In this talk, I will describe our efforts at Google Health on automated phenotyping of biobank-scale cohorts. Genome-wide association studies (GWAS) require accurate cohort phenotyping, but expert labeling can be costly, time-intensive, and variable. Here we develop a machine learning (ML) model to predict glaucomatous optic nerve head features from color fundus photographs. We used the model to predict vertical cup-to-disc ratio (VCDR), a diagnostic parameter and cardinal endophenotype for glaucoma, in 65,680 Europeans in the UK Biobank (UKB). A GWAS of ML-based VCDR identified 299 independent genome-wide significant (GWS; P≤5e−8) hits in 156 loci. The ML-based GWAS replicated 62 of 65 GWS loci from a recent VCDR GWAS in the UKB for which two ophthalmologists manually labeled images for 67,040 Europeans. The ML-based GWAS also identified 93 novel loci, significantly expanding our understanding of the genetic etiologies of glaucoma and VCDR. Pathway analyses support the biological significance of the novel hits to VCDR, with select loci near genes involved in neuronal and synaptic biology or known to cause severe Mendelian ophthalmic disease. Finally, the ML-based GWAS results significantly improve polygenic prediction of VCDR and primary open-angle glaucoma in the independent EPIC-Norfolk cohort.