type-1-diabetes/type-1-diabetes*, 2022.

[6] Healthline. A complete list of diabetes medications.

[7] Jana Sticht, Miguel Álvaro-Benito, and Stefan Konigorski. Type 1 diabetes and the hla region: Genetic association besides classical hla class ii genes. *Frontiers in Genetics*, 12:683946, 2021.

6. Appendix

Disease case identification and number of cases in UK Biobank

Phenotype	ICD-10 Codes	Self-Report Codes	Cases in UK Biobank (white British)
Prostate cancer	C61, D075	1044	13,806
Type 2 diabetes	E11.1-9	1223	30,507
Coronary artery disease	I2104,1219,I220, I221,I228,I232 I233,I235,I236 I238,I249,I252	1075	22,451
Breast cancer	C5.0-9, D05.0, D059	1002	18,588
Inflammatory bowel disease	K51	1461,1462 ,1463	5,959
Atrial fibrillation	148.0-4,148.9	1471,1483	22,472
Schizophr enia	F20.0-9, F21, F23.0-3, F23.8	1289	1,376
Alzheimer's disease	F00.0-2, F00.9, G30.0-1,8-9.	1263	2,547
Celiac disease	K900	1456	3,253
Bipolar disease	F31	1291	1,855
Type 1 diabetes	*	*	421

Table 4: Supplementary Table: How each disease case is defined in evaluating genetic risk scores in the UK Biobank

*Type 1 diabetes was defined as a combination the following inclusion and exclusion criteria:

- Self-diagnosed diabetes (any type)
- No self-diagnosed Type 2 diabetes
- Age of diabetes onset between 0 and 20 years
- Started insulin within one year of diagnosis of diabetes