Orchid Genetic Risk Score: Inflammatory Bowel Disease

Orchid has developed advanced genetic risk scores (GRS) for a variety of diseases. Here we present our data on our GRS of inflammatory bowel disease.

Written by the Orchid Team

1. Inflammatory Bowel Disease

Inflammatory bowel disease (IBD) is a term for two conditions (Crohn's disease and ulcerative colitis) that are characterized by chronic inflammation of the gastrointestinal tract [1]. Both forms of IBD are caused by systemic immune dysfunction, where the body's immune system attacks the GI tract rather than foreign viruses or bacteria. There is a substantial genetic component to IBD; having a first degree relative with the disease is the highest risk factor, and those with two affected parents have a 30% risk of developing the disease, compared to a population prevalence of 1.3% [2]. The heritability of IBD is approximately 70%, based on a literature review of several twin studies drawn from a number of European twin registries, including those of Denmark, Sweden, and Britain [2].

2. Clinical Impact and Prevalence

IBD affects over 3 million Americans, or roughly 1.3% of the US population [3]. It is usually diagnosed between 15 and 35, with 29 being the average age of diagnosis [4]. Typical symptoms of IBD include persistent diarrhea, abdominal pain, rectal bleeding, weight loss, and fatigue [1], [5]. While IBD is an immune disorder with no known cure, several medications may be used to treat the symptoms of IBD: aminosalicylates, corticosteroids (such as prednisone), and immunomodulators can be prescribed by a gastroenterologist [6]. People with IBD may have higher rates of colorectal cancer, though current treatments may reduce that risk somewhat [7].

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 IBD includes 1,097,724 variants and was developed based on the variants identified in a study that analyzed genomes of 34,652 individuals of European ancestry. The study included 12,882 cases (individuals with IBD) and 21,770 healthy controls [8]. The risk score model adjusted weights for linkage disequilibrium using the PRScs software.

Number of variants in genetic risk score	**1,097,724**
Discovery GWAS(n=34,652)	Cases:12,882 Controls: 21,770

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

4. Performant inflammatory bowel disease risk stratification

4.1. Validated using a large cohort of adults with known inflammatory bowel disease status

Individuals in the 99th percentile of genetic risk have a 5.8 percent absolute risk of inflammatory bowel disease, compared to the average of 1.46 percent.

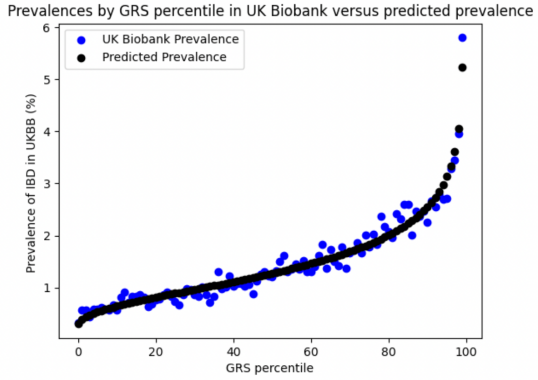

Figure 1: Risk gradient for inflammatory bowel disease. Each blue dot represents a percentile of Genetic Risk Score, with its percent prevalence in UK Biobank self-reported White British in the y-axis. The black line represents the predicted prevalence from a logistic regression derived from the data.

Validation in UK Biobank. In the UK Biobank, cases were identified using self-reported IBD, relevant ICD-10 diagnosis and death codes. See the phenotype supplement for full details. In the validation set, prevalence of the disease increased with GRS.

People at the tail end of GRS distribution were at an elevated risk (had higher odds) for developing the disease in comparison to the baseline rate. For each GRS threshold, we also computed the odds ratio. Baseline rate is the prevalence of the disease in the entire reference population.

4.2. Comparing Baseline and elevated risk for inflammatory bowel disease

Adults in the 99th percentile of genetic risk have a 5.80% risk of developing inflammatory bowel disease. The odds ratio for adults in the 99th percentile of genetic risk was 4.16.

Elevated Genetic Risk Definition	Prevalence	odds ratio
Baseline Prevalence	1.46%	1
Top 5% of distribution	3.84%	2.7
Top 3% of distribution	4.41%	3.11
Top 1% of distribution	5.80%	4.16
Top 0.5% of distribution	6.56%	4.74

Table 2: Disease prevalence and odd ratios in elevated genetic risk subgroups for white British individuals.

5. Comparison to Published Benchmarks

Orchid's model achieves comparable stratification performance with an AUC of 0.654 compared to the benchmark at 0.618.

We compared the performance of our model as validated on the UK Biobank with the performance of the best model in Khera et al. To make a comparison of models, we restricted our validation sample to those in Phase II of the UK Biobank release, as in Khera et. al. In the first column, we give the results for our predictor with the phenotype as described above. In the second, we report the metrics for the best-performing predictor in Khera et. al using the same phenotype as ours.

Model	Orchid	Reference[1]
AUC of model with GRS, age, and PCs[2]	0.654	0.618

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

References

[1] Orchid Team. What is inflammatory bowel disease (ibd)? Available at `https://www.cdc.gov/ibd/what-is-IBD.htm`, 2020. [cited 4 Jan 2022].

[2] Hannah Gordon, Frederik Trier Moller, Vibeke Andersen, and Marcus Harbord. Heritability in inflammatory bowel disease: from the first twin study to genome-wide association studies. *Inflammatory bowel diseases*, 21(6):1428–1434, 2015.

[3] Orchid Team. Ibd data and statistics. Available at `https://www.cdc.gov/ibd/data-statistics.htm`, note = [cited 4 Jan 2022], year = 2021.

[4] Raina Shivashankar, William J Tremaine, W Scott Harmsen, and Edward V Loftus Jr. Incidence and prevalence of crohn's disease and ulcerative colitis in olmsted county, minnesota from 1970 through 2010. *Clinical Gastroenterology and Hepatology*, 15(6):857–863, 2017.

[5] Orchid Team. Crohn's & colitis foundation of america. Available at `https://www.crohnscolitisfoundation.org/sites/default/files/2019-02/Updated%20IBD%20Factbook.pdf`, 2022. [cited 4 Jan 2022].

[6] Katsuyoshi Matsuoka, Taku Kobayashi, Fumiaki Ueno, Toshiyuki Matsui, Fumihito Hirai, Nagamu Inoue, Jun Kato, Kenji Kobayashi, Kiyonori Kobayashi, Kazutaka Koganei, et al. Evidence-based clinical practice guidelines for inflammatory bowel disease. *Journal of gastroenterology*, 53(3):305–353, 2018.

[7] Ryan W Stidham and Peter DR Higgins. Colorectal cancer in inflammatory bowel disease. *Clinics in colon and rectal surgery*, 31(03):168–178, 2018.

[8] Amit V Khera, Mark Chaffin, Krishna G Aragam, Mary E Haas, Carolina Roselli, Seung Hoan Choi, Pradeep Natarajan, Eric S Lander, Steven A Lubitz, Patrick T Ellinor, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature genetics*, 50(9):1219–1224, 2018.

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